*Ethiop. J. Biol. Sci.* **18(2): 153–178, 2019**
 **ISSN: 1819-8678**

***RESEARCH ARTICLE***

# NEW INSIGHT INTO THE PHYLOGENY AND BIOGEOGRAPHY OF THE CYPRINID FISHES *LABEO* (CYPRINIDAE; CYPRINIFORMES) IN AFRICA WITH EVIDENCE FOR CRYPTIC DIVERSITY IN ETHIOPIA

Kebede A. Beshera[1,*] and Phillip M. Harris[2]

**ABSTRACT:** To investigate the molecular phylogeny and biogeography of African *Labeo* and infer and date cladogenetic events that led to its diversification, a time-calibrated phylogeny was constructed based on analysis of mitochondrial cytochrome *b* (cyt *b*) gene sequences obtained from 18 African and 14 Asian *Labeo* species. Phylogenies were constructed under the GTR nucleotide substitution model employing Bayesian inference (BI) and Maximum Likelihood (ML) methods. A time calibrated phylogeny was generated using the GTR model with the Birth and Death speciation tree prior. Both Bayesian and ML phylogenies rendered African *Labeo* non-monophyletic with South African species grouped along with some Asian species suggesting multiple independent in-to-Africa dispersals of *Labeo*. Two exclusively African clades were recovered within *Labeo*; the first includes samples from lower and middle Congo and the eastern segment of Nilo-Sudanic river drainages, while the second comprised samples from the Congo and Nilo-Sudanic drainages. The separation of the Eastern Nilo-Sudanic + lower and middle Congo (clade 1) and Eastern and western Nilo-Sudanic + Congo clades (clade 2) may have occurred around 13.8 MYA with subsequent diversification in these clades leading to extant species diversity of *Labeo* in Africa. Overall, this study provides new insights into the taxonomy and biogeography of African *Labeo* with evidence for cryptic diversity within Ethiopia and incongruence between traditional taxonomy and molecular phylogeny. In addition, the study found a close correlation between the biogeographic history of the region and genetic diversity of *Labeo* in Africa.

**Key words/phrases**: Africa, Biogeography, Cryptic diversity, *Labeo*, Phylogeny.

[1] Louisiana State University Eunice, 2048 Johnson Highway, Eunice, LA 70535. E-mail: kbeshera@lsue.edu
[2] Biodiversity and Systematics, Department of Biological Sciences, The University of Alabama, Tuscaloosa, AL 35487-0345, U.S.A. E-mail: pharris@ua.edu
* Author to whom all correspondence should be addressed.

## INTRODUCTION

The subtribe Labeoina (Cyprinidae) defined by Yang *et al*. (2012) and recently raised to a tribe level as Labeonini by Tan and Armbruster (2018), includes the Afro-Asian genus *Labeo* Cuvier 1816 and its Asian allies. With about 105 valid species, *Labeo* represents one of the largest and most diverse genera among the Cyprinidae (Yang *et al*., 2010). The type species of *Labeo* is an African species, *L. niloticus* Forskål 1775.

African *Labeo* along with fishes of the tribe Garrini (as defined by Tan and Armbruster, 2018) originated in Asia (Greenwood, 1972; Reid, 1985) and colonized Africa through multiple dispersal events (Tang *et al*., 2009; Yang and Mayden, 2010; Yang *et al*., 2012; Zheng *et al*., 2012). Fossil evidence indicates that *Labeo* existed in Africa as far back as the Middle Miocene (Stewart, 2001; 2009), suggesting that the diversification of these fishes in Africa probably occurred 11.6–16.0 million years ago. According to Stewart and Murray (2017), the late Miocene Bab el-Mandeb Strait land bridge has likely allowed the cyprinids *Labeo*, *Labeobarbus* and *Barbus* access to the African continent from Asia and Arabia. These immigrants are now generally widespread throughout most of Africa including the Nile, Congo, and Zambezi basins as well as eastern and western African lakes and rivers and have had a large impact on taxonomic diversity in African freshwater fish faunas (Stewart and Murray, 2017). With about 80 species, *Labeo* is the third most diverse cyprinid genus in the continent, after *Enteromius* Cope 1867 and *Labeobarbus* Ruppel 1835 (Reid, 1985; Skelton *et al*., 1991).

Historically, the morphology of the lips and associated structures of *Labeo* have been used as a basis for intrageneric taxonomy. African *Labeo* were reviewed based on morphology by Reid (1985), Tshibwabwa and Teugels (1995), and Tshibwabwa (1997), while a recent taxonomic examination (Van Steenberge *et al*., 2016) led to corrections of records and type designations for seven Congolese species. In his revision of African *Labeo*, Reid (1985) subdivided the genus into six species-groups, namely *L. coubie*, *L. forskalii*, *L. gregorii*, *L. niloticus*, *L. macrostoma*, and *L. umbratus* species-groups. However, he doubted that these subdivisions may not represent monophyletic groups underscoring the need to assess the monophyly of these groups using molecular data. The species he included in each of these groups are listed below:

1. *L. coubie* group (LCG): *L. coubie* Ruppel 1832, *L. barbatus* Boulenger 1898, *L. degeni* Boulenger 1920, *L. longipinnis* Boulenger 1898

2. *L. forskalii* group (LFG): *L. forskalli* Ruppel 1835, *L. ansorgii* Boulenger 1907, *L. cylindricus* Peters 1852, *L. darlingii* Boulenger 1911, *L. trigliceps* Pellegrin 1926, *L. victorianus* Boulenger 1901, *L. fuelleborni* Hilgendorf & Pappenheim 1903, *L. worthringtoni* Fowler 1958, *L. molybdinus* duplessi 1963, *L. lunatus* Jubb 1963, *L. parvus* Boulenger 1902, *L. ogunensis* Boulenger 1910, *L. obscurus* Lin 1981, *L. djourae* Blache & Miton 1960, *L. alluaudi* Pellegrin 1933, *L. lukulae* Boulenger 1902, *L. nunensis* Pellegrin 1929, *L. nasus* Boulenger 1899, and *L. sorex* Nichols & Griscom 1917
3. *L. gregorii* group (LGG): *L. bottega* Vinciguerra 1897 and *L. gregorii* Günther 1894
4. *L. macrostoma* group (LMG): *L. cyclorhynchus* Boulenger 1899, *L. macrostoma* Boulenger 1898, *L. greenii* Boulenger 1902, and *L. batesii* Boulenger 1911
5. *L. niloticus* group (LNG): *L. niloticus* Linnaeus 1758 (junior synonym of *L. vulgaris*), *L. horie* Heckel 1847, *L. senegalensis* Valenciennes 1842, *L. weeksii* Boulenger 1909, *L. mesops* Günther 1868, *L. lineatus* Boulenger 1898, *L. rosae* Steindachner 1894, *L. altivelis* Peters 1852, and *L. ruddi* Boulenger 1907.
6. *L. umbratus* group (LUG): *L. umbratus* Smith 1841, *L. capensis* Smith 1841, *L. seeberi* Gilchrist & Thompson 1911, and *L. rubromaculatus* Gilchrist & Thompson 1913

In addition, morphological characters that have been widely used for species identification and phylogenetic reconstruction are highly variable confounding our ability to make accurate taxonomic and phylogenetic inference. In the past few decades, molecular markers such as mitochondrial DNA sequences (e.g., cyt *b* and cytochrome *c* oxidase gene sequences) have become increasingly popular as tools for phylogenetic and taxonomic studies. Researchers working on African *Labeo* only recently began to explore the potential use of molecular analysis to gain insight into the taxonomy of, and interrelationships within, *Labeo*. These include Lowenstein *et al*. (2011, for *Labeo* species from the Congo basin), Ramoejane (2016, for southern African *Labeo* species), and Yang *et al*. (2012, for African and Asian *Labeo* species). However, in the first two studies East African species were either underrepresented or not analyzed. Yang *et al*. (2012), on the other hand, analyzed more comprehensive data on *Labeo* but was focused mainly on establishing major relationships among

member genera and species within the framework of the tribe Labeonini. Consequently, other aspects of the phylogeny remain poorly resolved, notably monophyly of individual species and intraspecific genetic variation and phylogeography. These represent significant gaps in understanding the taxonomy, phylogeny, and biogeography of *Labeo* in Africa.

In Ethiopia, fishes of the genus *Labeo* are widely distributed in, and inhabit, all the major river basins including the Abay (Blue Nile), Awash, Baro-Akobo, Omo, Tekeze, and Genale-Wabishebele as well as the Rift Valley lakes Abaya and Chamo (Golubtsov and Mina, 2003; Tadele Awoke, 2015; Stewart and Murray, 2017). About 11 species of *Labeo* have been reported from the freshwaters of Ethiopia in early literature (Golubtsov *et al*., 2002; Golubtsov and Mina, 2003; Lemma Abera *et al*., 2014; Tadele Awoke, 2015). Reported species include *Labeo bottegi* Vinciguerra 1897, *L. boulengeri* Vinciguerra 1912, *L. brunellii* Parenzan 1939, *L. coubie* Rüppel 1832, *L. cylindricus* Peters 1852, *L. forskalii* Rüppel 1835, *L. horie* Heckel 1847, *L. niloticus* Linnaeus 1758, *L. pellegrini* Zollezi 1939, *L. parvus* Boulenger 1902, and *L. vulgaris* Heckel 1847. However, this is not an accurate estimate due to insufficient explorations of the taxon in the freshwaters of Ethiopia and the use of ambiguous or synonymous taxon names. *Labeo brunellii* for example, has long been recognized as a junior synonym of *L. horie* (Golubtsov *et al*., 2002). Besides being the second most abundant fish group (next to *Labeobarbus*) in Ethiopian rivers (Zeleke Berie, 2007; Simagegnew Melaku *et al*., 2017) *Labeo* contributes substantially to the inland fish catch in Ethiopia along with tilapia, catfish, Nile perch and *Labebarbus* (Gashaw Tesfaye and Wolff, 2014). Despite its ecological and economic significance, limited information is available on *Labeo* in Ethiopia. In this study we analyzed complete and partial cytochrome *b* (cyt *b*) gene sequences to 1) assess phylogeny and biogeography of *Labeo* in Africa; 2) investigate the timescale of lineage divergence within *Labeo* in Africa; 3) examine intraspecific genetic variation within, some Ethiopian species; and 4) test the monophyly of nominal species and Reid's (1985) species-groups within *Labeo*.

## MATERIALS AND METHODS

### Sampling

Because of logistic and financial limitations, we were not able to acquire samples representative of all African species. Taxon sampling was conducted to cover major Ichthyological provinces of Africa (Nilo-Sudanic, Congo, Zambezi, Lower-Guinea etc.) and include species representative of

most of the species groups proposed by Reid (1985). In this study 60 cytochrome *b* (cyt *b*) sequences (14 partial and 46 complete sequences) representing 18 *Labeo* species from Africa (*L. altivelis*, *L. batesii*, *L. cf. barbatus*, *L. capensis*, *L. coubi*, *L. cyclorhynchus*, *L. cylindricus*, *L. forskalii*, *L. horie*, *L. lineatus*, *L. longipinnis*, *L. nasus*, *L. parvus*, *L. senegalensis*, *L. sorex*, *L. umbratus*, *L. vulgaris*, and *L. weeksii*) and one specimen of uncertain taxonomic identity (*Labeo* sp.) were analyzed. Of these a set of 33 sequences represent five species (*Labeo cylindricus*, *Labeo forskalii*, *Labeo horie*, *Labeo parvus*, and *Labeo vulgaris*) derived from seven river drainages across Ethiopia. Twenty-two original sequences were generated for this study while the remaining sequences were retrieved from GenBank. Voucher specimens of samples collected for this study are stored at the University of Alabama Ichthyological Collections (UAIC). In addition, 21 species of the subtribe Labeonini (including 15 species of *Labeo*) from Asia were included in our analyses. Based on previous phylogenetic information (Yang *et al*., 2012), four cyprinid species were used as outgroup taxa. The list of samples and taxa examined in this study along with GenBank accession numbers are provided in Table 1.

Table 1. Taxa and samples examined in this study with locality information and GenBank accession numbers (cyt *b* gene).

| Taxon | Source | River Drainage | Accession # |
|---|---|---|---|
| **Ingroup taxa** | | | |
| *L. cylindricus* | | | |
| Haplotype 1 | Genale River | Genale-Dawa | MN073793† |
| Haplotype 2 | Genale River | Genale-Dawa | MN073794† |
| Haplotype 3 | Genale River | Genale-Dawa | MN073795† |
| Haplotype 4 | Genale River | Genale-Dawa | MN073796† |
| Haplotype 5 | Genale River | Genale-Dawa | MN073797† |
| Haplotype 6 (2) | Not available | Not available | AP011206, NC031536 |
| | Not available | Not available | |
| *L. horie* | Alwero River | Baro-AKobo | JX074288 |
| *L. forskali* | | | |
| Haplotype 1 | Alwero River, Ethiopia | Baro-Akobo | JX074287 |
| Haplotype 2 | Ghibe River, Ethiopia | Omo-Turkana | MN073791† |
| Haplotype 3 | Gojeb River, Ethiopia | Omo-Turkana | MN073786† |
| Haplotype 4 (2) | Dedessa River, Ethiopia | Nile | MN073787† |
| | Sanja River, Ethiopa | Tekeze | FJ196833 |
| Haplotype 5 | Dedessa River, Ethiopia | Nile | MN073788† |
| Haplotype 6 | Dedessa River, Ethiopia | Nile | MN073789† |
| Haplotype 7 | Abay River, Ethiopia | Nile | MN073790† |
| Haplotype 8 (2) | Sanja River, Ethiopa | Tekeze | FJ196831, FJ196832 |
| Haplotype 9 | Ghibe River, Ethiopia | Omo-Turkana | MN073786† |
| Haplotype 10 | Ghibe River, Ethiopia | Omo-Turkana | MN073792† |
| *L. parvus* | | | |
| Haplotype 1 | Baro River | Baro-Akobo | JX074286 |
| Haplotype 2 (2) | Baro River | Baro-Akobo | JX074285, MN073801† |
| Haplotype 3 | Benin, Niger bei Malauville | Not available | JX074292 |

| Taxon | Source | River Drainage | Accession # |
|---|---|---|---|
| Haplotype 4 | Not available | Not available | AP013339 |
| Haplotype 5 | Baro River, Ethiopia | Omo-Turkana | MN073798† |
| Haplotype 6 | Baro River, Ethiopia | Omo-Turkana | MN073799† |
| Haplotype 7 | Baro River, Ethiopia | Omo-Turkana | MN073800† |
| *L. vulgaris* | | | |
| Haplotype 1 (4) | Gojeb River | Omo-Turkana | JX074297, MN073805† MN073803† |
| | Ghibe River | Omo-Turkana | JX074298 |
| Haplotype 2 | Baro River | Baro-Akobo | JX074296 |
| Haplotype 3 | Baro River, Ethiopia | Baro-Akobo | MN073802† |
| Haplotype 4 | Baro River, Ethiopia | Baro-Akobo | MN073804† |
| *L. altivelis*1 | Congo: Odzala | | JX074294 |
| *L. altivelis*2 | Not available | | AP013322 |
| *L. lineatus*1 | Congo: Odzala | | AP012154 |
| *L. lineatus*2 | Not available | | NC022956 |
| *L. longipinnis* | Not available | | JX074290 |
| *L. coubie*1 | Benin: Pendjari National Park | | JX074261 |
| *L. coubie*2 | not available | | AP012149 |
| *L. cyclorhynchus*1 | Not available | | AP011359 |
| *L. cyclorhynchus*2 | Not available | | NC022949 |
| *L. senegalensis*1 | Benin: Oueme R. and Iguidi R. | | AB238968 |
| *L. senegalensis*2 | Not available | | NC008657 |
| *L. sorex* | Not available | | AY791415 |
| *L. weeksii* | Not available | | JX097079 |
| *L. cf. barbatus*1 | Not available | | AP013337 |
| *L. cf. barbatus*2 | Not available | | JX074278 |
| *Labeo* sp. | Not available | | AP013323 |
| *L. nasus*1 | Not available | | AP013333 |
| *L. nasus*2 | Not available | | NC029449 |
| *L. umbratus*1 | Val Dam, South Africa | Val-Orange | MF469476 |
| *L. umbratus*2 | Gariep Dam, South Africa | Upper Orange R. | MF469596 |
| *L. capensis*1 | Gariep Dam, South Africa | Upper Orange R. | MF469582 |
| *L. capensis*2 | Vaal River, South Africa | Orange-Val | MF469456 |
| *L. capensis*3 | Hardap Dam, South Africa | Orange River | MF469569 |
| *L. batesii*1 | Loa Loa, Gabon | | NC008656 |
| *L. batesii*1 | Loa Loa, Gabon | | AB238967 |
| *Labeo dyocheilus* | Not available | | JX074262 |
| *Labeo pierrei* | Cambodia, Kampong | | AP011200 |
| *Labeo yunnanensis* | Not available | | JX074282 |
| *Labeo calbasu* | Aquarium | | AP012143 |
| *Labeo chrysophekadion* | Cambodia: Market, Ta Khmau, Kandal | | AP011199 |
| *Gibelion catla*1 | Market | | JX083157 |
| *Gibelion catla*2 | Aquarium | | JX074273 |
| *Labeo fimbriatus* | India | | GQ853089 |
| *L. gonius* | GenBank | | HQ645093 |
| *Labeo stolizkae* | China: Ruili, Yunnan | | GU086574 |
| *Labeo rohita* | Cambodia: Landing port, Kampong | | AP011201 |
| *Labe dussumieri* | India | | JX074250 |
| *Incisilabeo behri* | Aquarium | | JX074248 |
| *Labeo bata* | N/A | | JX074260 |
| *Labeo bogutt* | GenBank | | HQ645089 |

| Taxon | Source | River Drainage | Accession # |
|---|---|---|---|
| *Labeo* sp. *salween* | Myanmar: Moei R. (Mae Nam Moei), near Sap Moei | | JX074230 |
| *L. rajasthanicus* | India | | DQ520921 |
| *Cirrhinus cirrhosus*1 | Vietnam: Lang Son Market | | JX074223 |
| *Cirrhinus cirrhosus*2 | N/A | | JX074249 |
| *Cirrhinus microlepis* | Cambodia: Market, Ta Khmau, Kandal | | HM536825 |
| *Bangana ariza*1 | Nepal: Upsteram of Koshi Barrage | | JX074227 |
| *Bangana ariza*2 | India | | JX074236 |
| *Bangana dero* | India | | JX074235 |
| **Outgroup taxa** | | | |
| *Osteocheilus nashii* | Aquarium | | AP011330 |
| *Catlocarpio siamensis* | Cambodia: Landing port, Kampong Chhnang | | HM536812 |
| *Gobio gobio* | GenBank | | AB239596 |
| *Tor sinensis* | China: Mengna, Yunnan | | HM536802 |

† Sequences generated for the present study.

## DNA isolation, PCR amplification, and sequencing

Genomic DNA was extracted from 20–30 mg ethanol preserved muscle tissue using QIAGEN DNeasy kit (QIAGEN Inc.) according to manufacturer's instructions. 1141 bp protein coding region of the mitochondrial genome representing cyt *b* gene was amplified using the primer pairs L15267 (5'-AATGACTTGAAGAAC CACCGT-3'), H16461 (5'-CTTCGGATTACAAGACC-3'). PCR amplification protocol and primer sequences were taken from Briolay *et al*. (1998). Prior to cycle sequencing PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN Inc.). Cycle sequencing reactions were performed using the primers H15891 (5'-GTTTGATCCCGTTTCGTGTA-3') and H16461 (Briolay *et al*., 1998). The reverse complements of sequences generated using the primer H15891 (internal primer) were used as forward sequences. Sequences were generated via dye terminator reactions and read on an ABI 3100 prism sequencer (www.lifetechnologies.com) and manually aligned using BIOEDIT software version 5.0.9 (Hall, 1999). DNA sequences were subsequently deposited in GenBank and assigned accession numbers (Table 1).

## Phylogenetic analysis and divergence date estimation

The best performing Maximum Likelihood nucleotide substitution models for 81 nucleotide sequences of our cyt *b* was determined in MEGA X (Kumar *et al*., 2018). The model with the lowest BIC (Bayesian Information Criterion) score is considered to describe the substitution pattern in the dataset the best.

Phylogenetic analyses were performed using Beast version 1.10.4 (Suchard *et al*., 2018) and Maximum Likelihood (ML) method as implemented in MEGA version X (Kumar *et al*., 2018). For all Bayesian analyses using Beast an XML input file was generated using Bayesian Evolutionary Analysis Utility (BEAUTI) version 1.10.4 software (Drummond *et al*., 2002–2018). Phylogenetic analysis employing Beast was run under the GTR nucleotide substitution, uncorrelated lognormal relaxed molecular clock, and Gamma + Invariant site rate heterogeneity models employing the Birth and Death speciation tree prior. For this analysis, three independent Markov Chain Monte Carlo (MCMC) simulations were run for 100 million generations, sampling trees every 5000 generations. The Effective Sample Size (ESS) values for each parameter and the stationarity of the likelihood values were evaluated in Tracer v.1.6 (Rambaut *et al*., 2018). Outputs of the three independent simulations were then combined using LogCombiner v.2.4.2 with a 10% burn-in. These outputs were used to reconstruct a maximum credibility tree in TreeAnnotator v.2.4.2. ML analysis was run applying the GTR nucleotide-substitution model with Gamma + Invariant sites along with rapid bootstrapping (1000 bootstrapping replicates) and ML heuristic Nearest Neighbor Interchange (NNI) tree inference option. We used Bayesian Clade Credibility (Posterior Probability values) for Bayesian analysis and maximum likelihood bootstrap values for ML analysis to evaluate support for resolved clades. The tree topologies constructed by ML and BI were compared visually.

Divergence times for principal clades in the phylogeny were inferred in Beast version 1.10.4 (Suchard *et al*., 2018) employing the GTR substitution, uncorrelated lognormal relaxed molecular clock, and the Inv-Gamma rate heterogeneity models and Birth and Death speciation tree priors. To calibrate our phylogeny, a generalized teleost cyt *b* substitution rate of 0.76–2.2% per million years (Berendzen *et al*., 2008) was used as a uniform prior. The earliest fossil record of *Labeo* in Africa is from the Mid-Miocene (11.6–16 MYA; Stewart, 2001; 2009). Therefore, the node comprising clades 1 and 2 was time scaled to 13.8 MYA in Figtree v.1.4.4 (Rambaut, 2006–2018) to provide additional calibration point for divergence time estimation. Three independent Bayesian Markov Chain Monte Carlo (MCMC) simulations each with 100 million generations were run, sampling trees every 5000 generations. Results from the three simulations were combined using Logcombiner version 1.10.4 (Rambaut and Drummond, 2002–2018) discarding the first 10% of the samples as burn-in. In all analyses involving Beast, convergence of chains was checked using Tracer

version 1.7 (Rambaut *et al*., 2018). Tree Annotator version 1.10.4 (Rambaut and Drummond, 2002–2018) and Figtree version 1.4.3 (Rambaut, 2006–2016) were used to summarize the posterior tree distribution and visualize annotated maximum clade credibility (MCMC) tree, respectively.

### Genetic variation and population differentiation

The number of unique haplotypes (H), haplotype diversity (h), and nucleotide diversity (Pi) were calculated for our mtDNA sequence data using DnaSP version 6.12.03 (Rozas *et al*., 2017). *L. forskalii* and *L. parvus* were also further examined using TCS analysis implemented in PopART, Population Analysis with reticulate Trees version 1.7 (Leigh and Bryant, 2015) to assess genealogical relationships among unique mtDNA haplotypes by grouping haplotype sequences into a minimum connecting parsimony network. The maximum numbers of mutational steps that make parsimonious connections between haplotype sequences were calculated with 95% confidence. Mean genetic distances between species and sub-group pairs were estimated by maximum composite likelihood model using the bootstrap variance estimation method as implemented in MEGA X program (Kumar *et al*., 2018).

## RESULTS

### Phylogenetic relationships

Sequences for the partial and complete cyt *b* gene ranging in length from 730 to 1141 bp were aligned for a total 60 individuals representing 18 *Labeo* species from Africa. The data set contained 278 invariable (monomorphic) and 253 variable (polymorphic) sites of which 219 were parsimony informative. Transition/Transversion bias (R) estimated under the Tamura and Nei (1993) model was 4.44. The nucleotide frequencies were A = 30.4%, T = 27.6%, C = 28.3%, and G = 13.7%. Of the 24 different substitution models evaluated in MEGA X, the nucleotide substitution model that best fit our dataset was GTR+G+I with Maximum Likelihood value (lnL) and BIC score of -10841.7 and 23608.4, respectively.

The resulting Bayesian (Fig. 1) and ML phylogenies are largely congruent except for variation in support values: some nodes received poor ML bootstrap support (all nodes whose support values are not shown in Fig. 1 have BS<70% and PP<75%). Both phylogenies do not support the monophyly of African *Labeo* because two South African species (*L. capensis* and *L. umbratus*) were allied to some Asian *Labeo* species (clade III, Afro-Asian clade). They, however, recover a clade inclusive of the

remaining African species (clade I + clade II) with moderate support, Posterior Probability (PP) = 85%. Overall, African representatives of *Labeo* were resolved in three clades: clade I corresponds to the *L. forskalii* species group (Bootstrap = 82%, PP = 100%) and Ethiopian (Nilo-Sudanic) and Middle and lower Congo species, clade II, which is the sister group to clade 1, comprises *L. coubie, L. macrostoma,* and *L. niloticus* species-groups (PP = 80%), and clade III included *L. umbratus* species group, which is allied to some Asian species (Afro-Asian clade).

Both our model-based Bayesian (Fig. 1) and ML analyses (results not shown) recognized the *L. forskalii* and *L. umbratus* Reid's (1985) species-groups as distinct lineages. However, taxa Reid (1985) included in *L. coubie*, *L. macrostoma*, and *L. niloticus* species-groups did not form monophyletic groups. These species-groups were rendered non-monophyletic as member species were always resolved in disparate clades (Fig. 1). On the other hand, 14 of the 18 African representative species of *Labeo* examined in this study were always resolved as monophyletic with strong nodal support (BS = ≥92%; PP = ≥96.4%). These include *L. altivelis*, *L. barbatus*, *L. batesii*, *L. capensis*, *L. coubie*, *L. cyclorhinchus*, *L. forskalii*, *L. lineatus*, *L. nasus*, *L. parvus*, *L. senegalensis*, *L. umbratus*, and *L. vulgaris*. On the other hand, *L. cylindricus* was rendered non-monophyletic because of the placement of two samples (GenBank accession #: NC031536 and AP011206) in subclade D within clade 1. However, *L. cylindricus* specimens from Ethiopia formed a strongly supported monophyletic group (subclade B; BS = 92, PP = 96.4%), which is the sister group to *L. forskalii*. The monophyly of *L. horie*, *L. longipinnis*, *L. sorex,* and *L. weeksii* could not be tested because only one representative sample was analyzed in the present study.

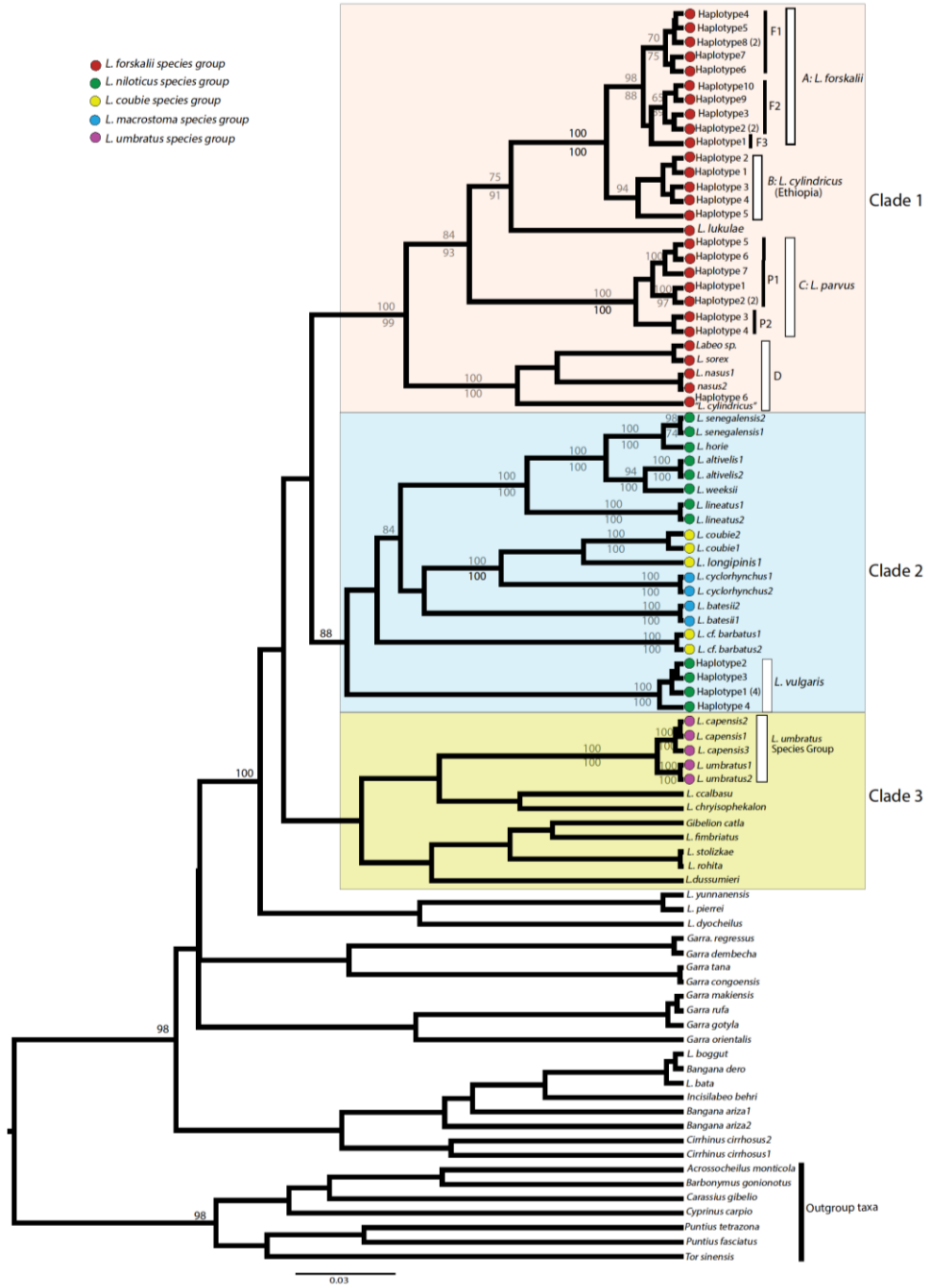


Fig. 1. Phylogenetic relationships of *Labeo* estimated from Bayesian analysis of partial and complete cyt *b* gene sequences. On the left side of the forward slash at the nodes are Maximum Likelihood bootstrap

values and on the right of the slash are Posterior Probability values derived from Bayesian analysis. Numbers at tip labels correspond to sample codes in Table 1 and numbers in parenthesis show number of samples for a given haplotype. Colour coded circles correspond to Reid's (1985) species-groups. The scale bar indicates base substitutions per site.

## Intraspecific diversity and interspecific genetic distances

Strongly supported genetic structuring exists within *L. forskalii* (clade A), which is further subdivided into at least three subclades (Fig. 1). One clade (clade F1), herein referred to as Abay-Tekeze clade, includes haplotypes 4, 5, 6, 7, and 8 originating from Abay (Blue Nile), Sanja (a tributary of Tekeze River in the northernmost part of the northwestern Ethiopian highlands) and Dedessa (tributary of Abay) Rivers. The second clade (clade F2), herein referred to as the Omo-Ghibe clade, includes haplotypes 2, 3, 9, and 10 from Ghibe and Gojeb rivers. Both clades received strong support (BS > 81 and PP >99%). One haplotype (haplotype 1) from Alwero River (which drains into the Baro River in western Ethiopia) formed its own clade (Baro-Akobo clade). Consistent with results from our phylogenetic analyses (Fig. 2), further examination of these clades in TCS recovered strong geographic partitioning within *L. forskalii*. These *L. forskalii* geographic clades were separated from each other by five to seven mutational steps. The overall Pairwise within group sequence divergence for *L. forskalii* ranged from 0.1± 0.003 to 1.2 ± 0.003%.

Within *L. parvus*, two geographically distinct sub-clades (or haplogroups) were resolved (Figs. 1 and 2): sub-clade P1, comprising haplotypes (haplotypes 1, 2, 5, 6, and 7) from Baro River in western Ethiopia; sub-clade P2 includes one haplotype (haplotype 3) from Benin, western Africa (GenBank #: JX074292) and another (haplotype 4) from unknown locality (GenBank #: AP013339). Consistent with our phylogenetic analysis, TCS analysis recovered similar haplotype clusters (Fig. 2B). The two sub-clades were separated from each other by five mutational steps. Monophyly of these sub-clades is strongly supported; BS = 83% and PP = 99.7% for subclade P1 and BS = 99% and PP = 99.9% for sub-clade P2. Our analysis also detected further structuring within subclade P2 (Fig. 1). *L. parvus* has a haplotype diversity (Hd) of 0.96 (±0.007 SD) and a nucleotide diversity (Pi) of 0.5%. The overall net genetic divergence between the two *L. parvus* sub-clades was 0.81%±0.003. According to our molecular dating, the mean age estimate for the split between *L. parvus* sub-clades was during the mid-Pleistocene (1.2 MYA).

Genetic diversity within the four Ethiopian species (*L. cylindricus*, *L. forskalii*, *L. parvus*, and *L. vulgaris*) examined in this study varied considerably. Among these species, we found a relatively high level of genetic diversity in *L. cylindricus*: Haplotype (gene) diversity (Hd) = 1.0 ± 0.13 and nucleotide diversity (pi) = 0.72%. Consistent with this, our phylogenetic analysis identified at least three putative clades (CL1, CL2, and CL3) within *L. cylindricus* from Ethiopia (Fig. 1). These clades received from moderate to high support values. Pairwise within group genetic distance in *L. cylindricus* ranged from 0.2 to 1.7% with a within group mean genetic distance of 1.0%.

Genetic divergence (results not shown) among African species ranged between 0.015 (*L. altivelis* vs *L. weeksii*) and 0.166 SE (*L. batesii* vs *L. nasus* and *L. batesii* vs *L. longipinnis*). The most divergent species among Ethiopian species is *L. vulgaris* separated from other Ethiopian species by large genetic divergences (*L. vulgaris* vs *L. cylindricus* = 0.10, *L. vulgaris* vs *L. forskalii* = 0.13, *L. vulgaris* vs *L. parvus*, = 0.11, *L. vulgaris* vs *L. horie* = 0.10).

**Divergence time estimations**

In our phylogeny South African species were grouped with some Asian species than with their African counterparts. A fossil record that can be an ancestor to both African and Asian lineages is not available. The earliest published fossil record of *Labeo* in Africa was from Lake Albert within the Nilo-Sudanic Ichthyological province (Stewart, 2003; 2009). Therefore, for divergence date estimation, we constrained the node comprising clades 1 and 2 in Fig. 3 at 13.8 MYA. Accordingly, the founding of clades 1 and 2 at about 13.8 MYA represents the most basal split within African *Labeo* (South African species excluded). Clade 1 comprises taxa from Eastern Nilo-Sudanic (subclades A + B + C) and lower and middle Congo (subclade D) River systems. Clade 2 is made by an assembly of species from the Congo and Eastern and western Nilo-Sudanic River drainages. The next divergence event at 9.9 MYA led to a phylogeographic split within clade 1 leading to the founding of two geographic clades, a clade containing subclades A + B + C (East Africa) and subclade D from the middle and lower Congo River systems. Subsequent cladogenetic events that started about 7.3 MYA produced significant phylogeographic structure within the east African clade (A + B + C). Events that diversified clade 2 of *Labeo* into nodes that subsequently gave rise to many species started around 12.6 MYA and continued until the late Pleistocene (0.5 MYA).

## DISCUSSION

### Phylogenetic and biogeographic inferences

Although some relationships, especially those at relatively deeper levels are unresolved, this study provides interesting insights into the taxonomy of African *Labeo* as it challenges the monophyly of *L. cylindricus* and recovers previously unknown cryptic diversity in Ethiopia. The study also provides evidence corroborating strong correlations between past historical events and extant distribution of genetic diversity within *Labeo* in Africa.

In contrast to previous morphological (Tshibwabwa, 1997) and molecular (Lowenstein *et al*., 2011; Yang *et al*., 2012; Ramoejane, 2016) studies, our phylogeny does not support the monophyly of African species, because South African species (*L. capensis* and *L. umbratus*) are embedded within Clade 3 along with some Asian species, (PP = 97.4; Fig. 1). The observed non-monophyly of *Labeo* in Africa may be attributed to multiple independent in-to-Africa dispersals of *Labeo*. Our phylogenetic reconstructions, however, reveal a moderately supported African clade (clade 1 + clade 2; PP = 85%) excluding South African species. This clade (mean age 13.8 My) splits into two mtDNA clusters: Lower and Middle Congo + Nilo-Sudan cluster (Clade 1 or *L. forskalii* species group) and Congo + Nilo-Sudanic cluster (clade 2). According to our molecular dating a mid-Miocene divergent event at about 9.9 MYA separates Clade 1 into two geographically concordant clades: clades A + B + C (Nilo-Sudanic clade comprised of Ethiopian taxa) and D in the Congo River system. The divergence event that isolated these two clades overlaps with the late Miocene separation of Congolese and Nilo-Sudanic lineages of *Hydrocynus* at about 6.8 MYA (CI: 3.2-10.8; Goodier *et al*., 2011) and African Crocodylus, dated at 8.13 My (CI: 5.24–12.64; Hekkala *et al*., 2011). The separation of the Nilo-Sudanic (clade A + B + C) and Congolese (clade D) clades concurs with the timing of tectonic uplifts associated with the Eastern and western flanks of the East African Rift System (ERS), which occurred since the Miocene (4.0–12.0 MYA, Stankiewicz and Wit, 2006; Roller *et al*., 2010; Speigel *et al*., 2010; Pinton *et al*., 2013).

Embedded within the Nilo-Sudanic clade (i.e., clade A + B + C), two dated nodes both representing relatively recent divergence events and comprising exclusively Ethiopian taxa are recovered with strong nodal support. The first Pliocene divergence event dated 7.3 MYA separated clade C (*L. parvus*) from clade A + B (*L. cylindricus* + *L. forskalii*). A more recent mid-Pleistocene (2.1 MYA) speciation event founded two sister species, *L.*

*cylindricus* (clade B, mean age 1.2 My) and *L. forskalii* (clade A, mean age 0.98 My). The Ethiopian plateau is dissected by Afar Depression and Main Ethiopian Rift (MER) into the northwestern plateau that harbours the Tekeze, Abay (Blue Nile), Baro-Akobo, and Omo-Ghibe River systems and southeastern plateau occupied by Wabi Shebelle and Genale River systems. All Ethiopian *L. cylindricus* specimens analyzed in this study were collected from the Genale River that originates in the South Eastern Ethiopian plateau east of the MER and flows south east into Somalia. On the other hand, all our *L. forskalii* specimens originated from River drainages in the north western plateau. According to Boccaletti *et al.* (1998) and Bonnini *et al*. (2005) the east-west directed rift extension in the Main Ethiopian Rift that resulted in the isolation of the Nubian and Somalian tectonic plates occurred at the Pliocene-Quaternary boundary. If the separation of the Nubian and Somalian tectonic plates contributed to the separation of clade A (*L. forskalii*) and clade B (*L. cylindricus*), then we expect the divergence time for these two lineages to date 1.8–3.6 MYA. The timing of geographical isolation between *L. cylindricus* and *L. forskalii*, which is constrained at 2.1 MYA in our phylogeny (Fig. 3), strongly concurs with the Pliocene-quaternary separation of the Nubian and Somalian plates, suggesting that Genale River's link to the Nile and Omo River systems was severed during this time. The timing of the cladogenetic event that isolated *L. cylindricus* and *L. forskalii* coincides with the founding of the Genale clade of *Labeobarbus*, dated at 2.3 MYA (CI: 1.6–3.0 MYA; Kebede Alemu *et al*., 2016).

Clade 2 (Mean age 12.6 My) is geographically a highly disjunct clade and comprises species from Congolese, Nilo-Sudanic, and Lower Guinea Ichthyological provinces. African species assigned by Reid (1985) to *L. coubie*, *L. macrostoma*, and *L. niloticus* species-groups are also embedded within this clade. In clade 2 species from geographically disjunct regions are closely related to each other than to species originating from the same geographic region. For instance, *L. altivelis* from the Congo River drainage is more closely related to *L. senegalensis* from the Nilo-Sudanic province than to *L. lineatus*. Similarly, *L. horie* from Ethiopia (Eastern Africa) exhibits a closer relationship with *L. senegalensis* (Western Africa) than with *L. vulgaris* from Ethiopia. The observed pattern of genealogy within clade 2 might be the outcome of repeated disruptions of geographic barriers between water systems within and across the Congo and Nilo-Sudanic provinces. This is consistent with the complex geological history of Africa, which is characterized by repeated uplift and rifting events since the

Miocene (Stankiewicz and Wit, 2006).

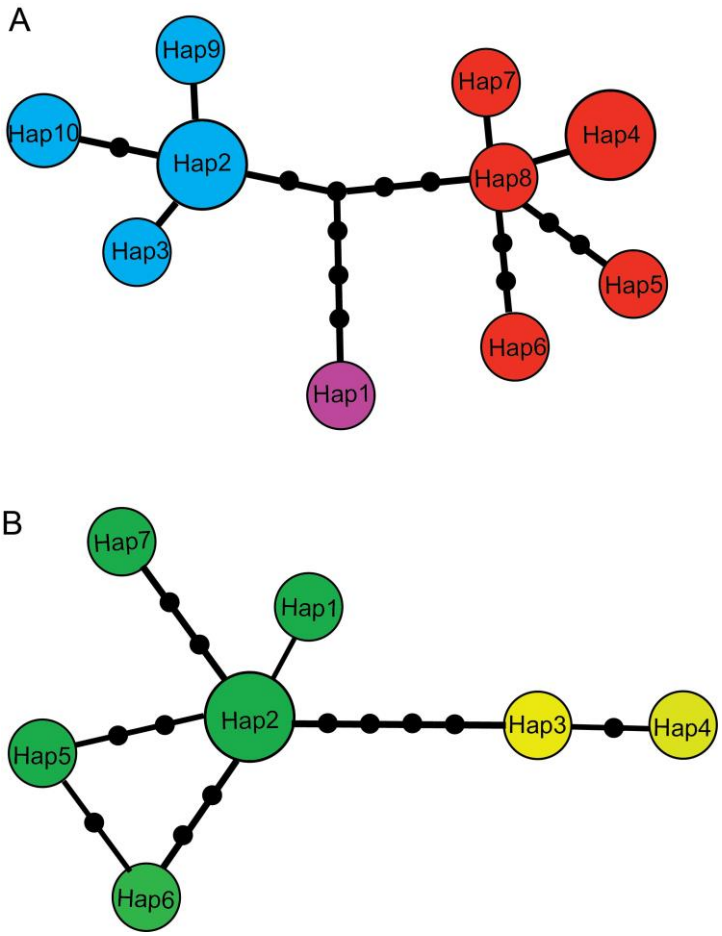


Fig. 2. Statistical Parsimony network for cyt *b* gene sequences of 10 and 7 unique haplotypes of *L. forskalii* (A) and *L. parvus* (B), respectively. Haplotypes are connected to each other with 95% confidence Interval. Circles correspond to unique haplotypes with numbers in circles indicating haplotype number while circle sizes reflect the frequency of each haplotype. Lines connecting circles (including the small dark circles) correspond to mutation steps. Geographic origins of haplotypes are color coded; A) *L. forskalii*: Abay-Tekeze River drainage, Omo-Ghibe River drainage, Baro-Akobo River drainage B) *L. parvus*: Baro River drainage, West African drainage (P2).

## Testing historical classification hypothesis

Reid (1985) subdivided African Labeo into six species-groups within African *Labeo*: *L. coubie*, *L. forskalii*, *L. gregorii*, *L. niloticus*, *L. macrostoma*, and *L. umbratus*. However, he conceded that his divisions may not be monophyletic. Our mtDNA study attempted to test the monophyly of

five of these subdivisions. Reid's (1985) *L. forskalii* (BS = 82%, PP =100%) and *L. umbratus* (BS = 99%, PP = 100%)) species-groups are monophyletic, whereas *L. coubie*, *L. macrostoma*, and *L. niloticus* species are rendered non-monophyletic since member species were always resolved in disparate clades (Fig. 1, indicated by coloured circles), challenging the validity of Reid's (1985) divisions. However, because we were unable to examine all the species within Reid's (1985) species-groups, further investigation incorporating more species is necessary to verify our results. On the other hand, fourteen of the 18 nominal African *Labeo* species examined in this study are always resolved as monophyletic each with strong nodal support (BS = ≥92%; PP = ≥96.4%). These include *L. altivelis*, *L. barbatus*, *L. batesii*, *L. capensis*, *L. coubie*, *L. cyclorhinchus*, *L. forskalii*, *L. lineatus*, *L. nasus*, *L. parvus*, *L. senegalensis*, *L. umbratus*, and *L. vulgaris*. The monophyly of the following African species could not be tested since only one representative sample was analyzed in the present study (*L. horie*, *L. longipinnis*, *L. sorex*, and *L. weeksii*).

From a taxonomic point of view, the most important finding of this study is the non-monophyly of *L. cylindricus* with one putative *L. cylindricus* haplotype (haplotype 6; GenBank accessions: NC031536.1 and AP011206.1) being grouped with *L. nasus*, *L. sorex* and an unidentified specimen, *Labeo* sp. (GenBank # AP013323), under clade D (BS = 99%, PP = 100%). This haplotype was highly divergent (mean genetic distance 9.6%) from *L. cylindricus* haplotypes originating from Eastern Ethiopia. The genealogical affiliation and high genetic divergence of haplotype 6 suggest a complete phylogeographic gap between this haplotype and the *L. cylindricus* clade of Ethiopia (clade B). Our time calibrated mtDNA tree (Fig. 3) shows that clade D diverged from the rest of the species within Reid's (1985) *L. forskalii* species group at approximately 9.9 MYA (Fig. 3). The timing of this divergence corresponds to the late Miocene divergence of Congolese and Nilo-Sudanic lineages of African *Hydrocynus* at 6.8 MYA, CI: 3.2–10.8 (Pinton *et al*., 2013) and Crocodylus at 8.13 MYA, CI: 5.24–12.64 (Hekkala *et al*., 2011). These cladogenetic events are consistent with tectonic events of the Miocene period that led to watershed isolation between Congo + Zambezi and Nilo-Sudan Ichthyological provinces (Pinton *et al*., 2013).

The geographic origins of all the samples clustered in clade D including the two *L. cylindricus* specimens (haplotype 6) are not provided in GenBank. However, available literature suggests that *L. nasus* (type locality Cataracts at Yelala, Lower Congo River; Reid, 1985) and *L. sorex* (type locality Lualaba, tributary of Congo River; Reid, 1985) are distributed in the lower

and middle Congo River basin, whereas the geographic range of *L. cylindricus* (type locality Tete, Zambezi River; Reid, 1985) spans Central, East, South Eastern and South Western Africa (Reid, 1985; Tshibwabwa, 1997; Skelton, 2001; https://www.iucnredlist.org/). These African subregions harbour major river drainages (e.g., Congo, Nile, Niger, Senegal, Genale, Limpopo, Zambezi etc.) that are known to have been isolated from each other for long periods of time. Judging from the genealogical distinctness of the two *L. cylindricus* specimens (NC031536.1 and AP011206.1) from *L. cylindricus* from Genale River (Eastern Ethiopia), their phylogenetic affinity with *L. nasus* and *L. sorex*, and regional biogeography of other fish taxa (Goodier *et al*., 2011; Pinton *et al*., 2013), the two *L. cylindricus* specimens are more likely to have originated from the Congo or Zambezi Ichthyological province than East African drainages of the Nilo-Sudanic province. The observed paraphyly in *L. cylindricus* is, therefore, most likely attributed to designating a single species for what could be phylogenetically distant and taxonomically distinct groups, suggesting that *L. cylindricus* as currently recognized may represent more than one species. Given the type locality of *L. cylindricus* (Tete, Zambezi River; Reid, 1985), the Ethiopian specimens identified as *L. cylindricus* could belong to a new species or other previously described species not included in our study. Therefore, the present phylogenetic study indicates that the taxonomic status of *L. cylindricus* as currently circumscribed is questionable and needs further investigation.

**Cryptic diversity within *L. forskalii***

Another important outcome of this study is the recovery of previously undescribed cryptic diversity within some species, namely *L. cylindricus*, *L. forskalii*, and *L. parvus* (Fig. 1) in Ethiopia. Because all our *L. cylindricus* embedded within subclade B originated from the same locality, there is no obvious geographic structuring among them. Therefore, herein we discuss only *L. forskalii* and *L. parvus* in more detail.

Two Pleistocene divergence events founded three subclades within *L. forskalii* in Ethiopia that geographically correspond to Abay (Blue Nile) and Tekeze River drainages (F1), Omo-Ghibe River drainage (F2), and Baro-Akobo River drainage, which is part of the White Nile River system (F3). The first main divergence event, dated 0.98 MYA, separated sub-clade F1 from sub-clade F2 + F3. According to Feibel *et al.* (1989) and Johnson and Malala (2009), Lake Turkana, into which the Omo-Ghibe River system drains, occasionally overflowed into the Nile drainage system 1.3 MYA.

Therefore, the timing of the phylogeographic split of *L. forskalii* into sub-clades F1 and F2 + F3 appears to concur with the timing of tectonic events in the basin that shifted the direction of Turkana overflow during humid times from a corridor to the Indian Ocean to the Nile drainage system. The second divergence event led to the geographical isolation of clade F2 from clade F3 (comprising haplotype 1) at about 0.79 MYA. Lake Turkana is believed to have been connected to the White Nile via the Lotagipi Swamps (Kenya, west of Lake Turkana), and the Sobat and Pibor Rivers (Baro-Akobo River system) in the early and middle Holocene (Grove, 1983; Feibel *et al*., 1989; Johnson and Malala, 2009). Our molecular clock results, however, provide a much older estimate of faunal connection (0.79 MYA, late Pleistocene) between Lake Turkana (and hence Omo-Ghibe River system) and the Nile system.

According to Kebede Alemu and Harris (2014), *Labeobarbus intermedius*, another cyprinid species occurring roughly in a similar range as *L. forskalii*, exhibits similar phylogeographic structure. In their extensive phylogeographic analysis, Kebede Alemu and Harris (2014) identified two sister lineages within *Labeobarbus intermedius* that have their distribution limits defined by Addis Ababa-Nekemt and Goba-Bonga Tectonic Lineaments; both *L. forskalii* and *Lb. intermedius* include genetically distinct northern (F1 in *L. forskalii*) and southern (F2 + F3 for *L. forskalii*) lineages distributed north and south of these lineaments. Further confirmation of this phylogeographic pattern is provided by Levin *et al*. (2020). The geographic isolation of *L. forskalii* haplogroups is constrained at about 0.98 MYA (Fig. 3), which points to the late Pleistocene founding of the three geographically delimited haplogroups of *L. forskalii*. A corresponding phylogeographic split in *Lb. intermedius* occurred 0.5 MYA (late Pleistocene; Kebede Alemu and Harris, 2014).

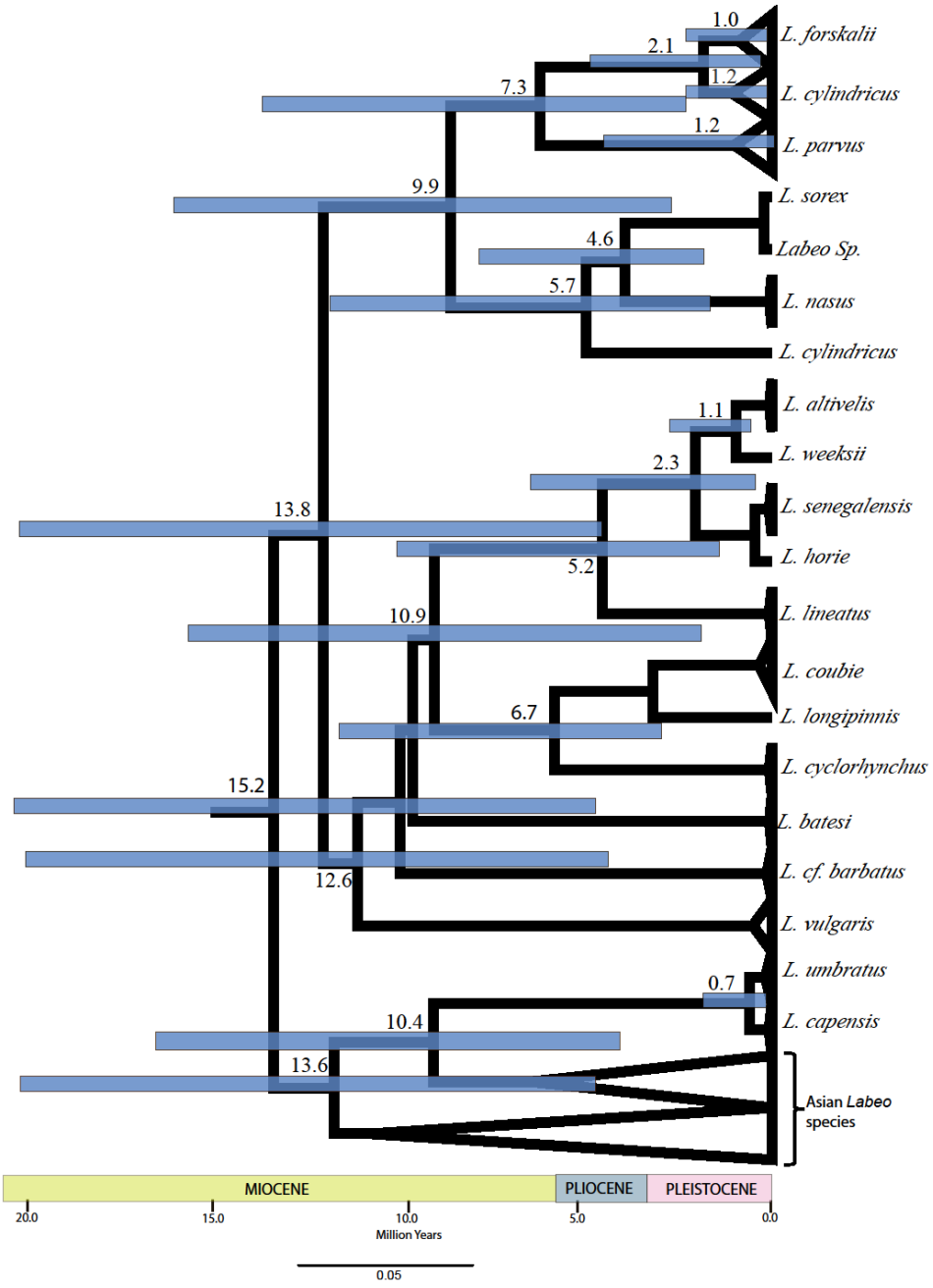


Fig. 3. Time calibrated tree of cyt *b* sequences of *Labeo* generated in Beast employing the GTR nucleotide substitution and uncorrelated lognormal relaxed molecular clock models with Birth Death Speciation tree prior. Numbers at nodes correspond to divergence date estimates of clades in million years. Error bars indicate the 95% HPD associated with node dates.

The observed phylogeographic concordance between *L. forskalii* and *Lb. intermedius* suggests that the two species may share geographic barriers or common history. Despite this however, the subclades of the two species showed different levels of genetic divergence; the *L. forskalii* haplogroups exhibit relatively lower levels of genetic divergence (0.5–0.6%) than *Lb. intermedius* lineages (1.5%, Kebede Alemu and Harris, 2014). This discrepancy may be attributed to differences in life history traits between the two species. Species-specific life history traits have been implicated as one of the major factors that shape extant phylogeographic patterns in fish populations (Bowen and Avise, 1990). In addition, the current study examined small sample size and limited geographic sampling (12 samples/10 haplotypes) of *L. forskalii* in contrast with Kebede Alemu and Harris (2014), who identified 45 unique haplotypes among 159 *L. intermedius* samples. Owing to the lower level of genetic divergence among recovered sub-clades and limited sampling, we refrain from drawing any taxonomic conclusions. It is likely that the analysis of larger and broader geographic samples of *L. forskalii* could dramatically influence phylogeographic inferences. Future phylogeographic studies of *L. forskalii* are certainly warranted.

## Cryptic diversity within *L. parvus*

Analysis of our mitochondrial DNA dataset showed that at least two genetically distinct subgroups are found within *L. parvus* (Fig. 1 and 2). The first group (P1) is composed of haplotypes from Baro River (western Ethiopia) and the second subgroup included two haplotypes, one from Benin in Western Africa (Nilo-Sudanic province) and another (GenBank #: AP013339) from an unknown locality. The lack of source information for the GenBank sequence (Accession #: AP013339), however, did not allow for a definitive assessment of whether these genetically distinct subgroups are geographically differentiated or not. Based on the results of this study, information on the geographic distribution of the species, and comparative phylogeography of other African freshwater fishes in the Nilo-Sudanic region, we speculate that the *L. parvus* sample (GenBank accession #, AP013339) probably originated from West African drainages. The geographic range of *L. parvus* includes western and central African drainages (Lévêque and Daget, 1984; Lévêque *et al*., 1990; Tshibwabwa, 1997) as well as the Nile River system in East Africa. Our study revealed that there was little genetic differentiation between the two individuals of *L. parvus* embedded within P2; they have a very high percent sequence identity (99.82%, based on Blast analysis in NCBI) and very low sequence

divergence (0.17%) for cyt *b* gene. Molecular dating of divergence between Nilo-Sudanic + Lower Guinea and Congolese lineages of African Tigerfish (Genus *Hydrocynus*) suggests that the geographic isolation of western African drainages and Congo River drainage occurred at around 6.8 MYA, when the youngest geologically-dated tectonic events modified the Cameroon Volcano-tectonic rift (Goodier *et al*., 2011). Given the long isolation of the Congo River basin from the Nilo-Sudanic drainages and the extremely low level of genetic differentiation between the two sequences embedded within haplogroup P2 of *L. parvus*, the sequence with unknown sampling locality (Accession #: AP013339) was probably withdrawn from West African drainages (if not from the same drainage) and less likely to have originated from the Congo River basin.

If the assumption that the mtDNA sequence (Accession #: AP013339) originated from west African drainages holds true, our results clearly support strong geographic differentiation between the western and Eastern Nilo-Sudanic populations of *L. parvus*. Our molecular dating indicated that the isolation of these populations occurred during the mid-Pleistocene (1.12 MYA), suggesting that a relatively recent connection existed between the eastern and western drainages of the Nilo-Sudanic Ichthyological province. This is consistent with the timing of the isolation of *Hydrocynus forskahlii* into Nilotic and west African clades at approximately 2.1 MYA, CI: 0.6–3.9 (Goodier *et al*., 2011), the existence of close phylogenetic affinities between Nilotic and West African Mormyrids (Levin and Golubtsov, 2017) as well as *L. horie* and *L. senegalensis* (current study; Yang *et al*., 2012), and panmixis in *Hydrocinus brevis* from Gambia and Niger Rivers and Nile River (Goodier *et al.*, 2011).

Overall, although preliminary, our study provides new insights into the phylogeny, biogeography and taxonomy of *Labeo* in Africa with evidence for cryptic diversity within Ethiopia and incongruence between morphology based taxonomy of *L. cylindricus* and molecular phylogeny. One limitation related to our study was that a moderate set of African species was analyzed. Therefore, we urge our readers to exercise caution in interpreting our results. Further study based on increased taxon sampling may improve phylogenetic resolution and thereby provide more compelling biogeographic and taxonomic inferences and divergence date estimates.

**ACKNOWLEDGEMENTS**

This research was funded by National Science Foundation (NSF), Cypriniformes Tree of Life Project (DEB 0431263 to PMH). We thank the University of Kansas Biodiversity Research Center for providing some tissue samples used in this study. Some mtDNA sequences used for this study were retrieved from Genbank. We are also grateful to Malcolm F. Vidrine for reviewing the manuscript and providing thoughtful suggestions. A sincere thank you goes to Ruth Alemu for her diligent proofreading of this manuscript.

**REFERENCES**

Berendzen, P.B., Gamble, T., and Simons, A.M. (2008). Phylogeography of the bigeye chub *Hybopsis amblops* (Teleostei: Cypriniformes): Early Pleistocene diversification and post-glacial range expansion. *J. Fish Biol*. **73**: 2021–2039. http://dx.doi.org/10.1111/j.1095-8649.2008.02046.x

Boccaletti M., Bonnini M., Mazzuoli, R., Bekele Abebe, Piccardi L. and Tortorici L. (1998). Quaternary oblique extensional tectonics in the Ethiopian Rift (Horn of Africa). *Tectonophysics* **287** (1–4): 97–116.

Bonnini, M., Corti, G., Innocenti, F., Manetti, P., Mazzarini, F., Tsegaye Abebe and Pecskay, Z. (2005). Evolution of the Main Ethiopian Rift in the frame of Afar and Kenya rifts propagation. *Tectonics* **24**: TC 1007, 10.1029/2004TC001680.

Bowen, B.W. and Avise, J.C. (1990). Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: Influence of zoogeographic factors and life-history patterns. *Mar. Biol*. **107**: 371–381. doi: 10.1007/BF01313418

Briolay, J., Galtier, N., Britto, R.M. and Bouvet, Y. (1998). Molecular phylogeny of Cyprinidae. Inferred from cyt *b* DNA sequences. *Mol. Phylogenet. Evol*. **9**: 100–108. doi: 10.1006/mpev.1997.0441

Drummond, A.J., Rambaut, A., Suchard, M.A. and Xie, W. (2002–2018). Bayesian Evolutionary Analysis Utility Version v1.10.4. Retrieved from http://github.com/beast-dev/beast-mcmc

Feibel, C.S., Brown, F.H. and McDougall, I. (1989). Stratigraphic context of fossil hominids from the Omo Group deposits: Northern Turkana Basin, Kenya and Ethiopia. *Am. J. Phys. Anthropol.* **78**: 595–622.

Gashaw Tesfaye and Wolff, M. (2014). The state of inland fisheries in Ethiopia: A synopsis with updated estimates of potential yield. *Ecohydrol. Hydrobiol*. **14**: 200–219.

Golubtsov, A.S. and Mina, M.V. (2003). Fish species diversity in the main drainage systems of Ethiopia: Current state of knowledge and research perspectives. *Ethiop. J. Nat. Resour.* **5**: 281–318.

Golubtsov, A.S. Dgebuadze, Y.Y. and Mina, M.V (2002). Fishes of the Ethiopian Rift Valley. In: **Biology of Inland Waters Series**. Backhuys Publishers, Leiden.

Goodier, S.A.M., Cotterill, F.P.D., O'Ryan, C., Skelton, P.H. and De Wit, M.J. (2011). Cryptic diversity of African Tiger fish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked Neogene geotectonic events. *PLoS One* **6**(12): 1–15. https://doi.org/10.1371/journal.pone.0028775

Greenwood, P.H. (1972). New fish fossils from the Pleiocene of Wadi Natrun, *Egypt. J. Zool. London*. **168**: 503–519. https://doi.org/10.1111/j.1469-7998.1972.tb01364.x

Grove A.T. (1983). Evolution of the physical geography of the East African Rift Valley Region, pp. 115–155. In: **Evolution, Time and Space: The Emergence of the Biosphere** (Sims, R.W., Price, J.H. and Whalley, P.S.E., eds.). Academic Press, London.

Hall, T.A. (1999). BioEdit; a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser*. **41**: 95–98. doi :10.14601/Phytopathol_Mediterr-14998u1.29

Hekkala, E., Shirley, M.H, Amato, G., Austin, J.D., Charter, S., Thorbjarnarson, J., Vliet, K.A., Houck, M.L., Desalle, R. and Blum, M.J. (2011). An ancient icon reveals new mysteries: Mummy DNA resurrects a cryptic species within the Nile crocodile. *Mol. Ecol*. **20**: 4199–4215.

Johnson, T.C and Malala J.O. (2009). Lake Turkana and its link to the Nile. In: **The Nile**, pp. 287–297. DOI: 10.1007/978-1-4020-9726-3_15

Kebede Alemu and Harris, P.M. (2014). Mitochondrial DNA phylogeography of *Labeobarbus intermedius* complex (Pisces, Cyprinidae) from Ethiopia. *J. Fish Biol.* **85**: 228–245. http://dx.doi.org/10.1111/jfb.12408

Kebede Alemu, Harris, P.M. and Mayden, R.L. (2016). Novel evolutionary lineages in *Labeobarbus* (Cypriniformes; Cyprinidae) based on phylogenetic analyses of mtDNA sequences. *Zootaxa* **4093**: 363–381. http://doi.org/10.11646/zootaxa.4093.3.4

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol*. **35**: 1547–1549. doi: 10.1093/molbev/msy096

Leigh, J.W. and Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods. Ecol. Evol.* **6**: 110–1116. doi: 10.1111/2041-210X.12410

Lemma Abera, Abebe Getahun and Brook Lemma (2014). Composition of commercially important fish species and some perspectives into the biology of the African Catfish *Clarias gariepinus* (Burchell), Lake Ziway, Ethiopia. *Int. J. Curr. Adv. Res*. **2**: 864–871.

Lévêque, C., Paugy, D. and Teugels, G.G. (1990). Faune des poissons d'eaux douces et saumâtres d'Afrique de l'Ouest. Tome 1. Ostrom.

Lévêque, C. and Daget, J. (1984). Cyprinidae. In: **Check list of the Freshwater Fishes of Africa (CLOFFA), Volume** 1, pp. 217–342 (Daget, J., Gosse, J.-P. and Thys van den Audenaerde, D.F.E., eds.). ORSTOM, Paris and Tervuren, MRAC.

Levin, B.A. and Golubstov, A.S. (2017). New insights into the molecular phylogeny and taxonomy of mormyrids (Osteoglossiformes, Actinopterygii) in northern East Africa. *J. Zool. Syst. Evol. Res*. **56**: 61–76. https://doi.org/10.1111/jzs.12186

Levin, B.A., Simonov, E., Dgebuadze Y.Y., Levina, M. and Golubtsov, A.S. (2020). In the rivers: Multiple adaptive radiations of cyprinid fshes (*Labeobarbus*) in Ethiopian Highlands. Scientific reports. 10: 7192. doi: 10.1038/s41598-020-64350-4

Lowenstein, L.H., Osmundson, T.W., Becher, S., Hanner, R. and Stiassny, M.L.J. (2011). Incorporating DNA barcodes into a multi-year inventory of the fishes of the hyperdiverse Lower Congo River, with a multi-gene performance assessment of the genus *Labeo* as a case study. *Mitochondrial DNA* **21**(S2): 1–19. https://doi.org/10.3109/19401736.2010.537748

Pinton, A., Jean-François A., Paugy, D. and 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. https://doi.org/10.1016/j.ympev.2012.12.009

Rambaut, A. (2006–2018). FigTree, Tree Figure drawing tool version 1.4.4. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A. and Drummond, A.J. (2002–2018). Tree Annotator v1.10.4. Retrieved from http://beast.community.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. and Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* syy032. doi:10.1093/sysbio/syy032

Ramoejane, M. (2016). **Genetic Diversity, Evolutionary Relationships and Conservation of Southern African *Labeo* Fishes in Relation to Water Management**. Doctor of Philosophy, Rhodes University.

Reid, G.M. (1985). A revision of the African species of *Labeo* (Pisces: Cyprinidae) and a re-definition of the genus. *Theses Zoologicae*. **6**: 1–322.

Roller, S., Hornung, J., Hinderer, M. and Ssemmanda, I. (2010) Middle Miocene to Pleistocene sedimentary record of rift evolution in the southern Albert Rift (Uganda). *Int. J. Earth Sci*. **99**: 1643–1661.

Rozas, J., Ferrer-Mata1, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. and Sánchez-Gracia, A. (2017). DnaSP v6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol*. **34**: 3299–3302. doi: 10.1093/molbev/msx248

Simagegnew Melaku, Abebe Getahun and Mulugeta Wakjira (2017). Population aspects of fishes in Geba and Sor rivers, White Nile System in Ethiopia, East Africa. *Int. J. Biodivers*. https://doi.org./10.1155/2017/1252604.

Skelton, P.H. (2001). **A Complete Guide to the Freshwater Fishes of Southern Africa**. Second edition. Struik Publishers, Cape Town.

Skelton, P.H., Tweddle, D. and Jackson, P.B.N. (1991). **Cyprinids of Africa: Cyprinids Fishes: Systematics, Biology and Exploration.** Chapman and Hall, London.

Speigel C., Kohn, B.P., David, X., Belton, D.X. and Gleadow, A.J.W. (2010) Morphotectonic evolution of the central Kenya rift flanks: Implications for late Cenozoic environmental change in East Africa. *Geology* **35**: 427–430. doi: 10.1130/G23108A.

Stankiewicz, J. and Wit, M.J.D (2006). Proposed drainage evolution model for Central Africa – Did the Congo flow east? *J. Afr. Earth Sci*. **44**(1): 75–84. https://doi.org/10.1016/j.jafrearsci.2005.11.008

Stewart, K.M. (2001). The freshwater fish of Neogene Africa (Miocene–Pleistocene): systematics and biogeography. *Fish Fish*. **2**: 177–230. https://doi.org/10.1046/j.1467-2960.2001.00052.x

Stewart, K.M. (2003). Fossil fish remains from the Pliocene Kanapoi Site, Kenya. In: **Geology and Vertebrate Paleontology of the Early Pliocene Site of Kanapoi, Northern Kenya**, pp. 21–39 (Dumont, H.J., eds.). Springer, Dordrecht.

Stewart, K.M. (2009). Fossil fish from the Nile River and its southern basins: In: **The Nile: Origin, Environments, Limnology and Human Use**, pp. 677–692 (Dumont H.J., eds.). Columbia University Press, New York.

Stewart, K.M. and Murray, A.M. (2017). Biogeographic implications of fossil fishes from the Awash River, Ethiopia. *J. Vertebr. Paleontol*. e126911516. http://dx.doi.org/10.1080/02724634.2017.1269115.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J. and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**(1): 1–5. vey016 https://doi.org/10.1093/ve/vey016

Tadele Awoke (2015). Fish species diversity in major river basins of Ethiopia: A review. *World J. Fish Mar. Sci.* **7**: 365–374. doi: 10.5829/idosi.wjfms.2015.7.5.95195

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512–526.

Tang, Q.Y., Abebe Getahun and Liu, H.Z. (2009). Multiple in-to-Africa dispersals of labeonin fishes (Teleostei: Cyprinidae) revealed by molecular phylogenetic analysis. *Hydrobiologia* **632**: 261–271. doi: 10.1007/s10750-009-9848-z

Tan, M. and Armbruster, J. (2018). Phylogenetic classification of extant genera of fishes of the order Cypriniformes (Teleostei: Ostariophysi). *Zootaxa.* **4476**: 006–039. https://doi.org/10.11646/zootaxa.4476.1.4.

Tshibwabwa, S.M. (1997). Systématic des espèces Africaines du genre Labeo (Teleostei, Cyprinidae) dans les regions ichthyogéographiques de Basse Guinée et du Congo. Vol I et II. Presses Universitaires de Namurp 530.

Tshibwabwa, S.M. and Teugels G.G. (1995). Contribution to the systematic revision of the African cyprinid genus *Labeo*: Species from the lower Zaire River system. *J. Nat. Hist.* **29**: 1543–1579. doi: 10.1080/00222939500770661

Van Steenberge, M., Snoeks, J. and Vreven, E. (2016). Lingering taxonomic confusion in *Labeo* (Actinopterigii: Cypriniformes: Cyprinidae): Correcting the records and basis of type designations for seven Congolese species. *Acta. Ichthyol. Piscat*. **46**(1): 1–8. doi: 10.3750/AIP2016.46.1.01

Yang, L., Arunachalam, M., Sado, T., Levin, B.A., Golubtsov, A.S., Freyhof, J., Friel, J.P., Chen, W.J., Hirt, M.V., Manickam, R., Agnew, M.K., Simons, A.M., Saitoh, K., Miya, M., Mayden, R.L. and He, S. (2012) Molecular phylogeny of the cyprinid tribe Labeonini (Teleostei: Cypriniformes). *Mol. Phylogenet. Evol*. **65**: 362–379. https://doi.org/10.1016/j.ympev.2012.06.007

Yang, L. and Mayden, R.L. (2010). Phylogenetic relationships, subdivision, and biogeography of the Cyprinid tribe Labeonini (sensu Rainboth, 1991) (Teleostei: Cypriniformes), with comments on the implications of lips and associated structures in the Labeonin classification. *Mol. Phylogenet. Evol.* **54**: 254–265. doi: 10.1016/j.ympev.2009.09.027

Yang, L., Mayden, R.L., Sado, T., He, S., Saitoh, K. and Miya, M. (2010). Molecular phylogeny of the fishes traditionally referred to Cyprinini sensu stricto (Teleostei: Cypriniformes). *Zool. Scr*. **39**: 27–550. doi: 10.1111/j.1463-6409.2010.00443.x

Zeleke Berie (2007). **Diversity, Relative Abundance, and Biology of Fishes in Beles and Gilgel Beles Rivers, Abay Basin, Ethiopia.** M.Sc. Thesis. Addis Ababa University, Addis Ababa.

Zheng, L., Yang, J. and Chen, X. (2012). Phylogeny of the Labeoninae (Teleostei, Cypriniformes) based on nuclear DNA sequences and implications on character evolution and biogeography. *Curr. Zool.* **58**: 837–850. https://doi.org/10.1093/czoolo/58.6.837