GENETIC CHARACTERIZATION OF THE GENUS '*BARBUS*' (CYPRINIDAE) IN THE LAKE VICTORIA DRAINAGE SYSTEM, KENYA

A Thesis

by

VIOLET M'MBONE NDEDA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Luis A. Hurtado
Co-Chair of Committee,	Mariana Mateos
Committee Member,	James B. Woolley
Head of Department,	David Caldwell

December 2018

Major Subject: Wildlife and Fisheries Sciences

ABSTRACT

The genus '*Barbus*' belongs to a speciose taxonomically complex and heterogeneous assemblage of cyprinid fish. In Lake Victoria drainage basin in Kenya, ten endemic species of '*Barbus*' are reported, which play a significant role in food security and socio-economic development of the local community. Although these species are identified using morphological characters, confusion may occur when trying to distinguish morphologically similar species. Recent molecular work in the region has suggested presence of introgression within certain '*Barbus*' species further complicating the taxonomy and species identification in the group. In this study, we obtained cytochrome *b* and GH-intron 2 gene sequences of nine '*Barbus'* species sampled in the Lake Victoria drainage basin in Kenya. We conducted Maximum likelihood and Bayesian phylogenetic analyses to establish their evolutionary relationships in relation to other '*Barbus'*. The results showed distinct lineages of '*Barbus*' species not subjected to introgression/hybridization. Herein, we present new sequences of cytochrome *b* and GH DNA for small African '*Barbus'*. We also report new sequences of cytochrome *b* for *Labeobarbus altianalis* sampled from the study site. The analyses further established *'B.*' *profundus* to be a sister to '*B.' anema* and not '*B'. radiatus* as originally described, a finding that compliments previous morphometric and meristic data suggesting *'B.*' *profundus* to be a distinct species and not a subspecies of '*B'. radiatus*. These results demonstrate that molecular markers can provide additional support to inferences derived from morphological evidences. We hope that the newly established sequences from this

study will enrich the online reference database and allow future molecular species identification of the African Barbs. In addition, this study contributes to a better understanding of phylogenetic relationships and diversity of '*Barbus*' in the Lake Victoria Basin and Africa in general.

ACKNOWLEDGEMENTS

Glory to God the almighty for giving me the grace, patience, wisdom, physical and mental health throughout my stay in the United States of America.

Special thanks are due to my academic advisor Dr. Luis Hurtado and co-chair Dr. Mariana Mateos for their invaluable contribution, guidance and support throughout my studies. Their continued interest, inspiration and encouragement gave me the morale to realize my research objectives. I thank Prof. James Woolley, my committee member, for his suggestions and insightful comments regarding this work. My deepest gratitude goes to the entire staff of Kenya Marine and Fisheries Institute and in particular James Achiya, a fish taxonomist for facilitating shipment of specimens used in this study.

I also offer my sincerest gratitude to Dr. Cecilia Smith and Billy (Texas A&M Univ. Libraries) for their intellectual input in the generation of the study area map used in this thesis. Many thanks to Amanda K. Pinion, Heather Prestridge and Dr. Kevin Conway (Department of Wildlife and Fisheries Sciences and Biodiversity Research and Teaching Collections, Texas A&M University) for assistance with photoimaging of some fish specimens used in this study.

I sincerely thank my academic sponsors: Fulbright academic exchange programme and American Association of University Women for funding my studies and supporting me throughout my academic training. On a more personal note, I would like to thank my family and all my siblings for their love and support.

I also want to express gratitude to the lab members of the Mateos-Hurtado lab, who supported me during my research and studies. Special thanks go to Humberto Martínez and Xochil Delarosa for their advice and friendship.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a thesis committee consisting of Dr. Luis A. Hurtado and Dr. Mariana Mateos of the Department of Wildlife and Fisheries Sciences and Professor James Woolley of the Department of Entomology.

The Kenya Marine and Fisheries Institute, Kenya provided tissue samples analyzed in this study. Staff from the Biodiversity Research and Teaching Collections at TAMU performed the photo imaging of fish species. Study area map in chapter 3 was created with assistance from Cecilia Smith and Billy (Texas A&M University Libraries). The analyses depicted in Chapter 4 were conducted in part by Dr. Luis A. Hurtado and Dr. Mariana Mateos of the Department of Wildlife and Fisheries Sciences.

The student completed all other work conducted for the thesis independently.

Funding sources

Graduate study was supported by fellowships from Fulbright academic exchange program and American Association of University Women. Dr. Luis Hurtado made research funds for this project possible.

TABLE OF CONTENTS

Page

1. INTRODUCTION*

Human-driven factors such as habitat destruction, fragmentation, overexploitation and non-native species introduction are recognized globally as major threats to biodiversity (Vorosmarty *et al.* 2010). Freshwater ecosystems in particular stand at risk of multiple stressors mostly because they serve as focal points for human settlement, intensive agriculture, industry and domestic activities in many nations (Dudgeon *et al.* 2006). According to World Wildlife Fund (2014), approximately 76% of freshwater biota has declined, while hundreds to thousands of species are being lost before they can be identified or described (Collen *et al.* 2009; Darwall *et al.* 2011; Galewski *et al.* 2011). This problem is further exacerbated in Africa, a region of high, but poorly known biodiversity, which faces a rapid rate of human population growth, anthropogenic impacts and climatic change (Collen *et al.* 2014). Lake Victoria, located in East Africa, ranks as the second largest freshwater lake in the world, and the largest in Africa (WWF 2014). This lake harbors rich biodiversity, but suffers significant problems catalyzed by human activities (Salzburger *et al.* 2014).

The Lake Victoria drainage Basin (LVB) covers approximately 194,200 km^2 and stands as the most critical economic resource linking five riparian countries: Tanzania (44% of LVB area occupied), Uganda (16%), Kenya (22%), Burundi (7%) and Rwanda (11%) (Balirwa *et al.* 2003; Tumwebaze *et al.* 2007).

---

*Reprinted with permission from "Evolution of African barbs from the Lake Victoria drainage system, Kenya" by Ndeda VM, Mateos M, and Hurtado LA, 2018. *PeerJ* 6: e5762. Copyright [2018] by Violet M. Ndeda

In the last two decades, however, the ichthyofaunal richness within the lake catchment area has come under extensive human exploitation leading to a decline in overall productivity and fish biodiversity (van Zwieten *et al.* 2015; Witte *et al.* 2013). Most notable was the dramatic extinction of indigenous fish species reported in the early 20th century (Kaufman 1993; Witte *et al.* 1991). Subsequently, IUCN (2010) reported approximately 45% of fish species in the basin as either threatened, endangered or thought to be extinct. Despite the biodiversity hotspot status of the lake, fish species diversity within the basin remains poorly documented and unexplored (Balirwa *et al.* 2003).

Barbus Cuvier and Cloquet, 1816 (also known as "barbs") has traditionally comprised a taxonomically complex and heterogeneous assemblage of cyprinid fish, with more than 800 species distributed across Eurasia and Africa (Berrebi *et al.* 1996; Skelton *et al.* 2012; Skelton *et al.* 1991). This genus, however, has been regarded as a polyphyletic assemblage, and some members are presently assigned to other genera (Tsigenopoulos *et al.* 2002; Yang *et al.* 2015). *Barbus* has been considered the most species-rich genus of African cyprinids, with an estimated number of > 300 species (Leveque & Daget 1984; Skelton 1988; Skelton 1993; Skelton *et al.* 1991). African barbs show three levels of ploidy (Berrebi *et al.* 1996), diploids (2n = ca. 48, 50), tetraploids (2n = ca. 100) and hexaploids (2n = ca. 150); and, based on their body size, have been divided into two groups, commonly referred to as the large and small barbs. Large African barbs are characterized by an adult body size > 20 cm standard length (SL), the presence of parallel or converging striae on their scales, and are, in general, either tetraploids or hexaploids (Ren & Mayden 2016). Adult small African barbs are usually less than 20 cm SL, have

divergent scale striae (Agnèse *et al.* 1990), and are diploids, in general (Ren & Mayden 2016). The taxonomy of the African barbs, however, is highly problematic (Berrebi *et al.* 2014; Schmidt & Bart 2015; Schmidt *et al.* 2017; Yang *et al.* 2015). In addition, morphological characters used for the identification of species may be highly variable, and the correct identification of species and even genera can be challenging (Berrebi *et al.* 2014; Ren & Mayden 2016; Schmidt & Bart 2015; Yang *et al.* 2015). These characters include coloration patterns, head and body lengths, number of pairs of barbels and the presence/absence of ossified and serrated rays in the dorsal fin (Banister 1973; Golubstov & Berendzen 2005; Greenwood 1962).

Molecular studies are shedding light on the taxonomy, diversity, and evolutionary relationships of African barbs. The large hexaploid African barbs are now classified within the genus *Labeobarbus* (tribe Torini), which has been shown to correspond to a well-supported clade (Tsigenopoulos *et al.* 2010; Yang *et al.* 2015). The large tetraploid African barbs are classified within the genus *Pseudobarbus*, in the tribe Smiliogastrini (Yang *et al.* 2015). The small African barbs have also been assigned to the tribe Smiliogastrini, and Yang *et al.* (2015) proposed inclusion of all of them within the genus *Enteromius*. This proposal, however, has been criticized due to poor resolution in the phylogeny of Yang *et al.* (2015), limited taxon sampling, failure to include the type species '*Barbus'* (*Enteromius*) *potamogali* in the analyses, absence of nuclear markers, and the incorporation of morphologically distinct presumed monophyletic genera within this group, such as the genera *Barboides*, *Clypeobarbus*, and *Pseudobarbus* (Schmidt & Bart 2015). A recent study of African diploid barbs found that '*Barbus*' and allies (*Systomus*,

Barboides, *Clypeobarbus* and African tetraploid barbs) form a strongly supported clade; however, '*Barbus*' is not resolved as monophyletic (Ren & Mayden 2016), and is composed of three well-supported clades. Thus, the proposal for grouping all small African barbs within *Enteromius* is not supported. Herein, we follow Schmidt *et al.* (2017) and refer to the African diploid and tetraploids as '*Barbus'*.

More recently, a multilocus study conducted in Kenya, including the Lake Victoria drainage Basin (LVD), revealed evidence of introgression involving three small barbs species, further complicating the taxonomy and species identification in this group (Schmidt *et al.* 2017). Use of the Growth Hormone (GH) intron 2, however, provided insight into the placement of heterospecific individuals. This study also uncovered high levels of genetic divergence within some recognized species (i.e., *B. kerstenii*, *B. paludinosus,* and *B. apleurogramma*). Another study examined genetic diversity of the large barb *Labeobarbus altianalis* in the Kenyan LVD, suggesting genetic differentiation among the four rivers draining the lake (Chemoiwa *et al.* 2013). In this study, I conducted genetic characterization and phylogenetic analyses of '*Barbus'* from the LVD, using Cytochrome *b* and GH intron 2 DNA sequences. The study area overlaps with that of Schmidt *et al.* (2017), but includes new localities and species. My results expand our understanding of the taxonomy and diversity of barbs in the LVD region.

2. MATERIALS AND METHODS

2.1 Study area

The Lake Victoria catchment area (LVD) covers a surface area of approximately 194,220 km^2, out of which 69,000 km^2 is the lake itself. The catchment area is shared by five East African riparian states as follows: Tanzania 44%(85,448km^2), Uganda 16%(31,072km^2), Kenya 22%(42,724km^2), Burundi 7%(13,594 km^2) and Rwanda 11%(21,362 km^2) (Figure 1) (Hecky & Bugenyi 1992). The Kenyan part of the LVD lies at an altitude of 1,134 m above sea level and falls within longitudes 34° 0' E to 35°53' E and latitude 0° 30' N and 1° 12' S (Fig. 1). Several rivers, dams and satellite lakes within the basin (Rivers: Nzoia, Yala, Awach-Seme, Nyando, Sondu-Miriu, Kisian; Satellite Lakes: Sare and L. Kanyaboli; Dams: Ugege**,** Ulanda, Ufinya, Kosigah, Uriri, Stella, Kokech, and Ratang) are known to harbor various species of '*Barbus'*, including: '*B.' altianalis, 'B.' apleurogramma, 'B.' profundus, 'B.' cercops, 'B.' nyanzae, 'B.' yongei, 'B.' kerstenii, 'B.' jacksoni*, '*B.' neumayeri,* and *'B.' paludinosus.*

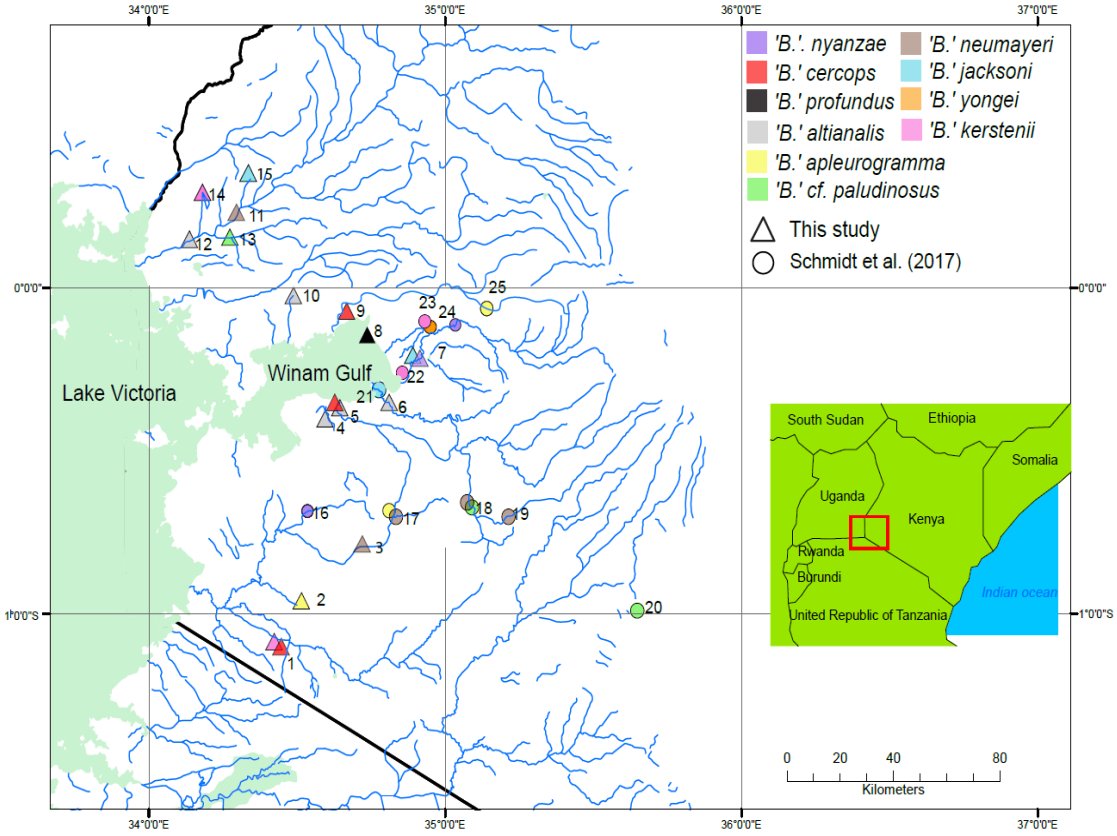

Figure 1. Localities of specimens sampled in this study (triangles) and those from Schmidt *et al.* (2017) (circles) (c.f. their Figure 1.) were georeferenced to a coordinate system. Colors indicate the putative species based on morphology and/or agreement of molecular data with those of Schmidt *et al.* (2017). The map was developed with ArcMap version 10.3—a part of the ESRI ArcGIS® Desktop suite.

2.2 Fish samples

Samples used in this study were collected and identified by fish taxonomists from Kenya Marine and Fisheries Institute (KMFRI). Approximately 400 specimens of *Barbus* were collected from seventeen localities in LVD (Fig. 1) and identified using morphological identification keys according to Greenwood (1962). A fin clip was taken from a subset of specimens and preserved in labeled 1.5 ml microtubes pre-filled with 96% ethanol. The remaining fish samples were preserved in 10% formalin and stored at KMFRI laboratories.

2.3 DNA isolation

Genomic DNA was extracted from ethanol-preserved fins of putative '*Barbus*' species using the DNeasy Blood and Tissue Kit (QIAGEN Inc.). The quality of the extracted DNA was examined by visualization on a 1.5% agarose electrophoresis gel, and quantification with a NanoDrop® ND-1000 spectrophotometer.

2.4 DNA amplification and sequencing

One mitochondrial (Cytochrome *b* ; Cytb; ~1140bp) and one nuclear gene (Growth Hormone Intron 2; GH; ~520bp) were PCR amplified from a total of 1- 4 specimens per locality. These genes have proved useful in the comparison of species within the African '*Barbus*' (De Graaf *et al.* 2007; Mayer *et al.* 1998; Mwita 2013; Mwita & Nkwengulila 2008; Schmidt *et al.* 2017). PCR was performed in a 25 µl reaction containing 19.9 µl ultrapure water, 0.5 µl dNTP mix (2.5 mM), 2.5 µl of 10X buffer, 0.5 µl of each 10µM primer, 0.1 µl *Taq* polymerase (OneTaq, New England Biolabs, Inc), and 1 µl of DNA template. Cytb was amplified with primers L15267 and H16461 according to Briolay *et*

al. (1998) while GH intron 2 was amplified using primers and protocols developed from (Mayden *et al.* 2009). These primers were Cytb L15267 (5'AATGACTTGAAGAACCACCGT3'), H16461 (5'CTTCGGATTACAAGACC3') and GH102F (5'TCGTGTACAACACCTGCACCAGC-3'), GH148R (5' TCCTTTCCGGTGGGTGCCTCA-3'). PCR amplification included a denaturation step of 2 min at 95°C followed by 35 cycles of 1 min at 95°C, 30 s at 58–60°C (Cytb)/ 55°C (GH) and 1 min at 72°C followed in turn by a final extension of 6 min at 72°C. Successful amplification was verified by running the PCR amplicons alongside a standard Lambda ladder on a 1.5% agarose gel stained with GelRed™ (Biotium Inc., Hayward, CA, USA). Products were sequenced bi-directionally using the above amplification primers in an ABI 3730 capillary sequencer.

2.5 Sequence assembly and alignment

Nucleotide sequences were assembled and edited with Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA). Newly generated Cytb sequences of '*Barbus'* were combined with publicly available sequences of African '*Barbus*' and their allies (*Systomus, Barboides, Clypeobarbus, Pseudobarbus, Labeobarbus*). We also included several sequences from Schmidt *et al.* (2017) provided by these authors prior to the public release of their sequences (i.e., seven species). Sequences were aligned with MAAFT v.6.0 (Katoh & Toh 2008). Aligned sequences were translated into amino acids to verify the alignments and to rule out the occurrence frameshifts and early stop codons that could be indicative of pseudogenes or sequencing errors. Species from the family Catostomidae were initially used as outgroups. Catostomidae represent tetraploids thought to have arisen

due to a hybridization event early (60 million years) in the history of the cypriniform fishes (Uyeno & Smith 1974). Following preliminary analyses of the above dataset, taxa were pruned to retain specimens relevant to this study: small barbs closely related to the taxa in this study; close relatives of '*B.*' *altianalis* (*Labeobarbus*), and four appropriate outgroup taxa (*Pethia ticto*, *Hampala macrolepidota*, *Puntigrus tetrazona*, and *Systomus sarana*; following Schmidt *et al.* (2017).

The GH dataset included newly generated sequences and representatives of seven *Barbus* species from the same region. Sequences of an appropriate outgroup were not publicly available for this gene.

2.6 Phylogenetic analyses

Phylogenetic analyses were performed using maximum likelihood (Stamatakis 2014), and Bayesian inference (Huelsenbeck & Ronquist 2001). Prior to employing the analyses, appropriate models of sequence evolution were determined using PARTITIONFINDER v2.7 (Lanfear *et al.* 2014) and JModeltest 2.1.9 (Darriba *et al.* 2012) under the Akaike Information Criterion (AIC), corrected AIC(c), and Bayesian Information Criterion (BIC) (Table 1).

Table 1. Description of characters and substitution models identified by model selection analyses. Best model selected by (a) JModelTest according to each criterion (AICc, AIC, BIC) and its corresponding weight, (b) the best partitioning scheme according to the BIC implemented in PartitionFinder (Lanfear *et al.* 2014)**.**

Gene	Non-redundant taxa	Characters used	Parsimony informative	Partitioning Scheme	AICc (weight)	AIC (weight)	BIC (weight)	Tree length	Consistency index (CI)	Retention index (RI)
GH	**14**	**211**	**71**	**1**	**JC (0.93)**	**TPM3uf+G (0.198)**	**TPM3uf (0.55)**	**181**	**0.86**	**0.98**
Cytb	**173**	**1023**	**494**	**1**	**TIM2+I+G (0.986)**	**TIM2+I+G (0.837)**	**TIM2+I+G (0.998)**	**4439**	**0.2**	**0.86**
				3 (by codon)						
				Codon 1			**SYM+I+G**			
				Codon 2			**HKY+I+G**			
				Codon 3			**GTR+G**			

Bayesian analyses were performed in MrBayes 3.2.6 (Huelsenbeck & Ronquist 2001) via the CIPRES Science Gateway (Miller *et al.* 2010). The analysis was run for 10,000,000 generations consisting of four independent Markov Chain Monte Carlo (MCMC) chains sampled at every 1000 generations. TRACER v1.6 was used to assess MCMC stationarity and to ensure adequate effective sampling size values (>200) were achieved. The first 25% of the sampled trees were discarded as burn-in, whereas the remaining sampled trees were summarized with "sumt" command implemented in MrBayes.

Maximum Likelihood (ML) analysis was implemented in RaxML v 8.2.6 (Stamatakis 2014) using rapid bootstrap and GTRGAMMA model via the CIPRES Science Gateway (Miller *et al.* 2010) to generate a maximum likelihood tree. Clade support was examined by a nonparametric bootstrap analysis of 200 replicates and summarized with 50% majority rule consensus tree computed using the SUMTREES script (v.3.3.1) (Sukumaran & Holder 2010).

3. RESULTS AND DISCUSSION

Phylogenetic reconstructions for the African small barbs using Cytb and GH DNA sequences are shown in Figures 2 and 3 included as separate/supplemental files. The sampling localities for the specimens characterized in this study are indicated by triangles (Fig. 1). Similarly, the approximate location of the specimens from Schmidt *et al.* (2017) included in our analyses are indicated by circles. Comparison of Cytb and GH trees does not suggest instances of introgression or hybridization for the samples characterized in this study. Distinct lineages were observed, for both Cytb and GH, for the nine species of small African barbs included in the analyses: *'B.' neumayeri*, *'B.' profundus*, *'B.' kerstenii*, *'B.' nyanzae*, *'B.' jacksonii*, *'B.' cercops*, *'B.' yongei*, *'B.' cf. paludinosus*, and *'B.' apleurogramma*. Of these, no specimens of *'B.' yongei* were collected in this study and the sequences used were from Schmidt *et al.* (2017).

This study is the first to report Cytb and GH DNA sequences for *'B.' profundus*, a species endemic to Lake Victoria (Greenwood 1970). Each of the four *'B.' profundus* specimens had a different Cytb haplotype (max. within clade divergence = 0.59% K2P). The Cytb phylogenetic tree (Fig. 2) shows this species in a highly supported clade with twelve other species reported from the Nilo-Sudan and Upper Guinea ichthyological provinces. Originally, *'B.' profundus* was described as a subspecies of *'B.' radiatus* (i.e., *Barbus radiatus profundus* Greenwood 1970). However, Stewart (1977) based on meristic and morphometric analyses concluded '*B'. profundus* is a separate species from *'B.' radiatus*. In this study, Bayesian and ML reconstructions did not find a sister

relationship between '*B'. profundus* and '*B'. radiatus*. Instead, we found a well-supported sister relationship between '*B'. radiatus* and a lineage comprised of '*B'. aspilus* and '*B'. cf. guirali*, congruent with Ren and Mayden (2016) who used the same sequences for these taxa in their study. In our study, the Bayesian and one of the ML analyses supported a sister relationship between *'B.' profundus* and *'B.' anema* from the Nile basin, suggesting an affinity between taxa of the East Coast and Nilo-Sudanian ichthyologic provinces, as proposed for other taxa (Roberts 1975). Phylogenetic analyses with additional species are necessary to assess the *'B.' profundus*-*'B.' anema* sister relationship.

This study is also the first to report Cytb DNA sequences for *'B.' cercops*. The Cytb sequences for the specimens identified as *'B.' cercops* clustered with a sequence from GenBank (AF180841) identified as *Barbus nyanzae* from Kenya (Tsigenopoulos *et al.* 2002). Therefore, specimen AF180841 probably represents a misidentification or an introgressed/hybrid individual. Maximum within-clade divergence was 0.39% K2P (including AF180841). The GH sequences of the individuals identified as *'B.' cercops* in this study formed a distinct cluster, thus, showing no indication of introgression with other species. Schmidt *et al.* (2017) reported that all specimens morphologically assigned to *'B.' cercops* in their study had a GH allele of *'B.' cercops*, but a Cytb haplotype of *'B.' neumayeri* or *'B.' cf. paludinosus*.

Levels of intraspecific sequence divergence in the Kenyan LVD varied among species. Maximum within-species divergence was 1.19% (K2P) for '*B.' kerstenii*, and 0.99% for *'B.' nyanzae* (max. within clade divergence = K2P). Similarly, the Cytb sequences obtained for *'B.' apleurogramma* in this study were 0.99% divergent (K2P)

from the available sequences reported by Schmidt *et al.* (2017). One of the two sequences obtained for *'B.' neumayeri* in this study was new and the other was identical to the one available from Schmidt *et al.* (2017). Maximum divergence within this species was 0.39% K2P. Three new haplotypes were discovered for *'B.' jacksonii*, but divergence within this species was low (max. within clade divergence = 0.20% K2P). Finally, the Cytb *'B.' cf. paludinosus* from this study were 1.59% divergent (K2P) from the ones available from Schmidt *et al.* (2017).

Compared to Cytb, GH sequences showed low levels of divergence. Only one allele was found for each of the following: *'B.' profundus*, *'B.' cercops*, *'B.' jacksonii*, *'B.' neumayeri*, *'B.' nyanzae* and *'B.' apleurogramma*. Two GH alleles were obtained for *'B.' kerstenii*. For *'B.' cf. paludinosus*, this study found only one GH allele, which is divergent from the two sequences available from Schmidt *et al.* (2017). For eight individuals of *'B.' nyanzae* only a partial clean sequence was obtained (194 bp; indicated by asterisks in Figure 2), which were identical to the corresponding fragment in individuals for which the whole sequence was obtained.

Poor resolution was obtained regarding the relationships among the African small barbs. The Cytb tree suggests that *'B.' kerstenii* and *'B.' nyanzae* are closely related, but probably not sister taxa. The Cytb tree also suggests the monophyly of *'B.' jacksonii* + *'B.' cercops* + *'B.' yongei* + *'B.' apleurogramma* + *'B.' cf. paludinosus*, but support for this clade is low. The GH tree is not in disagreement with the above relationships, but lack of adequate outgroups precludes stronger inferences. Slight discrepancies between the two genes are apparent. For example, '*B.' cercops* appears as sister to *'B.' yongei* in

the Cytb tree. In contrast, in the GH tree, *'B.' cercops* appears as sister to *'B.' jacksonii* (in agreement with Schmidt *et al.* (2017)), and *'B.' yongei* appears to be more closely related to '*B.' apleurogramma* and *'B.' cf. paludinosus.* Neither gene tree identifies a close relative of *'B.' neumayeri*.

Figure 4 included as separate/supplemental file, shows the Cytb phylogenetic reconstruction for *Labeobarbus*. This study is the first to report a Cytb sequence for *L. altianalis*. Only one haplotype was observed among the six *L. altianalis* individuals for which sequences were obtained, representing five localities in the Kenyan LVD. This is in sharp contrast with a previous study of this species in this area that reported high haplotype diversity for the mitochondrial control region Chemoiwa *et al.* (2013). The phylogenetic analyses indicate the Kenyan LVD *L. altianalis* haplotype reported in this study is at least ~2.5% divergent (K2P) from members of its sister lineage as shown in Figure 4. Banister (1973) proposed, based on morphology, two groups within *Labeobarbus*: the *Labeobarbus intermedius* complex (*L. intermedius*, *L. altianalis*, *'Barbus' acuticeps*, and *'B' rusae*) and the *Labeobarbus bynni* complex (*L. bynni*, *L. gananensis*, *L. oxyrhynchus*, and *'B' longifilis*). Cytb phylogenetic reconstructions in this study, however, do not support the monophyly of these groups. This is congruent with the findings of a previous phylogenetic analyses Beshera *et al.* (2016) that did not include *L. altianalis*.

4. SUMMARY AND IMPLICATIONS OF THIS STUDY

This study demonstrated distinct lineages of African *Barbus* from Lake Victoria basin that are not subjected to introgression. We identified new sequences of cytochrome b and GH DNA for small African '*Barbus*' species and also noted new sequences of cytochrome b for *Labeobarbus altianalis*. The newly identified sequences will be made available to the public through the GenBank online database to provide an invaluable reference tool for non-taxonomists and researchers to identify species. The analyses further established a sister relationship between '*B.' profundus* and '*B.' anema* and not '*B'. radiatus* as originally described, a finding that compliments previous morphometric and meristic data suggesting *'B.*' *profundus* to be a distinct species and not a subspecies of '*B'. radiatus*. These findings clearly demonstrate the effectiveness of molecular markers in complementing morphological data and contributes to the understanding of phylogenetic relationships and diversity of '*Barbus*' in the Lake Victoria Basin and Africa in general. We also detected errors in the species identities of small African *Barbus* deposited in the GenBank, suggesting taxonomic ambiguities in the online database due to misplacement of species or common errors attributed to mislabeling of original materials, contamination, or PCR-based errors. This study therefore reinforces the need for careful vetting of molecular databases and the integrated use of molecular markers and morphological evidences as an efficient and reliable tool to allow definitive conclusions about sample identity. Given the poor resolution within sister-group relationships demonstrated in this study, we suggest the use of the taxon name '*Barbus*' until more

completed studies on the group is undertaken. It is therefore hoped that the solutions to understanding the taxonomy and phylogenetic relationships among the African '*Barbus*' lies in an exhaustive sampling effort of the '*Barbus*' and use of additional loci.

REFERENCES

Agnèse J-F, Berrebi P, Lévêque C, Guégan J-F (1990) Two lineages, diploid and tetraploid, demonstrated in African species *Barbus* (Osteichthyes, Cyprinidae). *Aquatic Living Resources* **3**, 305-311.

Balirwa JS, Chapman CA, Chapman LJ, Cowx IG, Geheb K, *et al.* (2003) Biodiversity and Fishery Sustainability in the Lake Victoria Basin: An Unexpected Marriage? *BioScience* **53**, 703.

Banister K (1973) A revision of the large *Barbus* (Pisces, Cyprinidae) of east and central Africa. Studies on African Cyprinidae. Part II. *Bulletin of the British Museum (Natural History), Zoology* **26**, 1–148.

Berrebi P, Chenuil A, Kotlík P, Machordom A, Tsigenopoulos CS (2014) Disentangling the evolutionary history of the genus *Barbus* sensu lato, a twenty years adventure. *Professor Carlos Almaça (1934–2010)–Estado da arte em áreas científicas do seu interesse. Lisboa: Museu Nacional de História Natural e da Ciência*, 29-55.

Berrebi P, Kottelat M, Skelton P, Rab P (1996) Systematics of *Barbus*: state of the art and heuristic comments. Inst vertebrate biology as Cr kvetna 8, Brno 603 65, Czech Republic.

Beshera KA, Harris PM, Mayden RL (2016) Novel evolutionary lineages in *Labeobarbus* (Cypriniformes; Cyprinidae) based on phylogenetic analyses of mtDNA sequences. *Zootaxa* **4093**, 363-381.

Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular Phylogeny of Cyprinidae Inferred from Cytochrome b DNA Sequences. *Molecular Phylogenetics and Evolution* **9**, 100-108.

Chemoiwa EJ, Abila R, Macdonald A, Lamb J, Njenga E, *et al.* (2013) Genetic diversity and population structure of the endangered ripon barbel, *Barbus altianalis* (Boulenger, 1900) in Lake Victoria catchment, Kenya based on mitochondrial DNA sequences. *Journal of Applied Ichthyology* **29**, 1225-1233.

Collen B, Loh J, Whitmee S, McRae L, Amin R, *et al.* (2009) Monitoring change in vertebrate abundance: the living planet index. *Conservation Biology* **23**, 317-327.

Collen B, Whitton F, Dyer EE, Baillie JE, Cumberlidge N, *et al.* (2014) Global patterns of freshwater species diversity, threat and endemism. *Global Ecology and Biogeography* **23**, 40-51.

Darriba D, Taboada GL, Doallo R, Posada D (2012) JModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**, 772.

Darwall W, Smith K, Allen D, Holland R, Harrison I, *et al.* (2011) The diversity of life in African freshwaters: underwater, under threat. IUCN, Cambridge, UK and Gland, Switzerland.

De Graaf M, Sibbing F, Samallo J, Megens H-J (2007) Evolutionary origin of Lake Tana's (Ethiopia) small *Barbus* species: indications of rapid ecological divergence and speciation. *Animal Biology* **57**, 39-48.

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z-I, Knowler DJ, *et al.* (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**, 163-182.

Galewski T, Collen B, McRae L, Loh J, Grillas P, *et al.* (2011) Long-term trends in the abundance of Mediterranean wetland vertebrates: From global recovery to localized declines. *Biological Conservation* **144**, 1392-1399.

Golubstov A, Berendzen P (2005) A review of the small barbs (*Barbus*, Cyprinidae) with a serrated dorsal spine from Ethiopia, with a discussion of the origin of forms with a reduced dorsal spine from the White Nile and Omo River tributaries. *Journal of Ichthyology* **45**, S58.

Greenwood PH (1962) *A revision of certain Barbus species (Pisces, Cyprinidae) from East, Central and South Africa.* Bulletin of the British Museum (Natural History). Zoology.

Greenwood PH (1970) A revision of the cyprinid species *Barbus* (*Enteromius*) *radiatus* Peters 1853, with a note on the synonymy of the subgenera *Beirabarbus* and *Enteromius*. *Revue de Zoologie et de Botanique Africaines* **82**, 1–13.

Hecky RE, Bugenyi F (1992) Hydrology and chemistry of the African Great Lakes and water quality issues: problems and solutions. *Mitteilung Internationale Vereinigung fuer Theoretische und Amgewandte Limnologie*, 45-54.

Huelsenbeck JP, Ronquist F (2001) Mrbayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.

IUCN (2010) *African freshwater species threatened - livelihoods at stake.* https://www.iucn.org.

Katoh K, Toh H (2008) Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC bioinformatics* **9**, 212.

Kaufman L (1993) Evolutionary and conservation biology of cichlid fishes as revealed by faunal remnants in northern Lake Victoria. *Conservation Biology* **7**, 719-730.

Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC evolutionary biology* **14**, 1.

Leveque C, Daget L (1984) Cyprinidae. In J. Daget, J. P. Gosse & van den Audenaerde T. (Eds) Check List of the Freshwater Fishes of Africa, pp. 217–342, Paris: Tervuren & Orstom.

Mayden RL, Chen WJ, Bart HL, Doosey MH, Simons AM, *et al.* (2009) Reconstructing the phylogenetic relationships of the earth's most diverse clade of freshwater fishes--order Cypriniformes (Actinopterygii: Ostariophysi): a case study using multiple nuclear loci and the mitochondrial genome. *Molecular Phylogenetics and Evolution* **51**, 500-514.

Mayer WE, Tichy H, Klein J (1998) Phylogeny of African cichlid fishes as revealed by molecular markers. *Heredity* **80**, 702-714.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 November 2010, New Orleans, LA, pp. 1–8.

Mwita (2013) Phylogenetic relationships among small barbus, *Barbus* (actinopterygii: Cyprinidae) from malagarasi and Pangani river basins, Tanzania. *Natural Science* **05**, 1249-1254.

Mwita, Nkwengulila G (2008) Molecular phylogeny of the clariid fishes of Lake Victoria, Tanzania, inferred from cytochrome b DNA sequences. *Journal of Fish Biology* **73**, 1139-1148.

Ren Q, Mayden RL (2016) Molecular phylogeny and biogeography of African diploid barbs, '*Barbus*', and allies in Africa and Asia (Teleostei: Cypriniformes). *Zoologica Scripta* **45**, 642-649.

Roberts TR (1975) Geographical distribution of African freshwater fishes. *Zoological Journal of the Linnean Society* **57**, 249–319.

Salzburger W, Van Bocxlaer B, Cohen AS (2014) Ecology and Evolution of the African Great Lakes and Their Faunas. *Annual Review of Ecology, Evolution, and Systematics* **45**, 519-545.

Schmidt RC, Bart HL, Jr. (2015) Nomenclatural changes should not be based on equivocally supported phylogenies: Reply to Yang et al. 2015. *Molecular Phylogenetics and Evolution* **90**, 193-194.

Schmidt RC, Bart HL, Jr., Nyingi WD (2017) Multi-locus phylogeny reveals instances of mitochondrial introgression and unrecognized diversity in Kenyan barbs (Cyprininae: Smiliogastrini). *Molecular Phylogenetics and Evolution* **111**, 35-43.

Skelton PH (1988) *A taxonomic revision of the redfin minnows (Pisces, Cyprinidae) from southern Africa. Cape Provincial Museums.*

Skelton PH (1993) *Scientific and common names of southern African freshwater fishes*. *LB Smith Institute of Ichthyology* **56**, 34.

Skelton PH, Michael NB, Peter Brian Neville Jackson, Skelton PH (2012) *A Complete Guide to the Freshwater Fishes of Southern Africa*, First edn. Southern Book Publishers.

Skelton PH, Tweddle D, Jackson P (1991) Cyprinids of Africa. In: *Cyprinid Fishes,* pp. 211-239. Springer.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.

Stewart DJ (1977) Geographical distribution of *Barbus radiatus* Peters, a widely distributed African cyprinid fish. *Environmental Biology of Fishes* **1**, 13–125.

Sukumaran J, Holder MT (2010) DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**, 1569-1571.

Tsigenopoulos C, Kasapidis P, Berrebi P (2010) Phylogenetic relationships of hexaploid large-sized barbs (genus *Labeobarbus*, Cyprinidae) based on mtDNA data. *Molecular Phylogenetics and Evolution* **56**, 851-856.

Tsigenopoulos C, Ráb P, Naran D, Berrebi P (2002) Multiple origins of polyploidy in the phylogeny of southern African barbs (Cyprinidae) as inferred from mtDNA markers. *Heredity* **88**, 466-473.

Tumwebaze R, Cowx I, Ridgway S, Getabu A, MacLennan D (2007) Spatial and temporal changes in the distribution of *Rastrineobola argentea* in Lake Victoria. *Aquatic Ecosystem Health & Management* **10**, 398-406.

Uyeno T, Smith G (1974) Tetraploid origin of the karyotype of catostomid fishes. In: *Genetic Studies of Fish* (ed. Wourms JP), p. 112. Ardent Media, New York.

van Zwieten PA, Kolding J, Plank MJ, Hecky RE, Bridgeman TB, *et al.* (2015) The Nile perch invasion in Lake Victoria: cause or consequence of the haplochromine decline? *Canadian Journal of Fisheries and Aquatic Sciences* **73**, 622-643.

Vorosmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich A, *et al.* (2010) Global threats to human water security and river biodiversity. *Nature* **467**, 555-561.

Witte F, Goldschmidt T, Goudswaard PC, Ligtvoet W, Van Oijen MJP, *et al.* (1991) Species Extinction and Concomitant Ecological Changes in Lake Victoria. *Netherlands Journal of Zoology* **42**, 214-232.

Witte F, Seehausen O, Wanink JH, Kishe-Machumu MA, Rensing M, *et al.* (2013) Cichlid species diversity in naturally and anthropogenically turbid habitats of Lake Victoria, East Africa. *Aquatic Sciences* **75**, 169-183.

WWF (2014) *Living Planet Report 2014: species and spaces, people and places.*, WWF, Gland, Switzerland.

Yang L, Sado T, Vincent Hirt M, Pasco-Viel E, Arunachalam M, *et al.* (2015) Phylogeny and polyploidy: resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **85**, 97-116.

APPENDIX 1

Supplementary file figure legends

Figure 2. Cytochrome *b* bootstrap consensus tree (50% majority rule) of RaxML analyses. The "*Labeobarbus* and allies" clade has been collapsed (expanded in Figure 4). Bootstrap support values and Bayesian posterior probabilities are indicated next to the nodes. Analyses assuming a single data partition (for MrBayes: left value is GTR+G; right value is GTR+G+I) and three data partitions (= "best scheme") are indicated above the branches. Analyses assuming a single data partition (left) and 3 data partitions (right) for maximum likelihood are indicated below the branches. Boldfaced taxon labels are sequences generated in this study. Colors correspond to colors in other figures. Locality number indicated in taxon label within brackets for specimens from this study (see appendix 2) and those available from Schmidt *et al.* (2017). ^{GH} = Growth Hormone Intron 2 (GH) sequence was also obtained (shown in Figure 3). * = partial GH sequence was obtained (see text). [Supplementary file 1: RAxML_bootstrap Cytochrome b for small barbs.png]

Figure 3. Growth Hormone Intron 2 (GH) bootstrap consensus tree (50% majority rule) analyses based on maximum likelihood (RAXML) with a single data partition. Boldfaced taxon labels are sequences generated in this study. Bootstrap support values (numbers above branches) and Bayesian posterior probabilities (numbers below branches) are indicated next to nodes. Colors correspond to colors in other figures. Locality number indicated in taxon label within brackets for specimens from this study (see appendix 2) and those available from Schmidt *et al.* (2017). [Supplementary file 2: RAxML_bootstrap GH for small barbs.png]

Figure 4. Cytochrome *b* bootstrap consensus tree (50% majority rule) of RaxML analyses with a single data partition. The "African small barbs" clade has been collapsed (expanded in Figure 2). Bootstrap support values are indicated next to (most) nodes for analyses assuming a single data partition (left) and three data partitions (= "best scheme"; right). Boldfaced taxon labels are sequences generated in this study. Colors correspond to colors in other figures. Locality number indicated in taxon label within brackets. [Supplementary file 3: RAxML_bootstrap Cytochrome b for large barbs.png]

APPENDIX 2

List of samples used in this study, Locality numbers correspond to those in Figure 1.

Sample site/ locality	Locality number	Longitudes	Latitudes	Species collected	Specimen ID
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**BJ_2**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**Bspp_1**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**Bspp_2**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**BJ_8**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**BJ_9**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' jacksoni*	**BJ_12**
Yenga dam, Siaya, Kenya	15	34°12'39.21"E	0°12'54.78"N	*Barbus' jacksoni*	**Pld_1**
Yenga dam, Siaya, Kenya	15	34°12'39.21"E	0°12'54.78"N	*Barbus' jacksoni*	**Pld_2**
Yenga dam, Siaya, Kenya	15	34°12'39.21"E	0°12'54.78"N	*Barbus' jacksoni*	**Pld_3**
Yenga dam, Siaya, Kenya	15	34°12'39.21"E	0°12'54.78"N	*Barbus' jacksoni*	**Pld_4**
River Kisian , Kisumu, Kenya	9	34° 40.043'E	0° 4.274'S	*Barbus' cercops*	**KC3**
River AwachKendu, Homa Bay County, Kenya	5	34° 38.145'E	0° 22.817'S	*Barbus' cercops*	**AKC4**
River AwachKendu, Homa Bay County, Kenya	5	34° 38.145'E	0° 22.817'S	*Barbus' cercops*	**AKC3**
Kokech dam, Migori, Kenya	1	34° 26.748'	1° 5.746'S	*Barbus' cercops*	**KKC1**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' cercops*	**BCGH1**
Kokech dam, Migori Kenya	1	34° 26.748'	1° 5.746'S	*Barbus' cercops*	**KKDC3**
Kokech dam, Migori Kenya	1	34° 26.748'	1° 5.746'S	*Barbus' cercops*	**KKDC4**
River AwachKendu, Homa Bay County, Kenya	5	34° 38.145'E	0° 22.817'S	*Barbus' cercops*	**AKC2**

Ufinya dam, Siaya, Kenya	13	34° 16.874'E	0° 3.144' N	*Barbus' paludinosus*	**Pld1UFgh**
Ufinya dam, Siaya, Kenya	13	34° 16.874'E	0° 3.144' N	*Barbus' paludinosus*	**Pld2UFgh**
Ufinya dam, Siaya, Kenya	13	34° 16.874'E	0° 3.144' N	*Barbus' paludinosus*	**Pld3UFgh**
Uriri dam, Migori, Kenya	2	34° 30.842'E	0° 57.686'S	*Barbus' apleurogramma*	**Uraplr8**
Uriri dam, Migori, Kenya	2	34° 30.842'E	0° 57.686'S	*Barbus' apleurogramma*	**Uraplr9**
Uriri dam, Migori, Kenya	2	34° 30.842'E	0° 57.686'S	*Barbus' apleurogramma*	**Uraplr11**
River Kuja, Migori county, Kenya	3	34°20'41.78"E	0°54'54.63"S	*Barbus' neumayeri*	**Nmy2KJg**
River Kuja, Migori county, Kenya	3	34°20'41.78"E	0°54'54.63"S	*Barbus' neumayeri*	**Nmy3KJg**
Mauna dam, Siaya, Kenya	14	34° 9.445'E	0° 12.377'N	*Barbus' kerstenii*	**MN_apl2**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' kerstenii*	**krst_NW1**
Mauna dam, Siaya, Kenya	14	34° 9.445'E	0° 12.377'N	*Barbus' kerstenii*	**MN_apl3**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' kerstenii*	**Krst_NW3**
Mauna dam, Siaya, Kenya	14	34° 9.445'E	0° 12.377'N	*Barbus' kerstenii*	**MN_apl1**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' kerstenii*	**Krst_NW2**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' nyanzae*	**BNYZ_NW1**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' nyanzae*	**new_NZN2**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' nyanzae*	**ARC3**
Aquarium, KMFRI Kenya	Aquarium			*Barbus' nyanzae*	**AWRNZ5**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' nyanzae*	**ARC2**
River Awachrae, Kisumu County, Kenya	7	34° 26.748'E	1° 5.746'S	*Barbus' nyanzae*	**AWR_AMP27**
Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**Prof3gh**
Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**Prof2gh**
Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**DNG_PROF4 _4**

Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**DNG_PROF3 _3**
Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**DNG_PROF2 _2**
Dunga beach, Kisumu, Kenya	8	34°44'12.19"E	0° 8'41.20"S	*Barbus' profundus*	**DNG_PROF1 _1**

APPENDIX 3


E


F


G


H


I

Photographs showing phenotypic characteristics of '*Barbus*' specimens used in this study

A. Lateral view of '*Barbus' cercops* (Ethanol preserved) from river Kuja in Migori county, Lake Victoria drainage basin, Kenya
B. Lateral view of '*Barbus' kersternii* (ethanol preserved) from Kokech dam, Migori county, Lake Victoria drainage basin, Kenya
C. Lateral view of '*Barbus*' *nyanzae* (fresh specimen) from aquarium fish tank, Kenya Marine and Fisheries Institute, Kisumu, Kenya
D. Lateral view of '*Barbus*' *nyanzae* (Formalin preserved) from aquarium fish tank, Kenya Marine and Fisheries Institute, Kisumu, Kenya
E. Lateral view of '*Barbus*' *apleurogramma* (Formalin preserved) from Uriri dam in Migori county, Lake Victoria drainage basin
F. Lateral view of '*Barbus' neumayeri* (ethanol preserved) from river Kuja in Migori county, Lake Victoria drainage basin
G. Lateral view of '*Barbus' jacksoni* (ethanol preserved) from river Kuja in Migori county, Lake Victoria drainage basin
H. Lateral view of *'Barbus*' *paludinosus* (fresh specimen) from Kenya Marine and Fisheries Institute, Kisumu, Kenya
I. Lateral view of *'Labeobarbus*' *altianalis* (fresh specimen) from river Nzoia Siaya county, Lake Victoria drainage basin, Kenya