



**This electronic thesis or dissertation has been
downloaded from Explore Bristol Research,
<http://research-information.bristol.ac.uk>**

Author:
Ward, Sophie

Title:
**Morphological and Molecular Approaches for Assessing Diversity of East African
Freshwater Fishes**

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode> This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

MORPHOLOGICAL AND MOLECULAR APPROACHES FOR ASSESSING DIVERSITY OF EAST AFRICAN FRESHWATER FISHES

Sophie Ward



A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Masters by Research in the Faculty of Science.

July 2020

Student Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

S.Ward

Date: 31/07/20

Candidate Number: 37577

Supervisors:

Martin Genner, University of Bristol.

Christos Ioannou, University of Bristol.

THESIS ABSTRACT

Conservation and management of freshwater fish in Africa can benefit from improved knowledge of the diversity of species present. This thesis focusses on methods to delimit species using morphological measurements and DNA barcodes. In Chapter 1, I discuss known levels of freshwater fish biodiversity in Africa, threats to African biodiversity, the existence and complications of cryptic species, the current knowledge surrounding freshwater fish biodiversity in Tanzania, catfish biodiversity, and the use of DNA barcoding data in systematics. In Chapter 2, I present a study investigating the use of molecular data in taxonomy, specifically DNA barcoding, in studying Tanzania's freshwater fish biodiversity. Two different DNA markers (mitochondrial COI and 12S) were used to assess diversity among different barcoding genes and evaluate their utility in separating distinct taxa using neighbour-joining phylogenies. Both markers were equally successful in delimiting morphologically-identified specimens from the Ruaha-Rufiji-Kilombero river catchment into species groups. However, comparisons of these sequences to global reference sequence databases highlighted the absence of publically-available validated reference data for the species included in our study. In Chapter 3, I present a study investigating the use of linear morphological measurement in species identification, focussing on *Chiloglanis* catfish from South-Eastern Africa. The results support the use of linear measurement data in species delimitation, but also highlight the presence of genetically-distinct yet morphological similar "cryptic" species. Overall, the findings suggest morphological diversity within *Chiloglanis* has arisen from stabilising selection on the overall fluvial phenotype, potentially coupled with adaptation to specific characteristics of local river systems. In Chapter 4, I discuss the implications of these results for future freshwater ichthyofaunal research in Africa.

Acknowledgements

I would like to give my sincere thanks to many people for their support and involvement during my Masters by Research degree at the University of Bristol.

First to my supervisor, Professor Martin Genner, for his tremendous help and support throughout my project, especially in the transition to remote working under unprecedented circumstances.

I would also like to thank fellow lab colleagues who got involved and supported me in this project, especially Dr Rupert Collins who aided in my laboratory training and morphological identification of specimens.

Thanks also to Dr Chris Clements for his involvement in my six-month assessment and constructive review in the midpoint of my project.

Finally, I would like to acknowledge and thank all those who supported me through my academic development, specifically my family and the friends whom I met during my time in Bristol.

Contents

Student Declaration.....	i
Thesis Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
List of Appendices.....	ix

Chapter One

1.1 African freshwater biodiversity.....	1
1.2 Threats to African freshwater biodiversity	
1.2.1 Eutrophication.....	3
1.2.2 Climate change.....	4
1.2.3 Invasive species.....	5
1.3 Freshwater ichthyofauna of Tanzania.....	5
1.4 Catfish biodiversity.....	6

1.5 Morphological diversity of African freshwater fishes	
1.5.1 Sympatry.....	7
1.5.2 Allopatry.....	8
1.5.3 Cryptic species.....	9
1.6 DNA Barcoding and molecular diversity	
1.6.1 Background.....	10
1.6.2 Barcoding African freshwater fishes.....	11
1.6.3 Choice of DNA barcode.....	12
1.7 Morphological vs molecular data in fish phylogenetics.....	13

Chapter Two

2.1 Abstract.....	15
2.2 Introduction.....	16
2.3 Materials and methods.....	18
2.4 Results.....	23
2.5 Discussion.....	27

Chapter Three

3.1 Abstract.....	33
-------------------	----

3.2 Introduction.....	34
3.3 Materials and methods.....	38
3.4 Results.....	40
3.5 Discussion.....	47
 Chapter Four	
4.1 DNA barcoding for resolution of East African freshwater fish biodiversity.....	51
4.2 The future of Chiloglanis study.....	52
4.3 Concluding remarks.....	55
References.....	56

List of Tables

Table 2.1. Collection information for specimens.....	20
Table 2.2. DNA barcode genetic distances.....	24
Table 3.1. Summary details of morphological samples.....	37
Table 3.2. Genetic clades recovered by site.....	39
Table 3.3. Morphometric measurements for <i>Chiloglanis</i> specimens.....	41
Table 3.4. Logged residual measurements from sympatric clades.....	44
Table 3.5. Results of Anova (Type II).....	47

List of Figures

Figure 2.1. Map detailing the construction site of the Stiegler's Gorge.....	17
Figure 2.2. COI neighbour-joining tree phylogeny.....	25
Figure 2.3. 12S neighbour-joining tree phylogeny.....	26
Figure 2.4. Histograms of genetic distance for DNA barcodes.....	27
Figure 3.1. Maximum likelihood alongside representative specimens.....	36
Figure 3.2. Principal Component Analysis of all specimens.....	40
Figure 3.3. Principal Component Analyses of sympatric clades.....	43
Figure 3.4. Principal Component Analyses of allopatric populations.....	45
Figure 3.5. Principal Component Analyses of allopatric populations with overlaps.....	46
Figure 4.1. Diagram of a generalized <i>Amphilius</i> specimen.....	54

List of Appendices

Appendix 2.1. DNA barcoding database search results.....	67
Appendix 3.1. Collection sites for <i>Chiloglanis</i> specimens.....	70
Appendix 3.2. Illustration depicting the linear morphometric measurements.....	71

CHAPTER ONE: GENERAL INTRODUCTION.

1.1 African freshwater biodiversity

It is estimated that around 42% of the world's 32,500 fish species live in freshwater habitats, despite these ecosystems occupying only 0.8% of the Earth's surface (Darwall *et al.*, 2005; Nelson, 2006). Africa has at least 3,300 freshwater species, representing one of the most unique and diverse ichthyofaunas globally. Its species richness is only slightly less than that of Asia (~3,600) and South America (~4,200) (Nyboer *et al.*, 2019). Complex geological processes including uplift, volcanic activity, the formation of deserts and tectonic rifting led to substantial biogeographical barriers, and greatly shaped the genetic diversity of Africa's ichthyofauna and connectivity of populations (Basiita *et al.*, 2018). Around one sixth of these African freshwater species (~580) have previously been described as endangered or vulnerable to extinction by assessments in conjunction with the IUCN Red List (Nyboer *et al.*, 2019).

Among Africa's freshwater fish, Cichlidae, from the order Cichliformes, are by far the most species-rich group within overall African freshwaters, recorded as containing around 870 species in 1991, with estimates growing tremendously following recent molecular analyses (Lévêque, 1997), and it has been estimated that Africa may contain over 2000 cichlid species (Salzburger, 2018). However, most of these cichlid species are lacustrine, and thus species richness of African rivers instead tends to be dominated by members of the families Cyprinidae, Characidae and Siluriformes. The superorder of these fish, Ostariophysi, is thought to contain 68% of all freshwater species with around 48 families in tropical and southern Africa, of which 15 are endemic (Lévêque, 1997; Lévêque and Paugy, 2017).

Many of the fish species of Africa are regional endemics belonging to recognised adaptive radiations. The most famous and diverse of these radiations is that of the cichlid fishes in the rift valley lakes of East Africa (Nyboer *et al.*, 2019). Lake Malawi, the third largest lake in Africa, is reported to contain at least 800 species (Salzburger, 2018), which is more species of fish than any other lake in the world, with numbers increasing during each intensive survey (Ribbink, 2001; Weyl *et al.*, 2010). African freshwater lakes, such as Lake Malawi, are of special interest to evolutionary biologists, due to rapid speciation within species flocks, with strong divergent natural and sexual selection thought to be the major driving force behind their diversity (Plenderleith *et al.*, 2005). Several other species radiations in lakes have gained the attention of scientists in Africa, including those found in the Upper and Lower

Guinean ecoregions and the Congo River basin (Nyboer *et al.*, 2019). Additionally, a high diversity of fish are found in rivers, streams and small lakes. These systems are highly fragmented and isolated, potentially contributing to allopatric speciation and adaptive radiation (Seehausen and Wagner, 2014). This demonstrates that Africa's diverse freshwater ichthyofauna is of extreme importance, not only in its value to biodiversity but also in fuelling our understanding of evolutionary processes (Sayer *et al.*, 2019).

Additionally, the freshwater fishes of Africa also hold valuable socio-economic and cultural roles in African societies. Many species are harvested in commercial or artisanal fisheries, however some are also used in the aquarium trade (Nyboer *et al.*, 2019). Global economic value of inland freshwater fisheries catches is estimated to be approximately \$26 billion USD, with Africa a major contributor at around 22.2% of the total (Thorpe and Castillo, 2018).

1.2 Threats to African freshwater biodiversity

Several factors threaten the freshwater biodiversity of Africa. These include a number of anthropogenic threats to irreplaceable freshwater ecosystems, such as habitat loss, water extraction leading to hydrological disruption, the introduction of invasive species, pollution, and overexploitation of resources (Darwall *et al.*, 2011). Many threatened species are subject to several of these interacting stressors (Darwall *et al.*, 2005). Sayer *et al.* (2018) assessed the global extinction risk to 651 freshwater species native to East Africa's Lake Victoria (204 of which are endemic), including fishes, molluscs, and aquatic plants. They found that 20% are threatened with extinction, with 78% of endemics at risk.

Inland waters and the freshwater biodiversity that they contain are an extremely valuable natural resource, in terms of economic, cultural, aesthetic, and scientific uses (Dudgeon *et al.*, 2005). In order to fully protect these critical habitats, areas must be identified and declared as conservation reserves to mitigate loss of species diversity (Sarkar *et al.*, 2011). However, due to the extensive uses for freshwater systems, trade-offs between conservation of biodiversity and the use of ecosystem goods by humans may be required (Dudgeon *et al.*, 2005).

1.2.1 Eutrophication

Eutrophication, the excessive release of nutrients such as nitrogen and phosphorus into freshwaters from sewage and industries in urban areas, is commonly seen throughout Africa, with estimates suggesting that 28% of all lakes and reservoirs are eutrophic (Pollard *et al.*, 2008). Effects of eutrophication include the increase in phytoplankton, blooms of macrophytes and ultimate deoxygenation of the water column leading to the eradication of fish species. Following the loss of fish species, there can be shifts in food chains, impairment of fisheries production, and possible outbreaks of invasive species (Soesbergen *et al.*, 2019).

Otu *et al.* (2011) investigated long-term eutrophic effects in Africa's Lake Malawi and its subsequent change in water quality. Results indicated that Lake Malawi has experienced increased nutrients since 1940 associated with phosphorus runoff from land use. By 1980 the dominant diatom types in the lake had been replaced by taxa that prefer higher phosphorus levels relative to silica. Similar changes were observed in Lake Victoria, the largest of Africa's Great Lakes, which has suffered considerable biodiversity loss due to eutrophication (Otu *et al.*, 2011). It is thought to contain severe levels of eutrophication mainly due to agriculture, which supplies 22% of nitrogen and 50% of phosphorus loading (Soebergen *et al.*, 2019).

South Africa's most limiting natural resource is freshwater, thus making rivers and lakes pivotal for the provision of safe and sufficient drinking water (Van Ginkel, 2011). However, the growing human population of Africa has had a large effect on the eutrophication and nutrient pollution of freshwater systems, with many small towns without proper sewerage systems (Soesbergen *et al.*, 2019). With the possibility for decreased water quality to trigger large impacts on human and ecosystem health, mitigation following land use has been highlighted as a potential method to protect both water quality and subsequent biodiversity (Otu *et al.*, 2011). Van Ginkel (2011) highlights the importance of implementing adaptive management systems to monitor and handle eutrophication. This includes the testing and application of the appropriate methodology to monitor and reverse eutrophication, with continued monitoring following intervention. Management options include the treatment or re-use of wastewater, increased regulation of polluting industries, the improvement of agricultural systems, and the restoration of floodplain wetland (Tickner *et al.*, 2020).

1.2.2 Climate change

Exacerbation of eutrophication of river systems can occur following changes in precipitation patterns as a result of climate change and the alteration of hydrological cycles and increased discharge into lake and river systems (Nyboer *et al.*, 2019). Likely trends predicted by global climate change models include changes in the magnitude and amplitude of both rainfall and air temperature, which in turn can result in changes in water availability (Dallas and Rivers-Moore, 2014). Studies into the scale of impact of these changes suggest that climate change could eventually surpass habitat destruction in becoming the greatest global threat to biodiversity over the coming decades (Leadley, 2010). Climate change threatens almost 40% of African freshwater fishes. Given the threat of climate change as a key factor in driving biodiversity loss across all global ecosystems, freshwater fishes are thought to be among the most vulnerable taxa owing to many species being highly specialised in their habitat and life-history requirements (Nyboer *et al.*, 2019).

All physiological functions of freshwater fish are determined by surrounding water temperature, however the main factors include growth, food consumption and metabolic rate, reproductive success, and the ability to successfully maintain a stable internal homeostasis in the face of a variable external environment (Fry 1971; Ficke *et al.*, 2007). Therefore, possible impacts of thermal changes may involve increases or decreases in local species abundance, or species-level range changes. Those species that are endemic or reside in fragmented habitats will be more at risk as they are less able to follow changing thermal isolines over time (Ficke *et al.*, 2007).

Loss of fish populations will not only affect overall biodiversity levels across Africa, but also cause inevitable damage to the fisheries sector for which 12.3 million people rely on for income, representing 2.1% of Africa's population (de Graaf and Garibaldi, 2014). It is predicted that by 2050 (under a high CO₂ emissions scenario), climate change will have reduced fish catches by 7.7% and subsequent revenues by 10.4% globally (Lam *et al.*, 2016). Climate change impacts on African lakes have already been observed, including Lakes Tanganyika, Malawi, Kivu and Victoria (Marshall, 2012), however the broader impacts of climate change on inland fisheries resources within Africa is unclear, in part because of a lack of reliable environmental data (Marshall, 2012).

1.2.3 Invasive species

Invasive species can shift the nature of ecosystems through ecological impacts such as the potential for predation, competition among species, habitat alteration, and the introduction of novel diseases. Freshwater ecosystems are considered especially vulnerable to invasion, and the introduction of invasive alien plant and animal species can have devastating effects (Darwall *et al.*, 2011; Bradbeer *et al.*, 2018). In Africa, a major threat comes from the spread of cultured fish species into non-native ecosystems, linked to sustained expansion of aquaculture. This is partly due to the overexploitation of many major capture fisheries, which in turn leaves little capability for existing fisheries to expand through other methods, such as technological innovation (Shechonge *et al.*, 2018). Tilapiine cichlids are among the most widely cultured groups of freshwater fish species, having been introduced to over 140 countries worldwide (Shechonge *et al.*, 2019).

Species are introduced to habitats for which they are non-native in order to develop aquaculture and capture fisheries. Previous species include the Nile tilapia (*Oreochromis niloticus*) and blue-spotted tilapia (*Oreochromis leucostictus*), both of which have been widely introduced across East Africa (Bradbeer *et al.*, 2018). The translocation of these *Oreochromis* species has led to hybridisation with native species, such as the Wami tilapia (*O. urolepis*), a species that is present in coastal rivers of Tanzania. This hybridisation has led to the loss of indigenous unique genetic diversity, and there is potential for extinction of all pure populations of the *O. urolepis* (Bradbeer *et al.*, 2018).

It is thought that the best approach to limiting these impacts is to prevent the introduction of non-native invasive species altogether, through increased biosecurity. Once these species become established in a habitat, control and eradication may only be possible through physical removal, chemical treatment or biological control (Tickner *et al.*, 2020).

1.3 Freshwater ichthyofauna of Tanzania

The mainland of Tanzania (885,800 km²) is located along the Indian Ocean coastline, with neighbouring Kenya to the north and Mozambique to the south. The African Great Lakes, Victoria, Tanganyika, and Malawi, cross through Tanzania's northern, western and southern borders, and are some of the world's largest inland water bodies (Seeteram *et al.*, 2019; Tanzanian Fisheries Sector, 2016). Major rivers in Tanzania include the Pangani, Wami, Ruvu,

Rufiji and Ruvuma rivers, that flow eastward and eventually drain into the Indian Ocean (Seeteram *et al.*, 2019).

As of April 2020, there were 1848 described fish species reported from Tanzania, including 877 freshwater species in its inland rivers and lakes, comprising a total area of 64,000km² (Tanzanian Fisheries Sector, 2016; Froese and Pauly, 2019). Most of these freshwater fish species are present in Tanzania's Rift Valley Great Lakes, which have been the focus of many more studies than Tanzania's rivers (Seeteram *et al.*, 2019). Despite its diverse freshwater ichthyofauna, scientific study of Tanzanian riverine biodiversity has been greatly limited (Seeteram *et al.*, 2019). This is a consequence of a lack of taxonomic expertise, a lack of funding for taxonomic research in Africa, and well as limited facilities for hosting biodiversity collections (Swartz *et al.*, 2008; Skelton & Swartz, 2011).

The lack of understanding regarding fish species richness and unrecognised endemism among species in these habitats can hamper efforts to manage biodiversity. The "Field Guide to the Freshwater Fishes of Tanzania" (Eccles 1992) remains the primary source of information regarding the identification of most of the fishes from outside the Great Lakes (Seeteram *et al.*, 2019). However for some groups the utility of the guide is very limited, for example the small barb species now assigned to the genus *Enteromius* (Utzinger and Charlwood, 1996). If Tanzania's freshwater fish fauna are to be protected from imminent threats such as invasive species, habitat destruction and river pollution, it is vital that its ecosystems are extensively sampled, and species distributions more fully established.

In 2014, inland fisheries accounted for about 85% of the national fish production in Tanzania. Lake Victoria and Lake Tanganyika are the most important lakes for fishery, both responsible for around 94% of total inland fish production each year (Tanzanian Fisheries Sector, 2016). The Rufiji River of Tanzania also plays an important role in the fisheries industry, with most riverine fishing activities taking place in its major tributaries, the Ruaha and Kilombero Rivers (Bwathondi and Mwamsojo, 1993). The Nile perch (*Lates niloticus*), Nile tilapia (*Oreochromis niloticus*), and the freshwater sardine (*Rastrineobola argentea*) are the main freshwater species of commercial interest in Tanzania (Tanzanian Fisheries Sector, 2016).

1.4 Catfish biodiversity

Catfish species constitute one third of the global freshwater fish fauna. They are an extremely diverse group of ray-finned fish, representing more than 3,000 species falling under 478 genera

and 36 families (Lalronunga *et al.*, 2011). Extant catfish species live within the freshwater or coastal waters of every continent except Antarctica, with most catfish operating as bottom feeders (Nelson, 2006).

Due to their worldwide distribution and large number of species, catfishes have been the subject of a grant of \$4.7 million from The National Science Foundation to the Florida Museum of Natural History (University of Florida) and four other institutions in 2004, aiming to discover and describe all catfish species worldwide. Recent rates of the description of new catfish species have been greater than ever before, with over 100 species named between 2003 and 2005, a rate three times faster than that of the previous century (Ferraris and Reis, 2005). However, catfish are not only significant in their role in global biodiversity, but many of the larger species are of considerable commercial importance, usually farmed or fished for food. The three largest documented extant species of catfish include the Mekong giant catfish (*Pangasianodon gigas*) of Southeast Asia, the wels catfish (*Silurus glanis*) of Eurasia, and the pirariba (*Brachyplatystoma filamentosum*) of South America.

1.5 Morphological diversity of African freshwater fishes

The morphological phenotype of a species is greatly influenced by its ecology and the physical characters of local habitats. In freshwater fish, functional morphology is strongly affected by environmental characteristics of the water bodies (e.g. flow velocity, oxygen content of water, light intensity) that interact with locomotion and food acquisition (Shuai *et al.*, 2018). Fish fauna of river ecosystems in particular can display a high level of morphological diversity, linked to a large array of available niches within the system. For example, fishes found in fast-flowing streams tend to exhibit a streamlined morphology to reduce drag and increase their ability to cope with prolonged, steady swimming against currents (Løstrem *et al.*, 2015). By contrast, those found in slow flowing rivers and floodplains are thought to develop deeper bodies that can maximize thrust and stability when swimming in bursts (Langerhans, 2008).

1.5.1 Sympatric Speciation

Divergent selection can result in the formation of species in the absence of geographical barriers (sympatry), often as a result of disruptive selection that can arise from intraspecific competition for resources. If genetically determined, the differential use of microhabitats and

variation in relevant traits can result in assortative mating, leading to reproductive isolation and subsequent formation of separate species (Kautt *et al.*, 2016). Sympatric speciation has been suggested to take place among many groups of freshwater fishes. For example, in Nicaragua the crater Lake Apoyo is thought to have been seeded only once by the cichlid fish *Amphilophus citrinellus*, leading to the evolution of a flock of six endemic species (Recknagel *et al.*, 2014) that are hypothesised to have diverged within less than 10,000 years (Barluenga *et al.*, 2006).

In areas of habitat where ecological opportunity is present, it is common for species to diverge and fill available niches to avoid interspecific competition through niche partitioning. Kautt *et al.* (2018) found that the depth of a crater lakes is positively associated with variation in body shape and the number of locally-evolved taxa, presumably by providing a larger number of ecological opportunities. Competition for resources is often quoted as a main factor affecting divergence between groups of freshwater fish living in benthic and open-water limnetic habitats, which is seen in fish taxa of both temperate and tropical environments (Kautt *et al.*, 2016). It is possible that greater morphological variation can be seen among sympatric populations of species than allopatric populations, due to interspecific competition (Adams *et al.*, 2007).

1.5.2 Allopatric Speciation

Allopatric populations of species can diverge morphologically to adapt to suit the specific characteristics of their local environment. Adaptive allopatric divergence occurs when natural selection acts differently on traits in relation to environmental conditions, resulting in individuals that have a higher fitness level in their local environment than individuals from other habitats (Williams, 1966). Eventually, prolonged geographic isolation of these populations may lead to the formation of separate species that are no longer able to interbreed, with partial reproductive isolation resulting as a by-product of local adaptation and lack of gene flow (Surget-Groba *et al.*, 2012).

This allopatric model of speciation is widely recognised, and populations may develop genetic or physical incompatibilities through genetic drift, adaptation to differing habitats, by sexual selection, or by the accumulation of incompatible mutations in mutual adaptation to similar habitats (Surget-Groba *et al.*, 2012). For instance, shallow water cichlid fishes of the African Great Lakes have been found to show localised divergence in male colouration among allopatric populations. This results in preferential mate choices in laboratory mate-choice trials,

with females choosing to mate with males from their own population and earning them an incipient species status (Knight and Turner, 2004).

1.5.3 Cryptic species

The term “cryptic species” was first defined by Stebbins (1950) as “population systems which were believed to belong to the same species until genetic evidence showed the existence of isolating mechanisms separating them”. Typically, these species exhibit low morphological but high molecular disparity. Studies into this phenomenon in Lepidoptera suggest that 10-20% of morphologically based “single” species would in fact be classified as two or more distinct species through DNA barcoding methods (Janzen *et al.*, 2017). Stabilizing selection has been commonly referred to as the main force maintaining similarity in species complexes, which may act to preserve ecologically-important phenotypic traits that promote survivorship, such as camouflage. While divergent selection often leads to sister species with notably different morphologies, in circumstances where cryptic species are inferred, it may only act on traits causing reproductive isolation (e.g. behaviour) (Rosser *et al.*, 2018). However, in some cases, the evolution of cryptic species may simply result from recently-diverged sibling species in which morphological differences have not yet accumulated (Florio *et al.*, 2012).

The underestimation of biodiversity is a big issue surrounding conservation efforts, often leaving large numbers of species unprotected and at risk (Palandačić *et al.*, 2017). Pessimism surrounding species richness in an area can lead to inaccurate ecological status and inefficient management methods (Chenuil *et al.*, 2019). It is possible for single “species” to be considered “least concern” when they possess a wide geographical range, but if that “species” is actually multiple endemic taxa each with a restricted distribution, these may deserve higher levels of conservation recognition and protection (Morris *et al.*, 2015). The largest contributing factors to undiscovered cryptic diversity are the lack of distinguishing features in morphological study, and absence of relevant molecular data (Jirsová *et al.*, 2019).

Both molecular and morphological data must be combined to study cryptic species, and the large number of cryptic species discovered only through additional molecular study suggests that traditional morphological techniques may in some cases be insufficient for taxonomic research due to low resolution (Lajus *et al.*, 2015). However morphological data must still be gathered, as often morphological traits sufficient for distinct identification may simply be overlooked (Lajus *et al.*, 2015). For example, as recent as 2001, the African elephant was

classified as two distinct non-interbreeding species, the African bush elephant (*Loxodonta africana*) and the smaller African forest elephant (*Loxodonta cyclotis*) on the basis of both molecular and morphological differences (Roca, 2001). In catfishes, Cooke *et al.* (2012) unearthed five deeply divergent cryptic lineages among the Amazonian catfish *Centromochlus existimatus* using biogeographic predictions and genetic data. The cause of this apparent lack of morphological divergence between these cryptic lineages was speculated to be a result of contrasting hydrochemistry among sites, and the presence of non-visual mating cues (Cooke *et al.*, 2012).

Cryptic species may be more common in tropical regions than temperate regions (Bickford *et al.*, 2007). Hence, it is possible that there is a high diversity of cryptic species in relatively poorly studied species groups, such as the stream fishes of sub-Saharan Africa (Kadye and Moyo, 2008). If there are high levels of unrecognised endemism, this may make stream fishes among the most threatened groups of vertebrates worldwide due to multiple human impacts on these systems (Collares-Pereira and Cowx, 2004). Impacts from pollution events, water abstraction and dam construction all have the potential to lead to the decline in stream-specialist taxa, such as cyprinids and catfishes (e.g. *Chiloglanis* spp. and *Zaireichthys* spp.) (Kadye and Moyo, 2008). With catfishes comprising 32% of freshwater fish diversity globally, it is vital that the biodiversity among these species is fully documented and understood in order to minimise the threat to their survival (Cooke *et al.*, 2012).

1.6 DNA Barcoding and molecular diversity

1.6.1 Background

The classification of organisms is a crucial step in managing the biodiversity, breeding, conservation, and development of animal populations (Vu and Le, 2019). DNA barcoding in ecology is a common tool used for biodiversity inventories (Bhattacharya *et al.*, 2015). The main steps include i) DNA extraction, ii) polymerase chain reaction (PCR) amplifying a “barcode” gene, iii) sequencing of the barcode amplicon, iv) comparison of the derived sequence with those available in global bioidentification databases (Ward *et al.*, 2005). Hebert *et al.* (2003) first suggested that a single gene sequence could be sufficient in differentiating the majority of animal species on Earth. They proposed the mitochondrial DNA cytochrome-c oxidase subunit (COI) gene as most suitable, and this marker is still widely used as the mainstay of animal DNA barcoding initiatives, with reference sequences available on online DNA

sequence libraries such as on the Barcode of Life Data System (BOLD; <https://www.boldsystems.org/>) or GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>).

For identification to be successful, sequence comparisons within-species must have greater similarity to each other than when compared to sequences from other species. Since mitochondrial DNA is typically maternally inherited and non-recombining, then hybridisation events among species can lead the two species to both share the same barcode, potentially leading to incorrect identification (Ward *et al.*, 2005). Moreover, it is possible for very closely related to species to have insufficient divergence in their sequences, preventing reliable species identification (Bhattacharya *et al.*, 2015). Hubert *et al.* (2008) were able to distinguish 93% of 190 Canadian freshwater fishes using DNA barcodes, with recent speciation or introgressive hybridization proposed as explanations for the inability for the COI barcodes to distinguish the remaining species. In such cases where identification is found to be inaccurate or require further clarification, it may require sequence analysis for other genes, such as the more rapidly evolving mtDNA control region, or the use of other molecular markers (Bhattacharya *et al.*, 2015).

Despite some limitations, DNA barcoding has become a prime resource in species identification and discovery. For example, Bingpeng *et al.* (2018) demonstrated the applicability of the COI gene for the identification of fish at the species level in the Taiwan Strait. They confirmed high efficiency of species identification in their study by DNA barcoding and concluded that COI sequencing is successful in differentiating fish species. It has also been found that DNA barcoding techniques can reveal hidden diversity within species and provide more accurate species richness estimates for ecosystems. Additionally, DNA barcoding has found applied use as a tool for testing accuracy of fish labelling in fisheries and restaurants. Using the COI barcode, Staffen *et al.* (2017) diagnosed 30% of samples as mislabelled in fisheries, and 26% of samples as mislabelled in restaurants.

1.6.2 Barcoding African freshwater fishes

Back in 2011, barcode coverage on a continental scale was reported as almost complete for North American freshwater fishes (April *et al.*, 2011). In stark contrast, Africa's freshwater fish biodiversity is poorly understood, with taxonomic keys rarely revised and subsequently proving inadequate to identify specimens to the species level (Swartz *et al.*, 2008). The Fish Barcode of Life Initiative (FISH-BOL) was launched in 2005, a global research project with

the ultimate goal of collecting and standardising DNA barcode sequences for use in the molecular identification of all fish species. The project aimed to produce a fast, precise, and cost-effective method of molecular identification for the world's ichthyofauna. This project seemed promising in solving Africa's taxonomic issues and offered hope that the extensive biodiversity could be documented. However, a status report published in 2011 found that only 13.5% of African fish species were recorded as barcoded by FISH-BOL (Becker *et al.*, 2011), with a strong bias towards marine species. Thus, recommendations were made that sampling in the following years should focus on the collection of freshwater species in Africa. However, in recent years, DNA barcoding efforts in Africa and biodiversity-rich regions such as Tanzania have been few and far between. The absence of comprehensive DNA data for many African lineages limits our understanding of how these fishes might be affected by increasing threats to freshwater habitats, and the ongoing extinction risk (Adeoba *et al.* 2019).

1.6.3 Choice of DNA barcode

A gene region must satisfy three criteria in order to be suitable for use as a DNA barcode. The region must contain significant species-level genetic variability and divergence, possess conserved flanking sites that can be used in the development of PCR primers for universal use, and have a short sequence length to fit current capabilities of DNA barcoding methods (Kress and Erickson, 2008). Although the COI gene is the standard marker commonly used for animal DNA barcoding, its use is limited in identifying some species, and fragments may not always efficiently PCR amplify using universal primers for some taxa (Ly *et al.*, 2014). The inability for COI to work as a barcode in plants also required botanists to find an alternative marker (Kress and Erickson, 2008).

Choice of the most suitable region for barcoding varies between species groups. In the identification of hydrozoan species, studies suggest that the 16S rRNA gene is a 'better' marker for discriminating taxa at the genus level, and in some cases at the family level (Zheng *et al.*, 2014). Use of the 16S and 12S rRNA genes have also been investigated as potential universal markers for the identification of commercial fish species. Cawthorn *et al.* (2012) found that both the 16S and 12S sequences in fish contained nucleotide variation to make them suitable for identifying the majority of the examined specimens from South Africa.

Ly *et al.* (2014) investigated the rates of correct sequence identification for the markers COI, 16S rDNA, ITS2 and 12S rDNA. All four were found to have very high (>96%) accuracy,

indicating that COI was not significantly more accurate than other markers in its rate of correct sequence identification, in ticks at least. This suggests that COI does not necessarily have to be the first choice in animal species identification. This is important given the development of new high-throughput environmental DNA metabarcoding approaches for quantifying community composition. Studies in fish, for example, have demonstrated that eDNA primers developed for the 12S region of fish have much high taxonomic fidelity than equivalent eDNA primers developed for COI (Collins *et al.*, 2019). Thus, the use of 12S primers results in a much higher proportion of fish reads in eDNA metabarcode data than is possible with COI (Collins *et al.*, 2019). Therefore, it may be valuable to barcode individuals at both COI and 12S genes to enable the use of barcode data for both identification of sampled specimens, and the successful assignment of metabarcode eDNA reads to species.

1.7 Morphological vs molecular data in fish phylogenetics

Traditionally, morphology was the main source of data employed when formally describing taxa. Most of what we know today about systematics and classification is founded on morphological data, using methods such as linear measurements of the external anatomy of specimens (Scotland *et al.*, 2003). These methods have now evolved into image-based geometric morphometric approaches, that use landmarks to study variations in shape (Farré *et al.*, 2016). With increasing accessibility of DNA sequences for phylogenetic reconstruction of extant taxa, phenotypic data is becoming increasingly marginalised for phylogenetic inference, coming under scrutiny for its ambiguity and difficulty to model (Scotland *et al.*, 2003). Alongside the development of bioinformatics tools have developed, DNA barcode data have become increasingly important for identification and systematics. Much of this is because of the ease with which DNA sequence data can be compared to reference sequences on online DNA sequence libraries, such as on the Barcode of Life Data System (BOLD) or GenBank.

It has been stated that modern morphologists need to improve their methods and use phenotypic data in a fundamentally different way if morphology is to remain to play an integral part in systematic research (Lee and Palci, 2015). Moving forward, undoubtedly morphological evidence will remain a key first step for delimiting potentially distinct taxa. However, genetic information will largely replace morphological data for generation of meaningful phylogenies of extant taxa. Subsequently, morphological, behavioural and ecological traits can be associated with the structure of a molecular phylogeny. This can enable identification of

phenotypic novelties that characterise molecular groupings (Lee and Palci, 2015), as well uncovering hidden levels of biodiversity which are often highly underestimated as a result of phenotypic similarities among reproductively-isolated species (Palandačić *et al.*, 2017).

Morphology is thought to be poor in the discrimination among cryptic species, and insufficient diagnostic characters in larval fish often results in misidentification and difficulty to key to the genus or species level and inconsistencies between different laboratories (Ko *et al.*, 2013). Instead, DNA barcoding has been found to be a powerful tool in detecting cryptic species diversity and revealing taxonomic inconsistencies in many taxonomic groups (Decru *et al.*, 2015). However, in cases of incomplete lineage sorting or introgressive hybridisation, clustering in the same genetic lineage can occur in single-locus phylogenies of morphologically distinct species, as found in COI-based studies of fish from the north-eastern Congo basin (Decru *et al.*, 2015). If DNA barcoding is thus used on samples that have not yet been morphologically examined, species richness could be vastly underestimated. This suggests that the combination of morphological and molecular data will allow the best delimitation of species and increase reliability of the study of African ichthyodiversity.

Studies in recent years have highlighted the value of combining genetic and morphological data in taxonomy, and this “total evidence” approach can be helpful when genetic data is limited in quantity and/or cases where phenotypic traits are known to associate strongly with phylogenetic history (Palandačić *et al.*, 2017). In particular, such combined approaches are useful for identifying cases where phylogenetic inference from sequence data misalign with phenotypic data (Stepanović *et al.*, 2015). Such discrepancies can arise from cases where individual gene trees based generate misleading phylogenetic hypotheses, due to prevalent incomplete lineage sorting and/or historic hybridisation events (Palandačić *et al.*, 2017). An improved understanding of the prevalence of incomplete lineage sorting and hybridization has led to proposals for workflow methods, such as that proposed by Puillandre *et al.* (2012), where barcode data are first used to form a primary species-delimitation hypotheses, which are then tested as a secondary species hypotheses using additional information such as molecular markers, morphological and ecological data (Palandačić *et al.*, 2017).

CHAPTER TWO: DNA BARCODING THE RIVERINE FISH OF TANZANIA.

2.1 *Abstract*

Management and conservation of freshwater fish resources requires a comprehensive knowledge of the species community present. However, in many countries there is a lack of information on the species distributions. In principle, there is strong potential for DNA barcoding to inform biodiversity surveys. This study tested the utility of two DNA markers in identifying freshwater fishes in a Tanzanian river system. A total of 109 fish specimens from the Ruaha-Rufiji-Kilombero river catchment, that were identified using phenotypic traits, were sequenced for the COI and/or 12S barcoding genes. Both the COI and 12S genes generated monophyletic groupings of individual species in neighbour-joining phylogenies, and large barcoding gaps, demonstrating the potential for highly accurate species identification. However, there were low success rates for species identification when relying on publically-available reference library databases alone. Morphological species identification matched COI identification for 41% of specimens using Genbank, and 47% of specimens using the Barcode of Life database. Morphological species identification matched 12S identification for 12% of specimens using Genbank. These low assignment rates were due to largely incomplete barcode libraries for the species included in our study. The results suggest that future freshwater fish biodiversity research in Tanzania would benefit from a comprehensive DNA barcode reference library.

2.2 Introduction

DNA barcoding is becoming an increasingly prominent tool for construction of biodiversity inventories (Bhattacharya *et al.*, 2015; Ward *et al.*, 2005). It has capability to reveal hidden diversity within traditional taxonomic species, and overcome limitations to conventional biodiversity survey approaches to provide highly accurate species richness estimates. To date, the COI barcode gene remains the most popular choice genetic region for barcoding (Cawthorn *et al.*, 2012), but it is recognised that additional barcode markers may be valuable for confirmatory work, and to generate reference libraries for environmental DNA metabarcoding analyses that commonly use alternative barcode markers (Collins *et al.* 2019).

In 2005 a global research project, The Fish Barcode of Life Initiative (FISH-BOL), was launched with the goal of collecting and standardising DNA barcode sequences of all fish species found worldwide. The project aimed to aid in molecular identification by offering a fast, accurate and cost-effective system for global use. A status report published in 2011 found substantial bias towards marine species, with only 13.5% of African freshwater fish species recorded as barcoded by FISH-BOL (Becker *et al.*, 2011). The absence of comprehensive DNA reference data for these species limits our ability to use DNA barcode-derived methods for biodiversity assessment. In the long-run, this may elevate the vulnerability of fish species to threats such as habitat loss, habitat alternation, water extraction, hydrological disruption, introduced invasive species, pollution, and overexploitation (Darwall *et al.*, 2011; Adeoba *et al.*, 2019).

Currently, there are approximately 3300 described fish species known from African freshwater habitats, but many more undescribed species are known. Tanzania, in particular, has a rich and diverse ichthyofauna, although most of the research into these species has focused on the species of the Rift Valley lakes, and by contrast knowledge of riverine biodiversity is less extensive (Seeteram *et al.*, 2019). The “Field Guide to the Freshwater Fishes of Tanzania” (Eccles, 1992) remains the primary source of information when identifying and studying fish from this region. Although valuable, there is a need for this work to be updated with more complete biodiversity inventories and distributional maps. This is because accurate classification of organisms is typically a key primary step in the management of biodiversity, and outdated and limited knowledge of these species can lead to negative conservation outcomes (Vu and Le, 2019).

The Rufiji River basin is the largest river in Tanzania and home to fish species of great commercial value (Raphael, 1990; Hamerlynck *et al.*, 2011). A major tributary of the Rufiji River, the Great Ruaha River, is a vital resource for many rural households while also supporting the ecology of the Usangu wetlands and the Ruaha National Park (Kashaigili *et al.*, 2007). Occupying 20% of the Great Ruaha subbasin, the Kilombero River displays high levels of diversity in its tributaries (Muñoz-Mas *et al.*, 2019). The fish biodiversity of this catchment faces increasing threats from infrastructure development. Most notably, the Stiegler’s Gorge hydropower dam has begun construction across the Rufiji River within the Selous Game Reserve in 2019. This 2,100MW hydroelectric project will result in a reservoir that is the fourth largest in Africa, and ninth largest in the world at a size of 1,350sq kilometres (3% of the Selous Game Reserve’s footprint) (Dye and Hartmann, 2017).

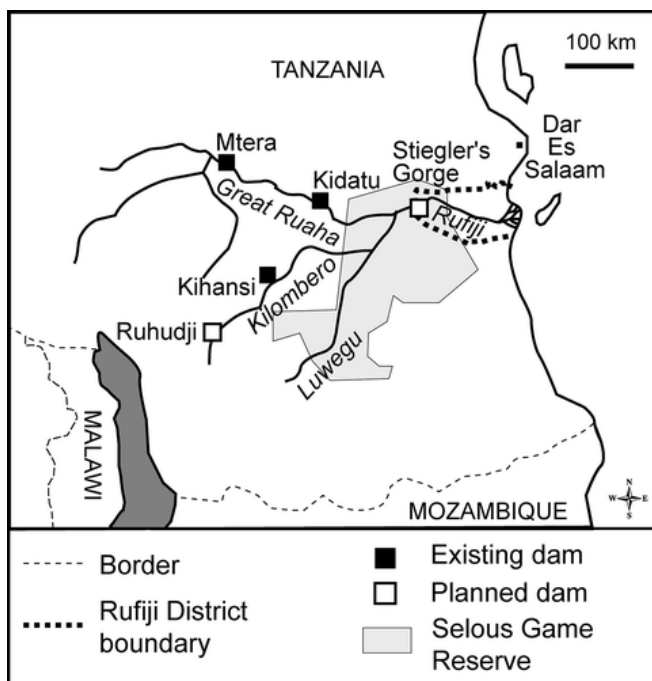


Fig. 2.1. Map detailing the construction site of the Stiegler’s Gorge project and its nearby river systems (from Hamerlynck *et al.*, 2011).

According to original environmental impact assessment reports published in 1988, water quality in the planned reservoir and downstream of the dam will be negatively affected and unfit for human and animal consumption (Mwalyosi, 1988). Further impacts include the potential for the dams to act as a barrier for the seasonal migration of fish up the Rufiji river to the Kilombero Valley Floodplain Ramsar site, while fragmenting rivers and degrading habitats (Rytwinski *et al.*, 2017; Dye and Hartmann, 2017). The hydroelectric facilities are also predicted to have serious consequences for fish populations owing to injury and mortality

resulting from entrainment and impingement. Alteration to fish populations could cause the potential collapse of the fish, prawn and shrimp fisheries found there (Dye and Hartmann, 2017). Nevertheless, despite the controversies surrounding this project, the hydropower dam will also bring benefits to Tanzania, such as providing more affordable power to the low-income citizens (Dye and Hartmann, 2017).

By monitoring fish catches in floodplain lakes that differ in their connectivity to the Rufiji river below Stiegler's gorge, Hamerlynck *et al.* (2011) were able to demonstrate that connectivity is a major determinant of catch composition in this area. If the dam construction reduces seasonal flows and therefore connectivity, it is therefore possible that species such as *Distichodus petersii*, *Citharinus congicus*, *Hydrocynus vittatus* and *Brycinus imberi* may likely decline in floodplain lakes. There are also increased likelihoods that lakes will dry out or become increasingly saline, leading to a substantial loss in human well-being due to a decrease in income from valuable fisheries. To truly understand the impacts of this dam on riverine species, however, requires "further detailed ecological research on fauna and flora in the Selous and downstream delta" (Dye and Hartmann, 2017).

Notably, an independent technical review by the IUCN (International Union for Conservation of Nature) of the Environmental Impact Assessment for Stiegler's Gorge published in April 2019 states some critical information as missing entirely, including information on aquatic biodiversity. The IUCN describe a high likelihood that there will be significant negative impacts on freshwater biodiversity in the Rufiji River, with great impacts on migratory fish. These effects may in turn result in potentially significant impacts on the livelihoods of thousands of people dependent on fisheries in the catchment, including the upstream Kilombero Valley Floodplain Ramsar Site (IUCN, 2019).

This study aimed to contribute to existing knowledge surrounding biodiversity of freshwater fish species in Rufiji-Ruaha-Kilombero catchment system, and investigate the applicability of DNA barcoding methods for species identification and delimitation for future biodiversity survey work.

2.3 Materials and Methods

Study areas

The Rufiji River Basin, which includes the Ruaha and Kilombero river systems, occupies an area equal to 20% of Tanzania, spreading across approximately 177,000km² (Raphael, 1990). The Rufiji is the largest river in Tanzania and has a mean annual flow of 800 m³/s (Hamerlynck *et al.*, 2011). Around one fourth of all the fish species in the Lower Rufiji floodplain are of commercial value, making the Rufiji River a major contributor to Tanzanian fisheries (Raphael, 1990). In 2016, an environmental flow assessment conducted on behalf of USAID documented 27 species of fish in the Rufiji River basin, around half of known species from the region (Smith, 2016). Data was also gathered on the longitudinal and latitudinal migrations of several species, including *Hydrocynus vittatus*, *Barbus macrolepis*, and *Brycinus affinis*.

The Lower Rufiji floodplain contains eight major floodplain-adjacent lakes near to Utete. The freshwater fish species composition of these lakes was investigated in 2011 by Hamerlynck *et al.* and was broadly split into four catch groups, cichlids, alestids, citharinids and several other families including mostly ‘catfish’. Notable species with high catch composition included *Oreochromis urolepis*, *Brycinus affinis*, *Brycinus imberi*, *Alestes stuhlmanni*, *Hydrocynus vittatus*, *Citharinus congicus*, *Distichodus petersii*, *Clarias gariepinus*, *Bagrus meridionalis*, and *Glossogobius giurus* (Hamerlynck *et al.*, 2011).

The Great Ruaha River is a major tributary of the Rufiji River, Tanzania, occupying 47% of the Rufiji Basin (83,970km²) and contributing to 15% of its runoff waters (Raphael, 1990). The river and its tributaries are vital for the livelihoods of many rural households, accounting for approximately 50% of the country’s hydropower production, and play a valuable role in supporting the ecology of the Usangu wetlands and the Ruaha National Park (Kashaigili *et al.*, 2007). In 1976, Petr found the fish family catch species composition in the middle course of the Great Ruaha River to be Cyprinidae (40.94%), Characidae (28.86%), Citharinidae (8.89%), followed by Mormyridae (4.16%) and several other families including Bagridae (2.24%), Mochokidae (1.59%), Clariidae (3.38%), and Cichlidae (1.04%). However, this species composition may have changed subsequently as a result of aquaculture and introduced species.

The Kilombero River, also known as the Ulanga River, occupies 20% of the Great Ruaha subbasin and contributes to 62% of the Rufiji Basin runoff (Raphael, 1990). The area is described as a sub-humid tropical climate with humidity levels ranging from 70 to 80% and with an annual rainfall of about 1200 to 1400 mm mainly falling within the rainy seasons of March to May, and October to December (Mombo *et al.*, 2011). Fish from the Rufiji River system are reported to migrate upstream to the Ulanga to spawn, usually taking place at the

beginning of November with the peak of spawning activity occurring in December (Utzing and Charlwood, 1997). In surveys conducted by Utzing and Charlwood in 1996, *Bagrus orientalis* was the most common fish caught in the Kilombero River, contributing to 49% of the total biomass. Other large fish of the region included *Clarias gariepinus*, *Hydrocynus* spp. and *Labeo congoro* (Utzing and Charlwood, 1996). More recent studies in 2019 found lotic communities in two tributaries of the Kilombero River to contain catfish (e.g. *Amphilius* spp., *Chiloglanis* spp. and *Schilbe* spp.), as well as cyprinids (e.g. *Labeo cylindricus*) and shellears (*Parakneria* spp.) (Muñoz-Mas *et al.*, 2019).

Sampling, DNA extraction, PCR and sequencing

A total of 122 specimens were sampled between 2012 and 2019 from 12 different sampling sites around the Great Ruaha, Kilombero and Rufiji river catchments (Table 2.1). Specimens were obtained either through opportunistic purchasing from artisanal fish farms or markets (if the source of the fish was known), or through the use of experimental seine nets and gill nets. Specimens were identified to the species level in the field, whenever possible, with the aid of a field guide for taxonomic studies (Eccles, 1992). The fish were then preserved in ethanol, and subsequently placed in 70% ethanol for long-term storage. For some specimens, genetic samples (fin clips) were taken and preserved in absolute ethanol separately in the field.

Table 2.1. Collection information for sampling sites and specimens analysed.

Site	Date	Location	Latitude (°S)	Longitude (°E)	n Sampled
S8	04-Sep-12	Great Ruaha River	8.860	34.025	15
S9	04-Sep-12	Great Ruaha River	8.855	34.086	9
S22	20-Aug-13	Rufiji River	7.858	38.962	19
KIL01	21-May-19	Kilombero River	8.138	36.679	26
KIL02	21-May-19	Kilombero River	8.189	36.693	11
KIL03	21-May-19	Kilombero River	8.178	36.694	6
KIL07	22-May-19	Kilombero River	8.140	36.673	14
MG10	20-May-19	Ruaha/Kilombero River	7.841	36.892	1
KIL08	22-May-19	Kilombero River	8.174	36.694	15
KIL09	23-May-19	Kilombero River	8.084	36.710	2
KIL10	23-May-19	Kilombero River	7.894	36.872	3
KIL11	23-May-19	Great Ruaha River	7.666	36.977	1

Total genomic DNA was extracted from a small quantity (~3mm³) of tissue from each specimen, using 200µL nuclei lysis solution (Promega, Madison USA) and 2µL proteinase K (Qiagen, Hilden Germany). The solution was warmed at 60°C until the tissue was broken down, and then 65µL protein precipitation solution (Promega, Madison USA) was added. The sample was centrifuged at 14,600rpm for 2 min, and the supernatant was poured into a new eppendorf. The DNA was precipitated with the addition of 600µL pure ethanol, before being centrifuged at 14,600rpm for 2 min. The liquid was then removed, leaving only a DNA pellet that was cleaned by adding 600µL of 70% ethanol, before being centrifuged at 14,600rpm for 2 min. The liquid was then removed, and the remaining DNA pellet was air dried, before being resuspended in 50µL of molecular biology grade H₂O.

The novel primers Aa22-12SF (5'-AGC ATA ACA CTG AAG ATR YTA RGA-3') and Aa633-12SR (5'-TTC TAG AAC AGG CTC CTC TAG-3') were used to amplify a 567-bp fragment of the 12S rRNA gene from all fish specimens. 12S PCR amplifications were performed in 20µL volumes with 5µL molecular biology grade H₂O, 10µL GoTaq Green MasterMix, 2µL of each primer (2µM), and 1µL genomic DNA. PCR conditions for 12S were as follows: an initial denaturation was performed at 95°C for 2 min, following by 40 cycles of denaturation at 95°C for 30s, annealing at 53°C for 30s, and extension at 72°C for 45s. A final extension step was performed at 72°C for 5 min.

COI primers FishF1 (5'-TCA ACC AAC CAC AAA GAC ATT GGC AC-3') and FishR1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3'), from Ward *et al.* (2005), were used to amplify a 655-bp fragment of the COI gene from mitochondrial DNA from all fish specimens. COI PCR amplifications were performed in 25µL volumes with 9.5µL molecular biology grade H₂O, 12.5µL GoTaq Green MasterMix, 1µL of each primer (10µM), and 1µL genomic DNA. PCR conditions for COI were as follows: an initial denaturation was performed at 95°C for 30s, following by 35 cycles of denaturation at 95°C for 30s, annealing at 52°C for 30s, and extension at 72°C for 1min. A final extension step was performed at 72°C for 5 min.

PCR success was verified by 1% agarose gel electrophoresis. A DNA template negative control reaction was included alongside all experiments to negate contamination. Sequences were obtained through Sanger sequencing of PCR products by Eurofins Genomics, using reverse primers for both markers. In total high quality 12S sequences were obtained for 83 individuals, while high quality COI sequences were obtained for 81 individuals, and 109 individuals had

one or both markers sequenced. All 109 individuals were identified to the genus level, with 74 visually identified to the species level by phenotypic characters.

Processing of DNA sequences

All successful, raw DNA sequences were all checked visually, and their ends trimmed, using Chromas 2.6 (Technelysium Ltd, Brisbane). Using MEGAX (Kumar *et al.*, 2018), the sequences were aligned using ClustalW (Thompson *et al.*, 1994) under default parameters. Separate alignments were also constructed for the COI and 12S sequences that had confirmed visual identifications, with more than one species representative. COI sequences were checked against NCBI Genbank (using BLAST) (Altschul *et al.*, 1990) and the BOLD databases (via <https://www.boldsystems.org/>) for species identification, with the highest ‘Max Score’ and ‘Similarity’ result noted respectively. 12S sequences were entered into the NCBI Genbank database (using BLAST) (Altschul *et al.*, 1990) for species identification, with the highest ‘Max Score’ species result recorded.

Neighbour-joining reconstruction of COI and 12S sequence relationships

To visualise relationships among derived sequences, we constructed a neighbour-joining tree for COI and 12S datasets (Saitou and Nei, 1987). All samples for each marker were included in their respective phylogeny, regardless of confirmed visual species identification. Neighbour-joining trees were constructed using MEGAX (Kumar *et al.*, 2018) using p-distance as the distance measure. All other parameters were set as default settings.

Intraspecific and interspecific variation, and DNA barcoding gap

The ‘barcoding gap’ refers to the separation between mean intraspecific and interspecific sequence variability for barcoding sequences. A large barcoding gap is commonly accepted as a way of predicting high DNA-barcoding success, allowing for easy species discrimination (Meier *et al.*, 2008). To confirm whether species identification is possible using COI and 12S markers for the specimens of this study, we calculated the average intraspecific distance and minimum intra-specific distance for the COI and 12S datasets using R 3.6.2 (R Core Team, 2019), and the package `barcoding.gap` (Dai *et al.*, 2012).

2.4 Results

DNA barcoding

Among the specimens, 25 described species across 17 genera were identified using phenotypic traits to the species level: *Enteromius kerstenii*, *Enteromius apleurogramma*, *Enteromius macrotaenia*, *Brycinus affinis*, *Brycinus imberi*, *Brycinus lateralis*, *Kneria ruaha*, *Labeo congoro*, *Labeo cylindricus*, *Labeobarbus macrolepis*, *Distichodus petersii*, *Clarias gariepinus*, *Clarias theodora*, *Oreochromis urolepis*, *Hydrocynus vittatus*, *Synodontis rufigiensis*, *Synodontis matthesi*, *Synodontis rukwaensis*, *Citharinus congicus*, *Ctenopoma muriei*, *Alestes stuhlmannii*, *Pareutropius longifilis*, *Glossogobius giuris*, *Hemigrammopetersius barnadi*, and *Mastacembelus frenatus*. *Eleotris klunzingerii* was later identified following DNA information.

A total of 81 COI sequences were entered into the NCBI and BOLD databases for evaluation of their ability to identify species, while a total of 83 12S sequences were entered into the NCBI database. Full details of these results are found in Appendix 2.1.

Of the 81 COI sequenced individuals, 21 individuals across nine species were correctly identified using the NCBI Genbank database, while 11 fish species (*Alestes stuhlmannii*, *Clarias theodora*, *Enteromius apleurogramma*, *Enteromius macrotaenia*, *Kneria ruaha*, *Labeo congoro*, *Labeo cylindricus*, *Labeobarbus macrolepis*, *Pareutropius longifilis*, *Synodontis matthesi*, and *Synodontis rufigiensis*) did not have COI barcode for their species on the NCBI database. Six individuals across three species (i.e. *Enteromius kerstenii*, *Oreochromis urolepis*, and *Hydrocynus vittatus*) were misidentified when using the NCBI database, despite a reference COI barcode being present for their correct species, as the suggested identification was incongruent with the specimen phenotype.

Of the 81 COI sequenced individuals, 24 individuals across 11 species were correctly identified using the COI barcode sequence by BOLD. Eight fish species (*Alestes stuhlmannii*, *Enteromius apleurogramma*, *Kneria ruaha*, *Labeo congoro*, *Labeobarbus macrolepis*, *Pareutropius longifilis*, *Synodontis matthesi*, and *Synodontis rufigiensis*) did not have a reference sequence on the BOLD database. Eight individuals across five species (*Enteromius kerstenii*, *Brycinus imberi*, *Clarias theodora*, *Labeo cylindricus*, and *Brycinus lateralis*) were misidentified when using the BOLD databases, despite a reference COI barcode being present on the BOLD database for their species, again due to the suggested identification being incongruent with the specimen phenotype.

Of the 83 12S sequenced individuals, seven individuals across four species were correctly identified by NCBI using the 12S barcode sequence. A reference 12S barcode sequence of 14 species (i.e. *Alestes stuhlmannii*, *Brycinus lateralis*, *Ctenopoma muriei*, *Distichodus petersii*, *Enteromius apleurogramma*, *Enteromius kerstenii*, *Enteromius macrotaenia*, *Hemigrammopetersius barnadi*, *Kneria ruaha*, *Mastacembelus frenatus*, *Oreochromis urolepis*, *Pareutropius longifilis*, *Synodontis rufigiensis*, and *Synodontis rukwaensis*) was not yet been recorded in the NCBI database. 14 individuals across five species (i.e. *Brycinus affinis*, *Clarias theodora*, *Hydrocynus vittatus*, *Labeo congoro*, and *Labeo cylindricus*) were misidentified when using NCBI, despite a reference 12S barcode present on the NCBI database for their correct species.

Genetic divergence

The neighbour-joining trees produced from 81 COI sequences (Fig. 2.2) and 83 12S sequences (Fig 2.3) showed that all sequences from phenotypically-identified species were monophyletic, and reliably clustered according to their families. Both genes revealed the presence of a barcoding gap between intraspecific and interspecific divergences, nevertheless, there were substantial intraspecific divergences observed in both loci, with the greatest intraspecific divergences observed for the COI gene, commonly associated with allopatric populations (Fig 2.4).

Table 2.2. COI and 12S inter- and intra-specific genetic distances between fish species with phenotypic identification and more than one species representative.

	COI	12S
<i>Average inter-specific distance</i>	0.243	0.241
<i>Minimum inter-specific distance</i>	0.100	0.050
<i>Maximum inter-specific distance</i>	0.317	0.300
<i>Average intra-specific distance</i>	0.006	0.001
<i>Minimum intra-specific distance</i>	0.000	0.000
<i>Maximum intra-specific distance</i>	0.078	0.025

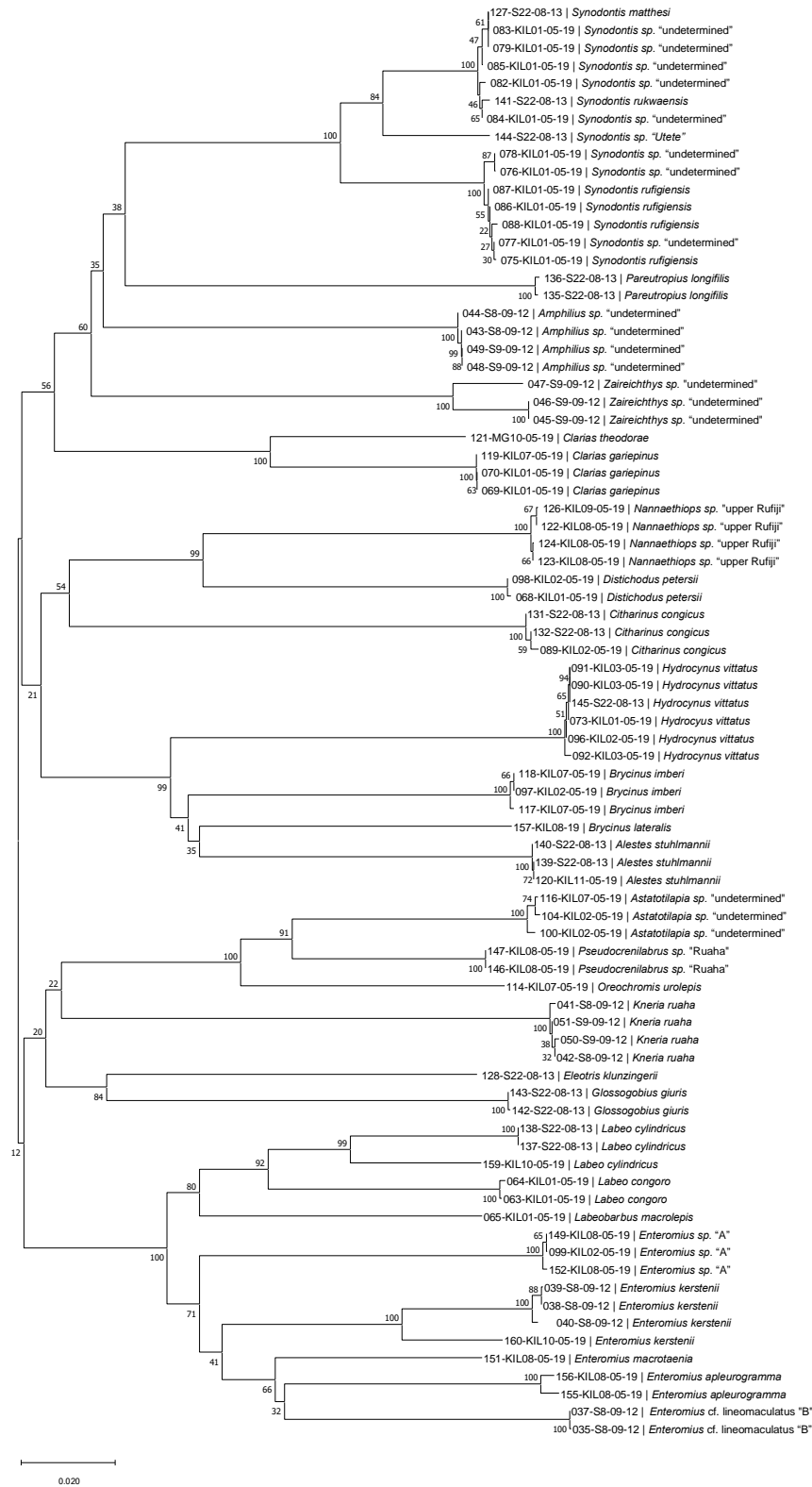


Fig. 2.2. Neighbour-joining tree produced using the COI barcoding gene for 81 specimens. Bootstrap values are indicated at tree nodes as percentages of 500 replicates. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

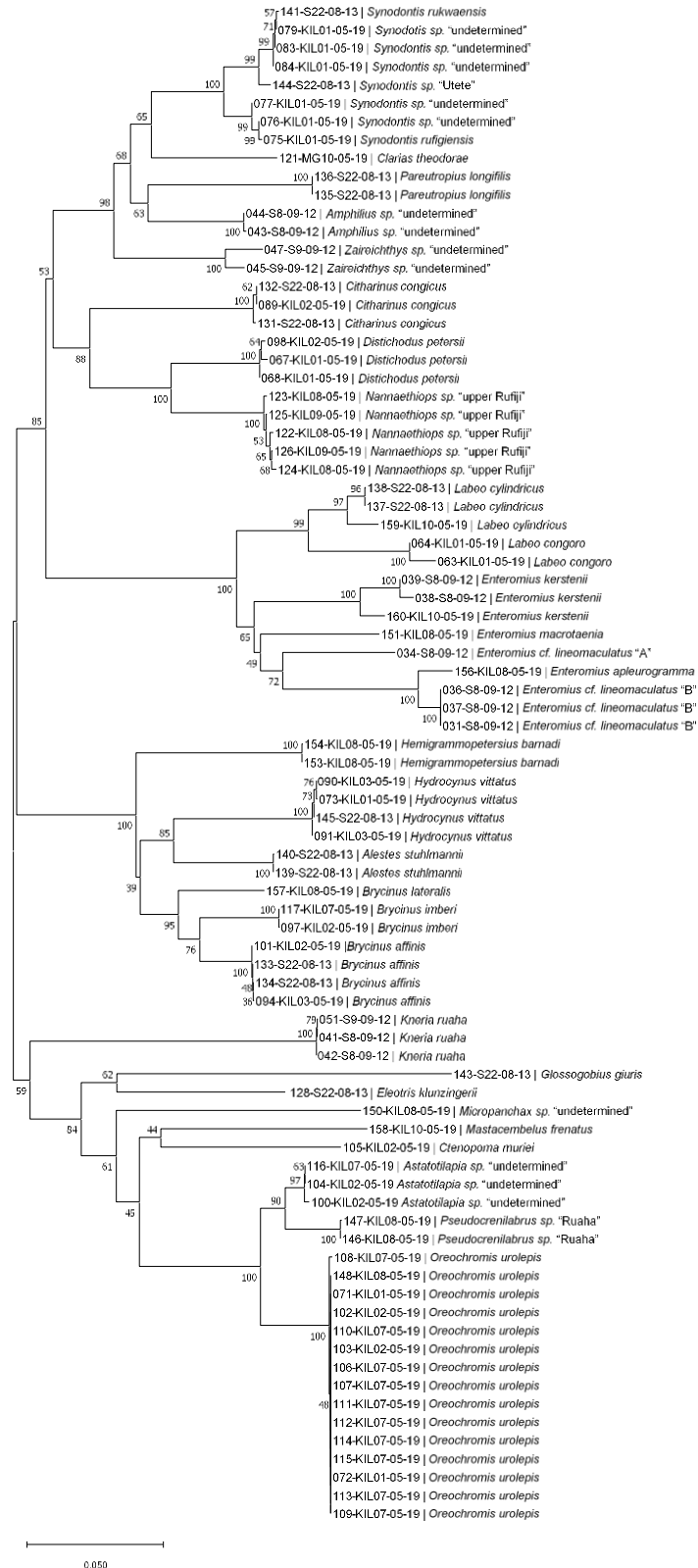


Fig. 2.3. Neighbour-joining tree produced using the 12S barcoding gene for 83 specimens. Bootstrap values are indicated at tree nodes as percentages of 500 replicates. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

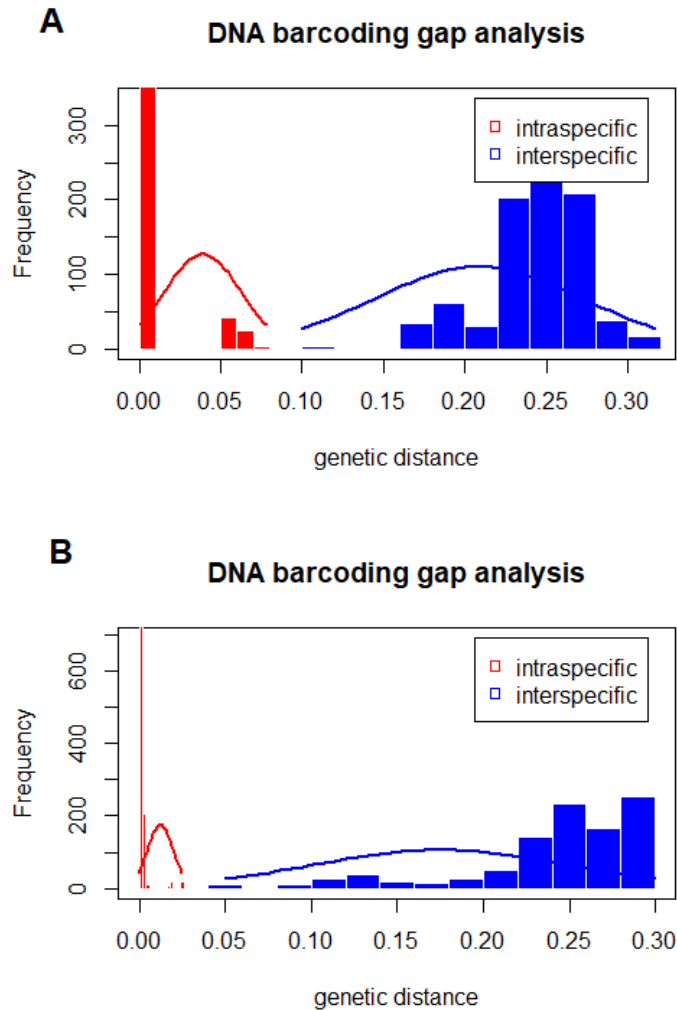


Fig. 2.4. Histograms of intra- and inter-specific pairwise genetic p-distance between single-gene barcodes for freshwater fish species. A) COI barcode. B) 12S barcode.

2.5 Discussion

In this study, the DNA barcoding approach generated clear monophyletic clusters of conspecific individuals that had been phenotypically identified to described species, suggesting that the method can be used successfully to assign species identities if suitable reference databases exist. Further evidence for their use comes from the large barcoding gaps seen in both markers between their respective intraspecific and interspecific genetic distances. A large barcoding gap is commonly seen as a reliable way of predicting high DNA-barcoding success and easy species discrimination, suggesting that both genes should be almost equal in their success to delimit species in identification studies.

However in the absence of suitable reference data on either NCBI Genbank or BOLD, then the ‘rate of success’ was low for both markers. This is in stark contrast to studies that have returned 95.6% barcoding success rates in Nigeria, and 93.0% barcoding success rates for Canadian freshwater fish (Iyiola *et al.*, 2018; Hubert *et al.*, 2008). In total 11 phenotypically-identified species in this study did not have a representative complete COI barcode sequence recorded on NCBI, while eight species do not have a recorded COI barcode on BOLD. Additionally, 14 species do not have a representative complete 12S barcode sequence recorded on NCBI. This is despite all missing species having formal species descriptions.

The high number of species without barcode information highlighted in this relatively small sample is a cause for concern for Tanzania’s ichthyofauna conservation efforts. The facilitation of species identification is crucial when Africa’s freshwater fish species are currently under threat from factors such as climatic change and water pollution. Without extensive DNA information, species may be misidentified or under-represented when recording threats to ecosystems, and biodiversity levels measured through DNA approaches may be underestimated. Thus, the development of a complete DNA barcode reference library for Tanzania’s ichthyofauna is required to facilitate taxonomic study. However, information must be continuously updated and include new discoveries of taxonomic diversity within river systems.

In this study, we also observed cases where species identification using the DNA barcoding approach did not match those assigned using morphological methods. This is despite a record of the correct species complete barcode sequence on the databases. For example, *Oreochromis urolepis* specimens were assigned as *Oreochromis* sp. ‘red tilapia’ using Max Score under the COI barcode in NCBI, with high ID scores of around 99%. This is despite the presence of relevant *Oreochromis urolepis* barcode sequences on the NCBI database. It is plausibly the case that *Oreochromis* sp. ‘red tilapia’, a hybrid variant of *Oreochromis* developed for aquaculture, has mitochondrial DNA originally derived from *Oreochromis urolepis*.

Misidentification of individuals through the use of DNA barcoding can potentially have multiple sources. Errors might occur through morphological misidentifications of voucher specimens submitted to DNA databases (Iyiola *et al.*, 2018). Equally, there are likely to be places where either incomplete lineage sorting or hybridization leads to individuals from different species sharing similar mitochondrial haplotypes. This demonstrates the requirement for both morphological and molecular methods in the study of Tanzania’s fish species. The

combination of this work will help to avoid matching source sequences to taxonomically uninformative entries (such as *Barbus* sp. CTOL 051).

DNA barcodes highlight novel species diversity

In this study, three samples (099-KIL02-05-19, 149-KIL08-19, and 152-KIL08-19) were all identified as *Enteromius apleurogramma* by NCBI with ID scores of 99-100% using COI barcode sequences. However, these specimens, all phenotypically assigned to an undescribed taxon *Enteromius* sp. “A”, do not match the typical phenotypic characters of *Enteromius apleurogramma*. This may be due to morphological misidentifications of the voucher *Enteromius apleurogramma* specimen recorded on the NCBI database for which the COI barcode was sequenced. Morphological misidentifications are common in literature and have been described as ‘particularly rampant’ in public databases, with figures reaching well over 5% (Čandek and Kuntner, 2014). This highlights the need for voucher collections to be carefully assessed and taxon identities updated. *Enteromius*, in particular, is a complex genus with most predicted species remaining undescribed and only a few specimens available for study for many described species (Schmidt *et al.*, 2018). *Enteromius* sp. “A” could therefore represent a new, undescribed species, or perhaps it may represent a known species but without recorded barcode sequences or extensive morphological information.

Sample 157-KIL08-19 was phenotypically identified as the striped robber, *Brycinus lateralis*. This identification was further supported using the COI barcode sequence in NCBI. This appears to be the first record for this species in this region of Tanzania. Previous records of this species distribution include Angola, Botswana, Malawi, Mozambique, Namibia, South Africa, Zambia, and Zimbabwe (Tweddle *et al.*, 2019). The overall population trend for this species is unknown, however continuous habitat degradation in South Africa is suspected to cause a decline in population size. Further records of this species across different regions of Africa increases its likelihood of survival against threats to biodiversity.

Two samples 146-KIL08-19 and 147-KIL08-19 from the Kilombero river system, were phenotypically identified as *Pseudocrenilabrus* spp., an identification supported by the COI barcode data. This represents the first record of the genus from this sector of the Rufiji basin. It is likely these specimens are *Pseudocrenilabrus* sp. “Ruaha”, which although similar in appearance to the widespread *P. philander*, has several distinctive markings along the dorsal fin and a characteristic melanin pattern is thought to distinguish this species as separate from

others (Morgenstern, 2008). Further samples will enable the species status of this species to be resolved.

Many species assigned to the *Synodontis* genus in this study were described as “undetermined” due to difficulty surrounding phenotypic identification of representatives of this genus in the field. Identification of many of these species is notoriously difficult, even with the assistance of field ID guides, not least due to the variation in morphological characters during ontogeny (Day *et al.*, 2009). Here, sample 144-S22-08-13 was provisionally assigned the phenotype ID of *Synodontis* spp. “Utete”, reflecting the location in which the specimen was collected. Based on gross phenotypic characters, it is possible that this specimen is *Synodontis orientalis*. However *S. orientalis* does not have a record of a reference COI barcode sequence on NCBI and BOLD, nor a reference 12S barcode sequence on NCBI for comparison. This highlights the need for linking detailed morphological study with extensive barcode library for Tanzania’s freshwater fish species.

The assigned identification for one specimen was amended retrospectively following the results of DNA analysis. Sample 128-S22-08-13 was originally identified as *Eleotris melanosoma*, however was later reassigned to *Eleotris klunzingerii* following the high ID matches (>99%) of its COI and 12S barcode in the NCBI and BOLD databases. Field identification of *Eleotris* species is difficult due to apparent overlap of meristic characters across several species, and separation of *E. melanosoma* from *E. klunzingerii* can only be achieved using DNA barcoding (Mennesson and Keith, 2017). This highlights the importance of DNA barcoding in taxonomy, and shows how species richness of a habitat may be underestimated in the absence of genetic information to facilitate the identification of species.

Biodiversity assessment

A total of 26 described species from 18 genera were included in this study across 74 phenotypically identified samples. In total 129 species have been recorded from the Kilombero, Ruaha and Rufiji river systems, across multiple studies (Eccles, 1992; Doody and Hamerlynck, 2003; Utzinger and Charlwood, 1996; Hamerlynck *et al.*, 2011; Van der Knaap, 1994; Msangameno and Mangora, 2016; Froese and Pauly, 2019). Thus, this study only provides a snapshot of the vast freshwater fish biodiversity in the catchment. Moreover, it is notable that 35 specimens were defined as either undetermined (due to unclear identification in the field) or unable to be identified from pre-existing taxonomic references. These specimens could

represent currently undescribed species from these regions, rendering the full diversity of this system even greater than that currently appreciated in academic texts.

With populations of species such as *Distichodus petersii*, *Citharinus congicus*, *Hydrocynus vittatus* and *Brycinus imberi* all predicted to decline due to construction of the Stiegler's Gorge dam, it is likely many more freshwater species will be unknowingly impacted due to restricted research in Tanzania (Hamerlynck *et al.*, 2013). WWF stated in their 2017 commissioned research of Stiegler's Gorge that "further detailed ecological research on fauna and flora in the Selous and downstream delta is needed to enable fuller knowledge of the complete species and ecological dynamics present" (Dye and Hartmann, 2017). This study highlights new records of species to these regions and emphasises the potential for formally undescribed species to be present. Thus, many more discoveries are likely to be made in wider and larger studies of these river systems. Results from these studies should then be considered in any further updates to the environmental impact assessments of Stiegler's Gorge, which at present do not fully accommodate knowledge of aquatic biodiversity. Without these assessments, and without the implementation of adequate mitigation procedures, it is likely that the fish of these river systems will face great impacts from the construction of the dam. Additionally, fish in these regions also face threats from several other factors including the impacts of climate change, eutrophication of their waters and the introduction of invasive species for commercial use.

Concluding remarks

Our study finds that the COI gene and 12S gene are both suitable barcoding regions for the identification of Tanzania's freshwater fish species, but the absence of a comprehensive DNA barcode reference library limits the use of DNA barcoding as a useful tool in evaluating fish biodiversity. Notably, new eDNA-based methods for surveying fish biodiversity also require comprehensive and well curated reference libraries. Recent work has emphasised the ability of 12S as a reliable taxon-specific marker for eDNA-based fish metabarcoding biodiversity studies (Collins *et al.*, 2019). Therefore, it would be useful if future work built reference libraries for fish of the region sequence both COI and 12S, enabling contributions to both global biodiversity barcoding indicatives, and enabling practical eDNA field research.

A key result from this work is that many, as yet, undescribed or previously unrecognised species are likely to be present in the Rufiji basin system. Many species uncovered in our study could face substantial negative impacts following the construction of Stiegler's Gorge project.

The findings invite further advanced biodiversity research to help conserve Tanzania's freshwater fish species.

CHAPTER THREE: MORPHOLOGICAL DIVERSITY IN *CHILOGLANIS* CATFISH POPULATIONS OF SOUTH-EAST AFRICA.

3.1 Abstract

Species are considered to be “cryptic” if multiple taxa cannot easily and reliably be distinguished on the basis of phenotypic characteristics alone. Such patterns can arise where there is strong stabilising selection on ecologically-important phenotypic traits that promote survivorship, despite the evolution of reproductive isolation. This study focused on morphological diversity of riverine catfishes of the genus *Chiloglanis* from Malawi, Zambia and Tanzania, in the context of a recently developed phylogeny based on genome-wide nuclear markers. The phylogeny delimited 11 clades (putative species) with partially overlapping distributions. Here, the morphological diversity both within and among those 11 clades was quantified using linear morphological measurements. We found that clades had substantial morphological overlap overall, consistent with the presence of superficially cryptic species in the region. Nevertheless, measurements that are putatively taxonomically-informative could be identified for the separation of most sympatric clades. We also found that most allopatric populations of the same clade have diverged in morphology. On the basis of these results, we suggest there is potential for further study to confirm morphological measurements that separate co-occurring species within this group, but further study is required to confirm the utility of these traits. We consider it plausible that the morphological diversity within *Chiloglanis* has arisen from stabilising selection on an overall fluvial phenotype, coupled with adaptation to specific characteristics of local river systems, and niche partitioning to promote coexistence among sympatric reproductively isolated populations.

3.2 Introduction

Estimation of freshwater fish species richness within biogeographic regions is challenging. Without a complete knowledge of the species present, habitats of conservation importance can go unrecognised, and freshwater management decisions may fail to fully account for the potential impacts on biodiversity (Palandačić *et al.*, 2017). In particular, “cryptic” species that are most readily separated by genetic evidence can be overlooked (Jirsová *et al.*, 2019). Southern Africa’s stream fish are relatively poorly understood from a taxonomic perspective, leaving potentially large numbers of species unprotected and at risk (Kadye and Moyo, 2008). Surveys of both morphological and genetic characteristics are needed to fully appreciate the full diversity of these systems, given the high potential for the presence of cryptic biodiversity.

One of the species groups that may contain cryptic species are the mochokid catfishes, a group found naturally only in the freshwaters of Africa. They have an almost ubiquitous distribution across tropical and subtropical regions of sub-Saharan Africa and possess a high degree of morphological diversity that allows them to inhabit a range of habitats, from deep lakes to fast flowing streams (Day *et al.*, 2013). The second most species-rich genus within the family Mochokidae is the genus *Chiloglanis* (Schmidt *et al.*, 2015). There are 51 species formally recognised to date, but taxa within this genus are hard to differentiate on the basis of morphological characters, and diversity is likely to be higher than that currently recognised (Seegers, 2008; Schmidt and Barrientos, 2019). The genus is distributed across almost the whole of tropical Africa, from the Nile drainage in the north, to the Phongolo system in the south-east (Seegers, 2008). The genus typically inhabits fast-flowing large rivers or smaller streams, and species tend to avoid lakes, with the exception of Lake Malawi in which they populate the surge zone of the rocky shoreline (Seegers, 2008). The lips and barbels of *Chiloglanis* have been modified into a sucking disc that is used for clinging to substrates in these turbulent habitats, where they feed primarily on epilithic algae (*aufwuchs*) (Seegers, 2008).

Previous work has demonstrated strong genetic structure among *Chiloglanis* sampled regionally, and these patterns of spatial genetic structure suggest a high degree of endemism to their respective habitats. Moreover, applications of population genetic and phylogenetic methods have informed species delimitation using morphological characters, despite overall morphological similarities of *Chiloglanis* species (e.g. Schmidt *et al.*, 2015). Specifically, detailed investigations have shown external body measurements to be taxonomically

informative, as well as the number of premaxillary teeth, shape and size of the sucker, and barbel morphology (Schmidt and Barrientos, 2019). Schmidt *et al.* (2015) used Principal Component Analysis (PCA) on 45 linear morphometric measurements to distinguish between three genetically-distinct *Chiloglanis* species, identifying oral disc length, prepectoral length, and dorsal spine length as traits differing among species. Morris *et al.* (2015) used two-dimensional geometric morphometric analysis of landmark and semi-landmark data to investigate body and caudal fin shape across six clades of a single species *Chiloglanis anoterus*. They showed that although morphological body shape data was unable to effectively separate either the clades or sexes, caudal fin shape variation did show sexual dimorphism.

To explore molecular diversity within and among *Chiloglanis* catfish populations from Malawi, northern Zambia and southern Tanzania, Watson (2020) used genome-wide nuclear ddRAD markers. This phylogeny delimited eleven clades (putative species), but morphological data are required to test if representatives of the clades are distinguishable on the basis of phenotypic traits, or instead represent fully cryptic species. The phylogeny of Watson (2020) is notable because it contains multiple allopatric populations of the same clade. Thus, it is possible to quantify spatial variation within a species that may be linked to either phenotypic plasticity, genetic drift, or natural selection within their local environment (Williams, 1966). Additionally, the phylogeny includes several cases where representatives of different clades are present in sympatry. This enables the quantification of diversity between species, which may reflect niche partitioning to promote coexistence among reproductively isolated populations.

Here, the same *Chiloglanis* specimens studied by Watson (2020) are studied using linear morphological measurements. Morphological diversity within and between genetically-identified clades is quantified, and morphological traits that potentially separate sympatrically-occurring clades are identified. The results provide insight into methods of assessing diversity within this species group and enable discussion about the evolutionary mechanisms that have promoted that diversity.

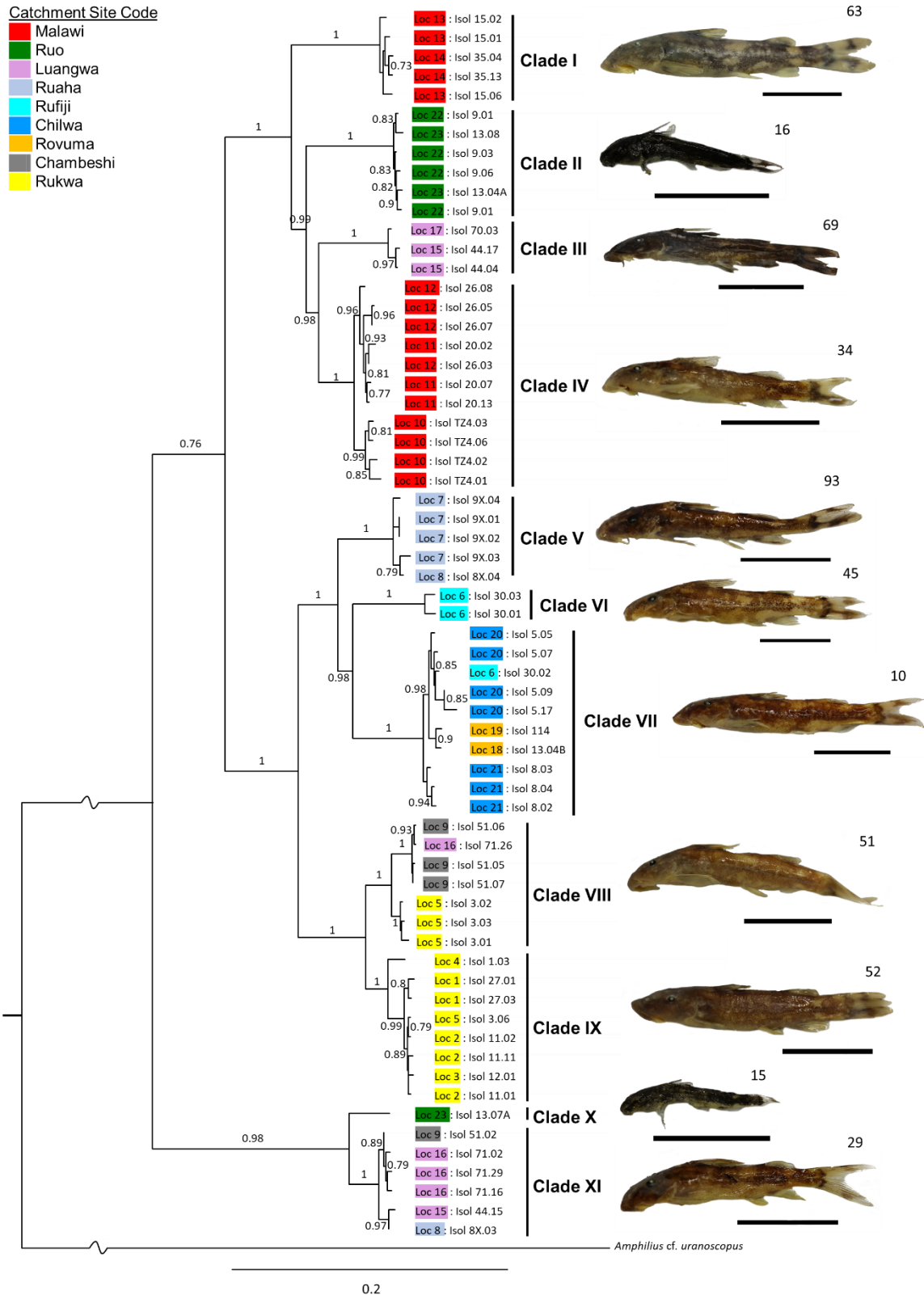


Fig. 3.1. Maximum likelihood phylogeny based on 4119 SNPs, from Watson (2020), alongside representative specimens for each clade. Numbers indicate proportional branch support values using the aLRT-SH-like approach in phyML. Scale bar represents the genetic distance on the basis of the GTR model. Location codes refer to map in Appendix 3.1.

Table 3.1. Summary of morphological samples analysed, including respective catchment area and sampling site (see Appendix 3.1 for site map).

Lab Code	Isolate	Clade	Catchment	Site (Map)	Site Name	Collection Date	Latitude	Longitude	Morphology
17	15.01	I	Malawi	13	Nkhotakota	23/05/2010	-12.8337	34.1623	Yes
18	15.02	I	Malawi	13	Nkhotakota	23/05/2010	-12.8337	34.1623	Yes
20	15.06	I	Malawi	13	Nkhotakota	23/05/2010	-12.8337	34.1623	Yes
61	35.04	I	Malawi	14	Linthipe	09/06/2010	-14.1795	34.1245	Yes
63	35.13	I	Malawi	14	Linthipe	09/06/2010	-14.1795	34.1245	Yes
5	9.01	II	Ruo	22	Chiradzulu	17/05/2010	-15.8463	35.1932	Yes
6	9.02	II	Ruo	22	Chiradzulu	17/05/2010	-15.8463	35.1932	Yes
7	9.03	II	Ruo	22	Chiradzulu	17/05/2010	-15.8463	35.1932	Yes
8	9.06	II	Ruo	22	Chiradzulu	17/05/2010	-15.8463	35.1932	Yes
14	13.04 (A)	II	Ruo	23	Muluvia	19/05/2010	-16.0004	36.3207	Yes
16	13.08 (A)	II	Ruo	23	Muluvia	19/05/2010	-16.0004	36.3207	Yes
54	70.03	III	Luangwa	17	Nzamane	20/06/2010	-13.7587	32.4498	Yes
69	44.04	III	Luangwa	15	Chitiwa	15/06/2010	-13.7868	28.9997	Yes
72	44.17	III	Luangwa	15	Chitiwa	15/06/2010	-13.7868	28.9997	Yes
1	TZ4.01	IV	Malawi	10	Impinda	16/07/2011	-9.3954	33.8273	Yes
2	TZ4.02	IV	Malawi	10	Impinda	16/07/2011	-9.3954	33.8273	Yes
3	TZ4.06	IV	Malawi	10	Impinda	16/07/2011	-9.3954	33.8273	Yes
4	TZ4.03	IV	Malawi	10	Impinda	16/07/2011	-9.3954	33.8273	Yes
21	26.03	IV	Malawi	12	Zebedia Shawa	29/05/2010	-12.2722	33.4878	Yes
22	26.05	IV	Malawi	12	Zebedia Shawa	29/05/2010	-12.2722	33.4878	Yes
23	26.07	IV	Malawi	12	Zebedia Shawa	29/05/2010	-12.2722	33.4878	Yes
24	26.08	IV	Malawi	12	Zebedia Shawa	29/05/2010	-12.2722	33.4878	Yes
34	20.02	IV	Malawi	11	Rumphi	26/05/2010	-11.0196	33.7857	Yes
35	20.07	IV	Malawi	11	Rumphi	26/05/2010	-11.0196	33.7857	Yes
36	20.13	IV	Malawi	11	Rumphi	26/05/2010	-11.0196	33.7857	Yes
76	9X.03	V	Ruaha	7	Chiradzulu	04/09/2012	-8.7693	34.3748	Yes
93	9X.01	V	Ruaha	7	Chiradzulu	04/09/2012	-8.76939	34.3748	Yes
94	9X.02	V	Ruaha	7	Chiradzulu	04/09/2012	-8.76939	34.3748	Yes
95	9X.04	V	Ruaha	7	Chiradzulu	04/09/2012	-8.76939	34.3748	No
96	9X.04	V	Ruaha	7	Chiradzulu	04/09/2012	-8.76939	34.3748	Yes
92	8X.04	V	Ruaha	8	Ruaha	04/09/2012	-8.85472	34.08608	No
45	30.01	VI	Rufiji	6	Kidodo	02/08/2017	-7.48484	37.0285	Yes
9	5.05	VII	Chilwa	20	Zomba	15/05/2010	-15.2791	35.4011	Yes
10	5.07	VII	Chilwa	20	Zomba	15/05/2010	-15.2791	35.4011	Yes
11	5.09	VII	Chilwa	20	Zomba	15/05/2010	-15.2791	35.4011	Yes
12	5.17	VII	Chilwa	20	Zomba	15/05/2010	-15.2791	35.4011	Yes
26	8.02	VII	Chilwa	21	Zomba	16/05/2010	-15.486	35.2364	Yes
27	8.03	VII	Chilwa	21	Zomba	16/05/2010	-15.486	35.2364	Yes
28	8.04	VII	Chilwa	21	Zomba	16/05/2010	-15.486	35.2364	Yes
46	30.02	VII	Rufiji	6	Kidodo	02/08/2017	-7.4848	37.0285	Yes
57	(18)114	VII	Rovuma	19	Kwitanda	17/08/2013	-10.8473	37.4736	Yes
77	13.04 (B)	VII	Rovuma	18	Maposeni	06/09/2012	-10.7017	35.396	Yes
31	71.26	VIII	Luangwa	16	Chipata	12/07/2010	-13.7064	32.4897	Yes
38	51.05	VIII	Chambeshi	9	Mwelalo	17/06/2010	-12.1599	31.2318	Yes
39	51.06	VIII	Chambeshi	9	Mwelalo	17/06/2010	-12.1599	31.2318	Yes
40	51.07	VIII	Chambeshi	9	Mwelalo	17/06/2010	-12.1599	31.2318	Yes
49	3.01	VIII	Rukwa	5	Vwawa	23/07/2017	-9.03482	32.949	Yes
50	3.02	VIII	Rukwa	5	Vwawa	23/07/2017	-9.03482	32.949	Yes
51	3.03	VIII	Rukwa	5	Vwawa	23/07/2017	-9.03482	32.949	Yes
41	11.01	IX	Rukwa	2	Itaka	25/07/2017	-8.9141	32.8204	Yes
42	11.02	IX	Rukwa	2	Itaka	25/07/2017	-8.9141	32.8204	Yes
52	3.06	IX	Rukwa	5	Vwawa	23/07/2017	-9.0348	32.949	Yes
58	27.01	IX	Rukwa	1	Legeza	30/07/2017	-7.3097	31.0602	Yes
60	27.03	IX	Rukwa	1	Legeza	30/07/2017	-7.3097	31.0602	Yes
68	11.11	IX	Rukwa	2	Itaka	25/07/2017	-8.9141	32.8204	Yes
73	12.01	IX	Rukwa	3	Mlowo	25/07/2017	-8.9132	32.8472	Yes
87	1.03	IX	Rukwa	4	Mbeya	01/09/2012	-8.8994	33.3265	Yes
15	13.07 (A)	X	Ruo	23	Muluvia	19/05/2010	-16.0004	36.3207	Yes
29	71.02	XI	Luangwa	16	Chipata	12/07/2010	-13.7064	32.4897	Yes
30	71.16	XI	Luangwa	16	Chipata	12/07/2010	-13.7064	32.4897	Yes
32	71.29	XI	Luangwa	16	Chipata	12/07/2010	-13.7064	32.4897	Yes
37	51.02	XI	Chambeshi	9	Mwelalo	17/06/2010	-12.1599	31.2318	Yes
71	44.15	XI	Luangwa	15	Chitiwa	15/06/2010	-13.7868	28.9997	Yes
91	8X.03	XI	Ruaha	8	Ruaha	04/09/2012	-8.85472	34.08608	No

3.3 Methods and materials

Sample collection

The *Chiloglanis* samples, as used by Watson (2020), were collected from rocky and reedy regions of flowing streams and rivers in Malawi, Zambia and Tanzania, using steel-framed D-shape handnets with 4mm black fryma mesh. The nets were placed downstream from disturbed substrates typical of *Chiloglanis* habitats. Following collection, fish were anaesthetised with clove oil and placed in absolute ethanol for long-term storage. A total of 24 sites throughout South-East Africa were sampled between 2010 and 2017 (Table 3.1).

Linear morphometric measurements

External morphology of specimens was measured using digital callipers to 0.01mm (results were rounded to nearest 0.1mm) under a light microscope. Linear morphometric measurements followed those described by Friel and Vigliotta (2008) on chiloglanidin species (Appendix 3.2) and were carried out by a single person. Measurements of the postcleithral process to occipital shield length (OSL), length of postcleithral process (LCP), occipital shield width (OSW), and mandibular tooth row width (MTRW) were not included due to common damage sustained on small and aged specimens. A total of 42 linear measurements were taken along the left side of each specimen. It was possible to measure 61 of the 64 specimens studied by Watson (2020); one specimen was damaged during the tissue removal for previous molecular work, two others were not located.

Measurements were \log_{10} transformed, and size-corrected standardised residuals were calculated from linear regressions of the corresponding variables on standard length (SL), using base R 3.6.2 (R Core Team, 2019). Of the 2562 measurements in the data, there were 11 missing values due to the individual specimens being damaged. To enable inclusion of these individuals, the missing data were interpolated using Bayesian PCA, in the *pcaMethods* package (Stacklies *et al.*, 2007) in R 3.6.2.

An initial analysis of morphometric data for all 42 traits across the 61 specimens were conducted using Principal Component Analysis (PCA) based on a correlation matrix in PAST 4.0 (Hammer *et al.*, 2001). PC scores were visualized using bivariate plots, and variables contributing most to PC variation were identified using their respective loadings. Pairs of clades which are known to co-occur (Table 3.2) were then studied separately using PCA, and

traits which did not overlap between clades were identified. Finally, individual clades found at multiple sampling sites were studied separately using PCA.

To summarise the extent that morphometric variation was due to differences among clades relative to allopatric populations within clades, an Anova (Type II) was conducted using PC1 and PC2 scores from the analysis across 61 specimens as response variables, using the `lm` function in base R and the `car` package (Fox and Weisberg, 2019) in R 3.6.2.

Table 3.2. Genetic clades recovered, by site (From Watson, 2020). See Appendix 3.1 for map.

Site	Latitude °	Longitude °	Catchment	Clades recovered
1	-7.3097	31.0602	Rukwa	IX
2	-8.9141	32.8204	Rukwa	IX
3	-8.9132	32.8472	Rukwa	IX
4	-8.8994	33.3265	Rukwa	IX
5	-9.0348	32.9490	Rukwa	VIII, IX
6	-7.4848	37.0285	Rufiji	VI, VII
7	-8.7694	34.3748	Ruaha	V
8	-8.8547	34.0861	Ruaha	V, XI
9	-12.1599	31.2318	Chambeshi	VIII, XI
10	-9.3954	33.8273	Malawi	IV
11	-11.0196	33.7857	Malawi	IV
12	-12.2722	33.4878	Malawi	IV
13	-12.8337	34.1623	Malawi	I
14	-14.1795	34.1245	Malawi	I
15	-13.7868	28.9997	Luangwa	III, XI
16	-13.7064	32.4897	Luangwa	VIII, XI
17	-13.7587	32.4498	Luangwa	III
18	-10.7017	35.3960	Rovuma	VII
19	-10.8473	37.4736	Rovuma	VII
20	-15.2791	35.4011	Chilwa	VII
21	-15.4860	35.2364	Chilwa	VII
22	-15.8463	35.1932	Ruo	II
23	-16.0004	35.3207	Ruo	II, X

3.4 Results

The PCA morphometric data of all 61 individuals (summarised in Table 3.3) categorised by clade showed substantial overlap among clades, albeit with some clades separating (Fig. 3.2). PC1 accounted for 16.71% of overall variation with head length (HL), head depth (HD) and snout length (SNL) contributing most to the variation observed along the axis. PC2 accounted for 10.86% of overall variation with anal-fin length (ANFL), anal-fin base length (ANFBL) and oral disc width (ODW) contributing the most to the variation observed along the axis.

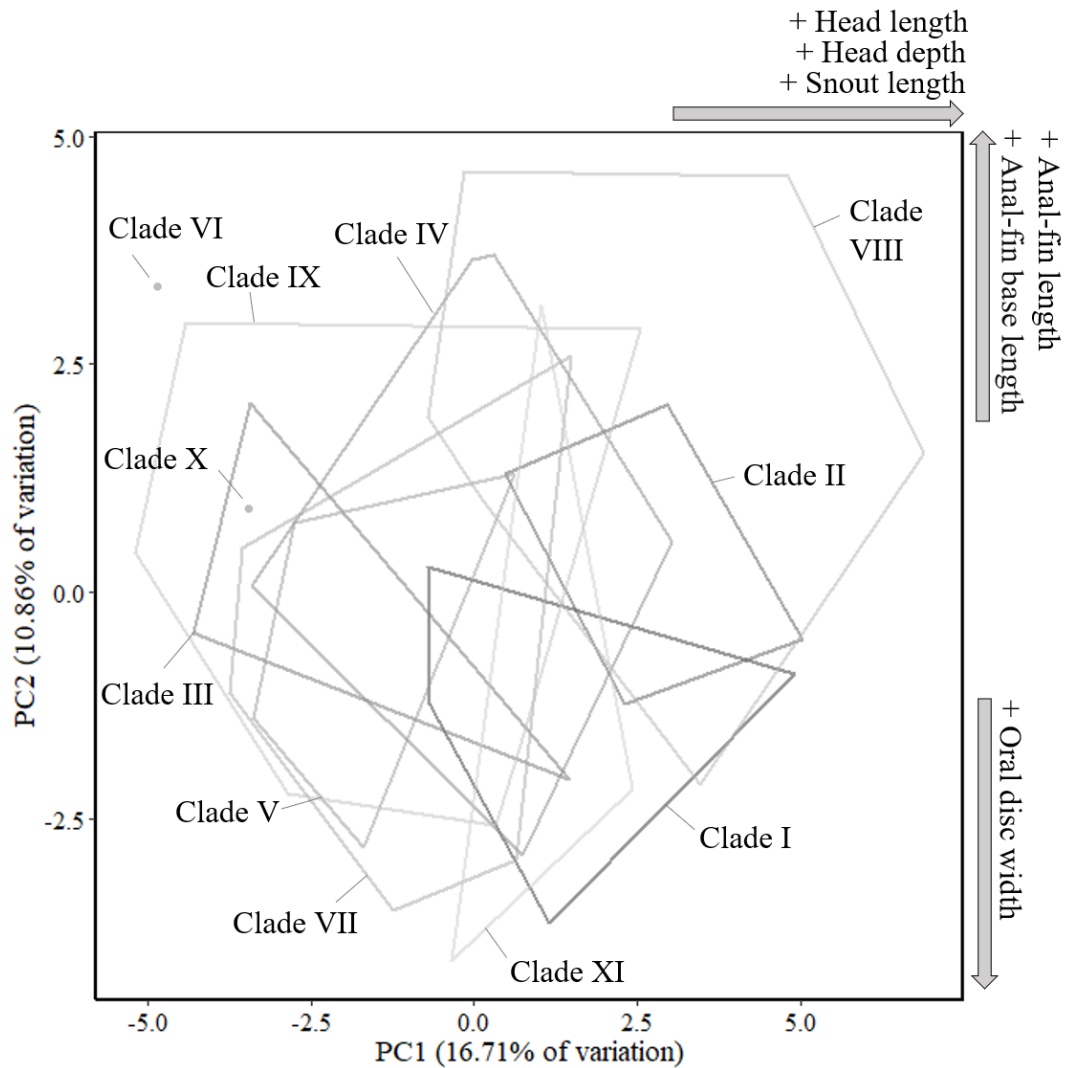


Fig. 3.2. Principal Component Analysis of 42 measurements from 61 specimens across 11 clades. Measurement variables contributing to most variation along PC1 and PC2 are included.

Table 3.3. Morphometric measurements for 61 *Chiloglanis* specimens. Standard length and total length expressed in mm. All other measurements are expressed in %SL.

Measurement	Clade I		Clade II		Clade III		Clade IV		Clade V		Clade VI		Clade VII		Clade VIII		Clade IX		Clade X		Clade XI	
	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean
Standard length (SL) (mm)	24.4-57.8	38.2	25.2-36.0	30.5	27.8-44.7	33.9	24.8-48.2	35.9	29.3-44.5	33.5	-	52.1	24.6-59.6	39.5	24.0-48.1	38.4	23.2-47.6	29.2	-	25.3	32.0-45.1	40.0
Total length (TL) (mm)	31.5-69.6	47.2	32.5-46.1	38.6	33.0-55.0	41.3	31.3-58.9	43.7	35.5-55.1	41.2	-	63.3	28.2-72.2	48.3	30.2-60.2	48.5	27.4-57.5	35.2	-	29.7	40.5-55.5	49.7
Head depth (HD) (maximum)	17.2-22.4	19.6	18.2-21.2	19.4	16.3-17.1	16.7	16.2-22.3	18.7	16.7-19.1	17.8	-	18.0	14.4-18.4	16.9	16.1-19.5	18.0	15.1-20.2	18.1	-	17.1	15.2-17.7	16.5
Snout length (SNL)	19.2-20.3	19.7	19.3-22.2	20.9	17.0-20.6	18.3	16.4-21.1	19.0	19.9-21.5	20.4	-	16.1	17.0-21.7	19.6	17.8-23.6	20.3	16.3-21.1	19.0	-	18.6	19.3-22.0	21.0
Head length to opercular membrane margin (HL)	30.8-35.2	32.4	30.4-36.1	33.2	30.3-33.0	31.4	30.2-35.2	31.9	31.5-33.9	32.3	-	27.6	30.4-34.0	31.9	30.8-37.7	34.0	30.5-35.0	32.8	-	30.5	33.1-35.6	34.2
Predorsal length (PDL)	36.0-40.0	37.3	37.2-40.2	38.7	32.3-40.4	35.5	33.1-40.8	35.9	33.6-38.1	36.0	-	34.9	35.6-39.3	37.6	36.7-42.4	39.1	33.1-39.0	36.3	-	38.4	37.7-41.9	40.4
Body depth at dorsal-fin insertion (BDDF)	18.0-21.0	20.1	20.4-23.2	22.1	18.0-19.8	19.1	17.9-23.0	20.5	17.4-20.6	19.0	-	19.6	14.3-20.1	18.2	17.8-21.9	19.7	17.2-23.0	20.0	-	21.9	17.2-21.1	19.1
Dorsal-fin base length (DFBL)	11.4-15.6	13.8	11.8-17.2	15.1	15.3-16.2	15.7	11.7-16.1	14.1	13.6-17.1	15.2	-	14.9	12.4-17.9	14.7	12.0-17.5	15.1	11.8-19.6	17.1	-	16.8	14.0-16.0	15.1
Dorsal-spine length (DSL)	14.6-18.2	16.7	15.0-17.5	16.4	15.6-18.8	16.8	14.0-20.7	17.5	15.5-19.4	17.3	-	20.3	13.1-18.8	15.6	14.0-19.8	16.8	10.5-17.3	14.8	-	-	16.8-21.3	18.8
Dorsal-fin length along longest ray (DFL)	18.2-23.3	20.5	17.8-23.3	21.0	22.3-24.0	23.3	17.9-22.8	19.7	17.6-20.5	19.0	-	21.5	14.7-22.7	17.9	17.0-21.9	19.5	16.4-22.9	19.8	-	16.6	18.2-22.1	20.2
Dorsal fin to adipose fin length (DF-ADFL)	21.1-26.5	24.3	23.9-29.1	25.6	22.9-29.1	25.0	20.3-29.0	25.6	25.8-26.9	26.5	-	25.8	22.1-27.7	25.0	23.1-29.3	26.4	21.5-30.5	25.8	-	23.3	22.4-28.3	24.2
Body depth at anus (BDA)	15.4-18.6	16.8	16.8-20.7	18.4	14.2-15.2	14.9	12.9-19.1	17.0	13.4-16.5	14.7	-	15.6	14.7-18.4	16.6	15.8-20.9	17.7	15.0-19.5	16.7	-	16.8	14.9-17.4	16.2
Adipose-fin base length (ADFBL)	15.5-21.2	18.6	15.4-20.7	18.0	14.1-19.1	16.2	14.4-21.2	18.3	13.5-22.7	18.6	-	14.9	13.7-22.7	17.8	13.9-19.0	16.8	17.2-22.3	19.3	-	18.0	15.8-17.3	16.5
Adipose-fin height (ADFH)	4.0-6.6	5.4	3.6-4.9	4.4	5.1-5.6	5.3	3.8-6.5	5.3	3.0-4.5	3.9	-	5.1	3.8-5.5	4.6	4.8-8.3	5.9	3.9-6.2	5.0	-	5.8	4.9-5.6	5.3
Anal-fin base length (ANFBL)	10.3-15.2	12.5	11.3-16.8	13.7	10.1-18.5	13.6	12.5-16.9	14.0	12.4-16.7	13.9	-	11.8	10.8-16.9	13.9	12.7-17.3	15.0	13.7-18.3	15.8	-	17.7	11.9-16.4	13.5
Anal-fin length along longest ray (ANFL)	12.9-19.7	15.3	13.0-19.3	16.3	15.2-15.3	15.2	13.2-19.9	16.1	14.1-17.9	15.7	-	15.3	12.9-19.3	15.7	14.5-18.9	16.7	13.8-19.2	16.3	-	14.9	12.8-16.5	14.5
Adipose fin to caudal peduncle length (AD-CPL)	13.0-18.1	15.5	15.9-19.7	17.3	14.7-18.3	16.8	14.6-17.9	15.8	16.3-18.5	17.0	-	17.2	15.3-18.6	16.3	16.3-20.7	18.6	15.0-18.7	16.8	-	17.7	13.1-18.5	16.3
Caudal peduncle length (CPL)	15.0-20.5	18.4	15.5-19.6	17.5	16.3-20.5	18.4	17.4-22.5	19.3	20.7-22.9	22.0	-	19.6	16.4-23.2	19.3	17.3-21.0	19.5	17.8-22.1	20.3	-	21.8	17.2-19.4	18.4
Caudal peduncle depth (CPD)	11.8-14.2	12.9	13.3-14.9	14.2	11.3-12.6	12.0	11.0-15.3	13.1	10.4-11.6	11.2	-	12.3	10.6-14.0	12.1	11.7-14.3	12.8	11.0-14.2	12.8	-	10.8	10.7-13.3	11.8
Caudal fork length (CFKL)	8.3-17.2	13.1	12.8-18.5	16.2	12.3-16.7	14.7	12.3-19.8	14.7	10.0-18.3	14.2	-	11.7	12.7-19.6	14.3	15.9-24.0	19.3	13.4-21.5	18.1	-	15.9	13.3-20.9	16.1
Caudal-fin length (CFNL)	22.5-27.0	25.3	28.1-33.8	30.5	22.8-27.8	26.1	25.1-30.5	27.0	24.5-28.9	26.3	-	24.4	19.4-30.5	26.1	27.6-32.8	29.8	21.6-27.5	25.8	-	24.4	24.5-30.9	28.6
Preanal length (PANL)	66.2-72.5	69.2	68.9-73.1	70.7	65.1-76.9	69.3	65.1-72.8	70.0	61.4-67.6	65.5	-	70.1	64.2-76.1	70.4	59.6-71.2	67.1	63.7-71.6	67.1	-	66.4	66.4-73.9	70.4
Prepectoral length (PPTL)	30.4-31.4	30.9	29.3-32.3	30.4	25.3-32.3	28.5	26.9-37.0	30.6	27.8-31.5	29.5	-	25.8	27.3-30.9	29.2	29.1-32.9	30.8	27.7-32.6	29.9	-	28.3	30.1-33.7	31.3

Table 3.3. continued. Morphometric measurements for 61 *Chiloglanis* specimens. Standard length and total length expressed in mm. All other measurements are expressed in %SL.

Measurement	Clade I		Clade II		Clade III		Clade IV		Clade V		Clade VI		Clade VII		Clade VIII		Clade IX		Clade X		Clade XI	
	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean
Width at pectoral-fin insertion (WPTFI)	21.2-31.2	25.3	20.3-26.4	24.0	21.1-25.4	23.8	19.2-24.5	22.3	20.1-25.5	22.5	-	20.3	21.1-23.9	22.2	23.2-28.0	24.9	19.6-26.9	23.2	-	22.5	21.2-24.7	22.9
Pectoral-spine length (PSL)	16.1-20.1	17.7	12.8-21.9	18.3	16.1-19.8	18.1	15.5-22.7	19.6	17.7-20.9	19.2	-	22.1	15.9-20.5	18.3	18.0-22.4	20.4	12.2-17.5	16.2	-	-	16.7-22.1	19.6
Pectoral-fin length (PTFL)	19.2-25.2	23.1	18.4-25.8	22.1	17.5-23.0	21.1	16.0-26.4	22.6	19.6-22.6	21.4	-	22.2	18.8-25.8	21.3	21.3-26.0	23.2	19.3-22.3	20.7	-	20.0	20.3-24.2	21.4
Prepelvic length (PPVL)	55.7-62.4	58.1	55.2-59.6	57.5	55.3-57.9	56.8	50.3-59.7	56.5	50.5-56.5	53.5	-	58.4	49.2-61.4	56.9	55.4-59.7	57.6	53.2-60.8	57.1	-	55.2	54.8-61.2	58.6
Pelvic-fin interspace (PVI)	4.4-7.2	5.4	5.5-7.8	6.2	3.8-4.7	4.2	4.0-5.4	4.8	4.0-5.2	4.8	-	4.3	4.0-5.6	4.8	4.6-7.1	5.5	4.7-6.0	5.2	-	5.1	4.0-5.5	4.7
Pelvic-fin length (PVFL)	12.2-14.7	13.7	12.7-15.1	14.0	12.8-14.9	13.8	12.7-16.7	14.7	13.3-15.1	14.3	-	13.4	12.3-17.9	14.5	13.8-16.1	14.9	11.0-13.4	12.2	-	13.4	13.9-15.1	14.4
Anterior nares interspace (ANI)	5.5-6.9	6.1	5.4-6.8	6.3	5.6-6.4	6.1	5.0-7.5	6.1	5.4-6.2	5.7	-	4.9	5.2-7.1	5.8	5.3-7.9	6.4	5.4-6.4	6.0	-	6.4	5.3-7.1	5.9
Posterior nares interspace (PNI)	5.1-6.5	5.6	4.8-6.8	6.0	4.3-5.6	5.0	4.5-6.8	5.5	5.2-5.8	5.4	-	4.9	4.2-6.3	5.2	4.8-7.3	6.0	4.9-6.2	5.5	-	5.7	4.7-6.5	5.4
Orbital interspace (OBI)	8.5-11.1	9.6	10.0-11.7	10.8	9.5-10.9	10.0	9.0-11.4	9.8	8.1-10.0	9.4	-	7.6	7.7-11.4	8.8	8.6-11.9	9.8	10.0-19.0	11.7	-	9.3	8.6-9.9	9.2
Eye diameter (horizontal axis) (EDH)	4.6-7.0	5.9	5.1-7.3	6.0	4.8-6.2	5.6	4.6-6.9	5.8	4.9-6.5	5.9	-	5.9	4.7-7.7	5.9	4.8-7.6	5.7	4.7-7.6	6.5	-	7.4	4.9-6.9	6.2
Eye diameter (vertical axis) (EDV)	4.3-6.7	5.4	4.7-7.0	5.5	4.5-5.7	5.2	4.0-6.3	5.2	4.4-6.0	5.5	-	5.2	3.7-6.6	4.9	4.0-6.7	5.2	4.2-6.5	5.7	-	6.6	4.5-5.9	5.3
Oral disc length (ODL)	16.5-20.8	17.5	13.1-19.3	16.3	15.1-16.9	16.0	14.8-19.0	16.1	15.0-17.4	15.9	-	11.8	13.6-19.4	16.4	14.2-21.4	17.1	11.9-18.1	15.4	-	15.9	15.6-20.9	17.6
Oral disc width (ODW)	16.7-21.5	19.2	17.1-22.0	19.1	18.3-19.7	18.9	15.2-19.1	17.5	16.5-21.1	19.2	-	15.5	17.4-19.7	18.7	15.1-19.0	17.7	15.2-20.8	17.7	-	14.6	16.5-19.2	17.8
Premaxillary tooth-patch width (PMXW)	13.2-16.3	14.8	13.9-15.8	14.8	14.1-15.2	14.7	12.1-15.7	14.1	13.5-16.3	15.2	-	12.2	14.4-15.7	15.2	12.4-16.0	14.5	12.1-15.3	13.8	-	11.6	11.9-14.5	13.8
Mouth width (MW)	8.3-10.2	9.5	8.3-11.1	10.2	9.2-10.3	9.6	8.0-13.1	10.0	8.6-11.2	10.5	-	8.7	8.5-10.9	9.6	9.6-11.5	10.6	8.7-10.8	9.9	-	11.8	10.2-11.1	10.7
Lower lip length (LLL)	6.1-9.4	8.2	6.1-8.8	7.5	5.3-8.5	6.6	6.0-9.7	7.4	8.7-10.4	9.3	-	6.4	5.3-8.7	7.8	5.9-8.7	7.4	6.9-8.8	7.6	-	5.6	5.4-7.6	6.7
Upper lip length (ULL)	3.5-5.1	4.4	4.1-5.0	4.5	4.1-4.8	4.5	3.2-5.4	4.2	3.6-5.0	4.1	-	2.7	3.2-4.9	4.1	3.4-4.8	4.2	3.8-5.1	4.7	-	4.6	3.8-4.9	4.3
Premaxillary tooth-patch length (PMXL)	4.6-5.6	4.9	4.1-5.5	4.8	4.4-5.2	4.8	3.7-5.9	4.9	4.5-5.2	4.9	-	4.0	4.6-5.4	5.0	4.4-5.6	4.9	4.2-6.4	5.4	-	5.5	3.9-5.5	4.9
Maxillary barbel length (MXBL)	4.6-12.1	7.8	5.6-8.3	7.3	6.0-7.1	6.7	6.0-10.2	7.9	8.7-9.5	9.2	-	8.2	5.3-9.0	7.0	5.0-9.1	7.1	5.9-11.1	8.3	-	6.4	4.2-11.2	7.5

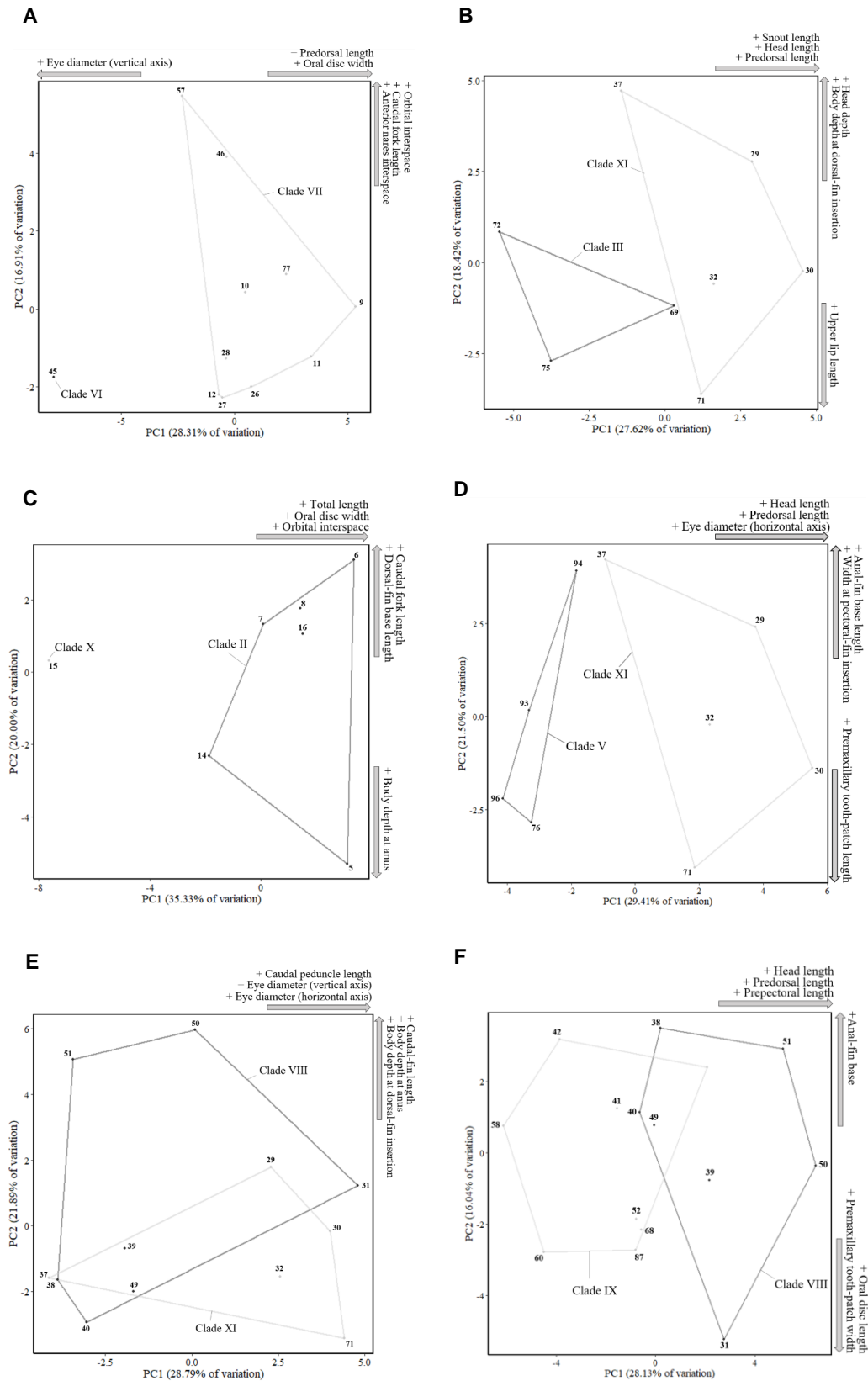


Fig. 3.3. PC1 versus PC2 from Principal Component Analyses of morphometric measurements of specimens from sympatric clades. A) Clades VI and VII. B) Clades III & XI. C) Clades II & X. D) Clades V and XI. E) Clades VIII and XI. F) Clades VIII and IX.

PCAs conducted on morphometric data of different clades that have sympatric representatives showed that most (four of six) sympatric clades were separated on the first two PC axes. These included specimens from Clades III and XI that co-occur in the Luangwa catchment, Clades II and X that co-occur in the Ruo catchment, Clades V and XI that co-occur in Ruaha drainage, and Clades VI and VII that co-occur in Rufiji drainage (Fig. 3.3). By contrast, PCA indicated more substantial morphological overlap between representatives of Clades VIII and XI that co-occur in the Chambeshi catchment, and between Clades VIII and IX that co-occur in the Luangwa drainage. In comparisons of individual measurements between clades that co-occur in sympatry, it was possible to identify non-overlapping measurements that have potential to enable separation of clades in most cases (Table 3.4). Only Clades VIII and XI that co-occur in the Chambeshi catchment had no non-overlapping measurements.

Table 3.4. Measurements that do not overlap between sympatric clades (size standardised residuals, \log_{10} transformed data).

Clade Comparison x y	Morphometric measurement	Clade x Range	Clade y Range
III XI	Body depth at anus (BDA)	-0.061 : -0.044	-0.041 : 0.022
	Head length (HL)	-0.035 : 0.014	0.015 : 0.038
	Mouth width (MW)	-0.042 : 0.0017	0.020 : 0.050
VIII IX	Caudal-fin length (CFNL)	0.019 : 0.093	-0.084 : 0.0010
	Pelvic-fin length (PVFL)	-0.0033 : 0.064	-0.11 : -0.024
V XI	Adipose-fin height (ADFH)	-0.19 : -0.062	0.023 : 0.066
	Caudal peduncle length (CPL)	0.047 : 0.066	-0.030 : 0.0085
	Lower lip length (LLL)	0.064 : 0.15	-0.063 : 0.0054
II X	Adipose fin to caudal peduncle length (AD-CPL)	0.001 : 0.033	-0.044
	Anal-fin base length (ANFBL)	-0.099 : 0.098	-0.113
	Body depth at anus (BDA)	-0.045 : 0.096	-0.061
	Body depth at dorsal-fin insertion (BDDF)	-0.018 : 0.053	-0.023
	Caudal peduncle depth (CPD)	-0.026 : 0.071	-0.036
	Dorsal fin to adipose fin length (DF-ADFL)	-0.028 : 0.063	-0.041
	Dorsal-fin base length (DFBL)	-0.099 : 0.025	0.064
	Dorsal-fin length (DFL)	-0.048 : 0.081	0.095
	Dorsal-spine length (DSL)	-0.045 : 0.040	0.058
	Lower lip length (LLL)	-0.151 : 0.019	0.052
	Oral disc width (ODW)	-0.076 : 0.018	0.038
	Preanal length (PANL)	-0.017 : 0.017	0.037
	Prepectoral length (PPTL)	-0.034 : 0.027	0.037
	VI VII	Adipose-fin height (ADFH)	0.054
Anterior nares interspace (ANI)		-0.056	-0.038 : 0.043
Body depth at dorsal-fin insertion (BDDF)		0.022	-0.109 : 0.020
Dorsal-fin length (DFL)		0.052	-0.121 : 0.50
Dorsal-spine length (DSL)		0.096	-0.098 : 0.050
Eye diameter (horizontal axis) (EDH)		0.087	-0.029 : 0.051
Eye diameter (vertical axis) (EDV)		0.085	-0.082 : 0.026
Head length (HL)		-0.060	-0.030 : 0.024
Oral disc length (ODL)		-0.153	-0.069 : 0.068
Oral disc width (ODW)		-0.066	-0.017 : 0.037
Predorsal length (PDL)		-0.034	-0.015 : 0.022
Premaxillary tooth-patch length (PMXL)		-0.050	-0.016 : 0.056
Premaxillary tooth-patch width (PMXW)		-0.072	0.000 : 0.037
Prepectoral length (PPTL)		-0.056	-0.035 : 0.011
Pectoral-spine length (PSL)		0.037	-0.065 : 0.017
Snout length (SNL)		-0.086	-0.061 : 0.044
Width at pectoral-fin insertion (WPTFI)		-0.056	-0.040 : 0.015

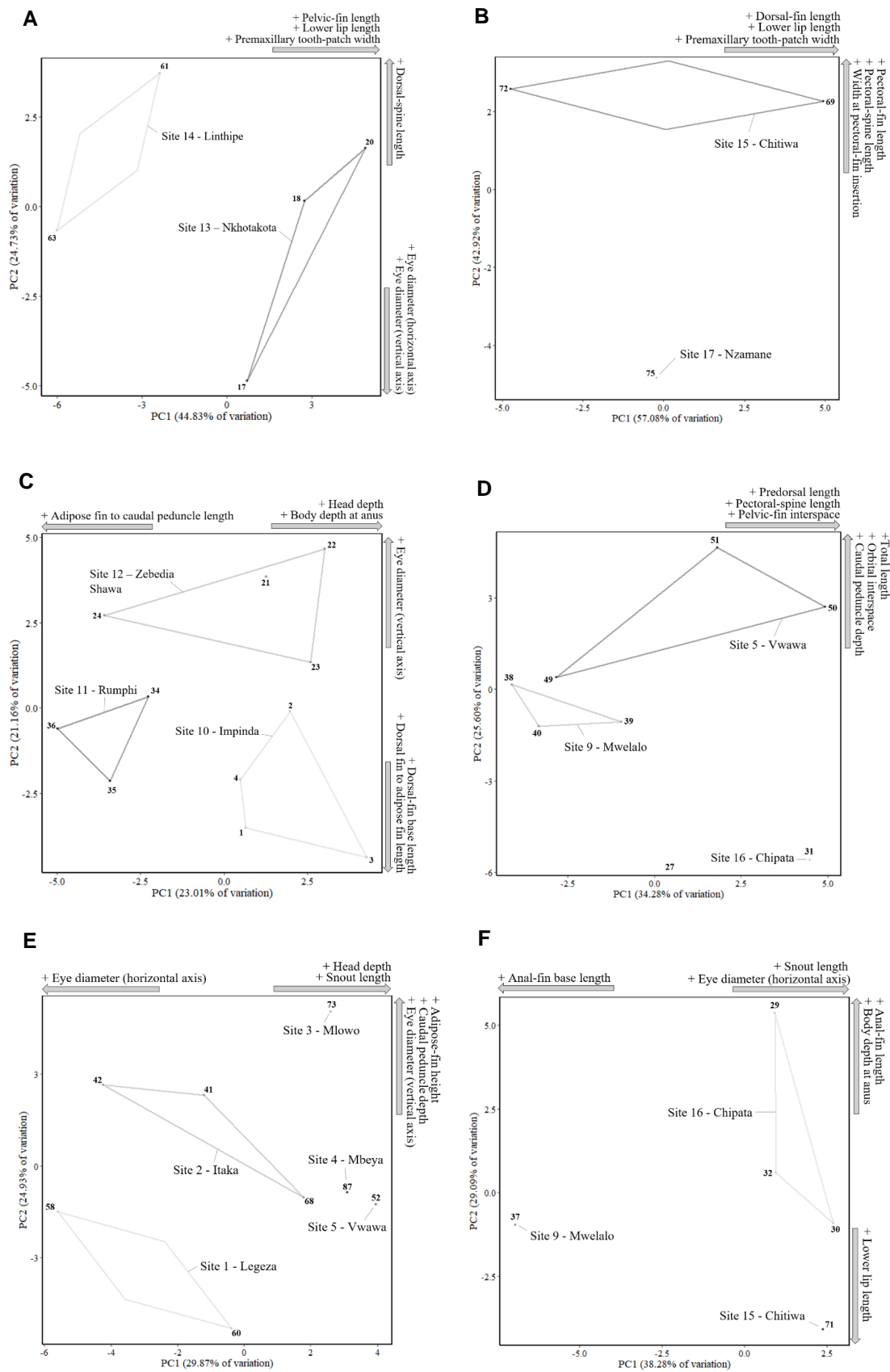


Fig. 3.4. PC1 versus PC2 from morphometric measurements of allopatric populations of the same clade. A) Clade I. B) Clade III. C) Clade IV. D) Clade VIII. E) Clade IX. F) Clade XI.

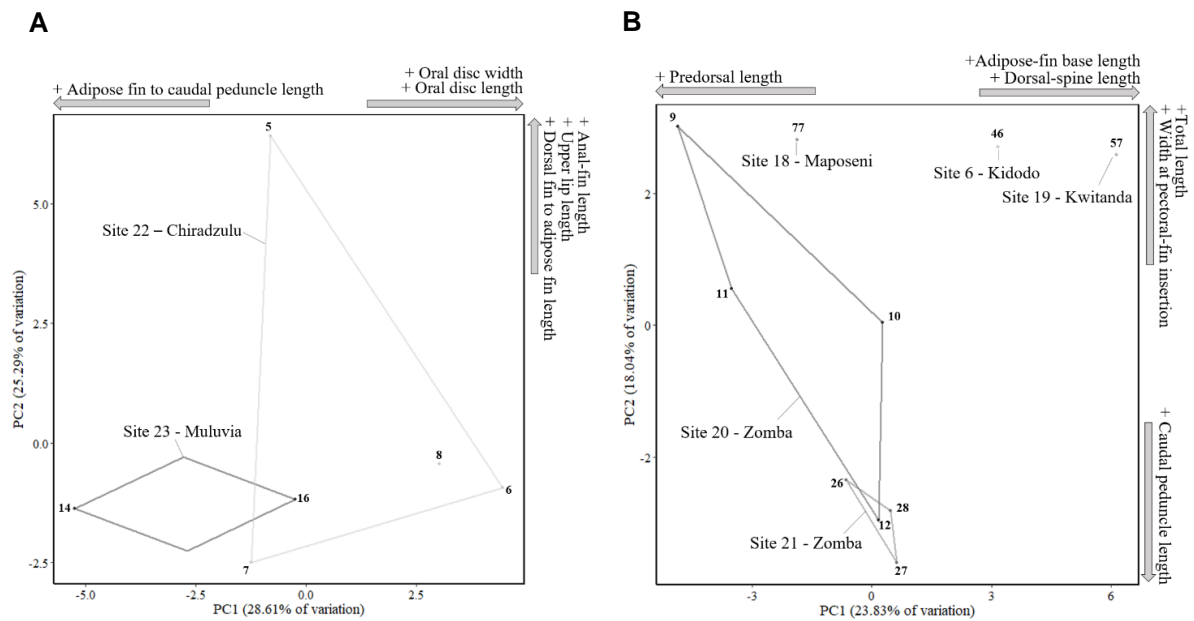


Fig. 3.5. Scatterplots of PC1 versus PC2 from Principal Component Analyses of morphometric measurements of allopatric populations of the same clade with overlapping morphological similarities. A) Clade II by sampling site. B) Clade VII by sampling site.

Eight of the clades included representatives from allopatric populations. In six of those clades the representatives showed evidence of morphological differences along Principal Component Axes 1 and 2 (Fig. 3.4). The remaining showed some evidence of morphological overlap among the allopatric populations from multiple sampling sites (Figure 3.5). However, it is notable that the allopatric populations that did show overlap in both cases were in extremely close proximity. Clade II overlapped in morphology between Site 22 and 23, which are 22km apart and within the same Ruo catchment, while Clade VII overlapped between site 20 and 21, which are 29km apart and within the same Chilwa catchment.

A summary Anova analysis, using PC1 and PC2 data from the PCA of all 61 individuals as response variables, demonstrated significant differences ($P < 0.05$) both among clades, and among sampling sites within clades (Table 3.5). The sum of squares from this analysis showed that 41.1% of the total variance was partitioned among clades, 33.2% of the total variance was partitioned among sampling sites within clades, and 26.6% residual variation (Table 3.5).

Table 3.5. Results of an Anova (Type II) conducted using PC1 and PC2 scores from the analysis on the morphometric data for all 42 traits across 61 specimens, quantifying the extent of variation observed i) among clades, and ii) among site populations within clades.

	<i>Degrees of freedom</i>	<i>Sum of Squares</i>	<i>F</i>	<i>P</i>
Clade	10	279.25	5.30	< 0.001
Clade: Site	17	225.21	2.52	0.011
Residuals	33	173.81		

3.5 Discussion

Morphological differences among clades

The results from this study clearly demonstrate a complex pattern where genetically-distinct *Chiloglanis* clades overlap substantially in external morphology, and there was no clear evidence for any clade being distinct on either of the two primary axes of morphological variation, PC1 and PC2. Therefore, these results support the concept that *Chiloglanis* species with extremely similar morphological traits are present in the study region. These results are similar to results from Schmidt *et al.* (2015) of *Chiloglanis* populations from Kenya, who found that while some species could be distinguished from other on the basis of morphological characters, many genetically distinct species have very similar morphology. Such patterns are not unusual for riverine catfishes, and similar conserved phenotypes have been reported from other rheophilic lineages including the African genus *Amphilius* (Schmidt and Pezold, 2011), and the South American genus *Microglanis* (Souza-Shibatta *et al.*, 2018.) It is plausible that the conserved phenotypes are maintained in these groups, despite speciation, as a consequence of strong stabilising selection on ecologically-important phenotypic traits that promote survivorship in flowing waters.

Despite the considerable overlap among many of the clades, a high proportion of the variance was partitioned among clades. Specifically, the Anova (Type II) conducted on PC1 and PC2 scores from the analysis of all morphometric data across 61 specimens shows that there is approximately 41% variance partitioned among clades. Differences in characters such as head length, head depth, snout length and oral disc width contributed most substantially to overall variation on axes PC1 and PC2. Notably, similar results suggesting craniofacial characters are important in distinguishing *Chiloglanis* species have been reported from sampling sites in

northwestern Tanzania (Friel and Vigliotta, 2011). Collectively, such differences may point to differences in the ecology of these species, if head morphological variation is functional in enabling occupation of different flow regime substrates or use of different food resources. To determine whether the observed variation is an adaptation to differing habitats would require information on stream characteristics and food resources.

There were multiple examples where there were sympatric representatives of different clades, each with distinct morphology. Given the extent of genetic divergence observed, and evidence that these were not sister taxa, then it appears likely these species have undergone secondary contact and remain reproductively-isolated. Ecological theory predicts that stable coexistence can only succeed following divergence in traits related to their ecology and mating systems (Kusche *et al.*, 2015), and therefore it is possible that morphological traits differing between sympatric clades may be linked to niche partitioning. Microhabitat niche differences between coexisting *Chiloglanis* in West African river species have been previously documented (Schmidt *et al.*, 2016). Specifically, *Chiloglanis* cf. *micropogon* was collected among fast flowing riffles over pebbles and cobble, whereas *Chiloglanis* sp. “Senegal/Niger” was collected among woody debris in a lower flow. Detailed study of microhabitat use, diet and trophic morphology are required to further explore niche divergence among both sympatric and allopatric *Chiloglanis* clades in this study.

Anal-fin length and anal-fin base length were also contributing traits to variation along PC2. Such traits may also be driven by differences in locomotory performance among habitats. However, it is known that adult males of the *Chiloglanis* genus commonly display elongate anal and/or caudal fins (Morris *et al.*, 2015), and the exaggeration of male fish fin size and colour is thought to be driven by sexual selection as opposed to natural selection (Morris *et al.*, 2015). In this study the individuals have not been sexed, and hence it has not been possible to quantify the extent of sexual dimorphisms present within clades. Future work would benefit from an information on both the gender and maturity of focal individuals.

Differences within clades

The results demonstrated individuals the same clade collected from different locations tend to have diverged in morphology characters. Population-level analyses suggest strong spatial genetic structure within *Chiloglanis* species (Morris *et al.*, 2015), so it is likely that the samples from geographically-distant sites exhibit low levels of gene flow. Hence, it is possible that any

morphological differences among populations is a consequence of either adaptation to local habitats, or a consequence of genetic drift. Alternatively, the differences may be a consequence of phenotypic plasticity. Intriguingly, the only clades exhibiting morphological overlap between sampling sites were from clades II and VII, and those sites showing overlap were in close geographic proximity (< 25km) and in the same catchments. This may suggest that within-species genetically-closer populations are more likely to exhibit similar phenotypes, or alternatively it is possible that the close geographic proximity may also indicate greater ecological similarity of habitats, favouring similar phenotypes. Experimental work studying development in controlled common-garden conditions will be required to disentangle the roles of rearing environment and genomic divergence for morphology (e.g. Vrtílek and Reichard, 2016).

Practical implications for identifying species

On the basis of the combined molecular and morphological results, it would be possible to formally describe the many of the eleven clades as distinct species. Recognition of these taxa, and their cryptic nature, has implications for their conservation and contribution to biodiversity. Based on morphological data alone, the *Chiloglanis* specimens in this study may in the field be classed as a single clade or species, with a broad distribution and a threat risk of “Least Concern”. However, descriptions based on both genetic and morphological evidence would highlight the restricted distributions of these clades, elevating their conservation status.

Further systematically focused work is now needed to strengthen the evidence for these clades representing distinct species. In the first instance it would be valuable to identify genetic “barcode” markers that could be used to reliably and rapidly used to distinguish sympatric taxa. These could be from the use of an existing barcode gene (e.g. COI), or mining RAD sequence data to identify sets of distinguishing nuclear markers. It would then be possible to further screen existing and new collections to boost sample sizes for morphological analysis, with a view to testing existing potentially diagnostic combinations of traits revealed in this study (e.g. Table 3.4), but also studying traits not directly considered in this study, including coloration, patterning and trophic morphology.

Concluding remarks

The morphometric traits and their linear measurements used in this study have provided strong support for the presence of cryptic species within *Chiloglanis* from South-East Africa. It is likely these cryptic species have developed as a consequence of strong stabilising selection on the overall fluvial phenotype which promotes survivorship in this genus. Local adaptation in allopatric populations and niche partitioning (and possible behavioural isolation) in sympatric populations may also have led to a decoupling of morphological divergence and phylogenetic divergence. Future research may benefit from more detailed investigations of variation of phenotypic traits not considered in this study, namely tooth and jaw morphology, and the complex colour patterning that is present among the sampled specimens.

CHAPTER FOUR: GENERAL DISCUSSION.

Through the course of this thesis, both molecular barcode and morphological methods have been assessed for potential contributions to our knowledge of the biodiversity of East African fishes. It was clear from the results that both methods have strong potential to enable an improved understanding of species diversity in the region.

4.1 DNA barcoding for resolution of East African freshwater fish biodiversity

From 122 specimens, 109 specimens were successfully sequenced in our study using the 12S or COI barcode. This means that 13 individuals were unsuccessful in the PCR or DNA sequencing process, plausibly as a result of DNA degradation of samples. However, it may also be possible that the primers used in the PCR process of these unsuccessful specimens were unsuitable for the respective species. This is commonly seen in barcoding studies, such as that conducted by Ward *et al.* when barcoding Australian fish species in 2005. In this study, a new internal forward primer was designed for *N. cepedianus* to obtain enough product to be sequenced. It may be possible that new primers would also enable sequencing of other species. In cases such as these, a reference database detailing the successful/unsuccessful use of a variety of primer sequences for individual species would prove valuable when saving time and resources in future DNA barcoding studies.

In total 81 specimens were successfully sequenced at the COI gene, while 83 specimens were successfully sequenced at the 12S gene, but these species were not always the same. This supports the notion that the most suitable barcode region for sequencing also differs between individual species. Future studies might investigate the variable success in barcode sequencing of known Tanzanian freshwater fishes using a larger range of markers. A reference database may then be produced to recommend the most reliable barcoding gene for a suspected species, after initial morphological identification. This could prove valuable in the fast and accurate bioidentification of freshwater fishes, both in conservation and also in the monitoring of commercial fisheries.

In order to allow sufficient mitigation procedures to be put in place as a result of infrastructure development in Tanzania, further research must be done to capture the full freshwater fish diversity. While general assumptions have previously been made concerning a decrease in

water quality in the planned reservoir and downstream of Stiegler's Gorge hydrodam, further research is necessary to further detail this impact on individual tributaries, and thus full species diversity present (Mwalyosi, 1988). Typically, impacts are assessed via Environmental Impact Assessments (EIAs), which evaluates the risks and effects likely to occur following a major project significantly affecting the environment (Jay *et al.*, 2007). However, while EIAs were carried out for Stiegler's Gorge, it appears they were not sufficiently detailed spatially nor in enough detail to capture the full impacts on fish biodiversity.

An emerging tool in fish conservation in recent years has been the use of eDNA for monitoring past and present biodiversity. The collection of eDNA data is described as an useful, non-invasive and easy-to-standardise alternative to traditional monitoring techniques, using genetic material obtained directly from environmental samples such as soil (Thomsen and Willerslev, 2015). With traditional methods often involving invasive surveying techniques, as well as displaying difficulties in the correct identification of cryptic species, eDNA provides a potential route for the future of biodiversity management at relatively low costs and effort. Use of eDNA in freshwater fish studies has led to previous uses in calculating fish biomass from freshwater samples (Takahara *et al.*, 2012), as well as estimates of species composition (Minamoto *et al.*, 2011). With eDNA proving successful in other studies, this technique should be considered in future large-scale biodiversity surveys of Africa's freshwater fishes. It is now clear that the most successful marker for fish eDNA metabarcoding is the 12S gene (Collins *et al.*, 2019), and this project has provided the core of a 12S reference library that would facilitate eDNA metabarcoding studies of Tanzanian freshwater fishes.

4.2 *The future of *Chiloglanis* study*

This study included only 61 *Chiloglanis* specimens, collected from across Zambia, Tanzania and Malawi. In order to study overall trends in morphological divergence in the *Chiloglanis* genus, a larger-scale study using samples from a greater variety of geographical locations would be required. As *Chiloglanis* is thought to have an almost ubiquitous continental distribution, this would allow comparisons to be made between different ecological habitat types, and perhaps demonstrate divergence in traits related to temperature, predator types and differing substrates (Day *et al.*, 2013). This is likely to occur due to the high endemism found among *Chiloglanis* species, which are exclusively rheophilic fishes that are unlikely to be able to disperse (Friel and Vigliotta, 2011).

During the study, it was apparent that there is scope for colour patterns and markings of *Chiloglanis* species to contribute to studies of phenotypic divergence. Similar conclusions were reached by Schmidt and Barrientos (2019), who found that colouration was a useful tool in their diagnosis of *C. productus*. While there was considerable variation among the representatives of the *Chiloglanis* clades of South-East Africa used in our study, the variability in these colour patterns would require intensive study and may benefit from the application of pattern analysis methods. The majority of *Chiloglanis* species formally described to date have very little colour information, with colours referred to as “mainly brownish with a lot of light & dark spots” (Seegers, 2008). It is likely that colour traits acting as a form of camouflage against predators (e.g. aerial), and that there may be convergence in patterns and coloration among species found in similar habitats and substrates.

An alternative to linear measurement is geometric morphometrics, which can offer considerable insight into variation of shape of fish species. Studies have found that species classification has the greatest success when using data gathered through geometric morphometrics compared to other methods, which is particularly useful in detecting small and localized changes in shape (Schmieder *et al.*, 2015). However, specimens to be used for geometric morphometrics of body shape require specimens to be carefully preserved to avoid bending and other shape distortion during preservation.

Use of traditional morphometrics alone, using linear measurements as used in this study, can miss key aspects of variation that may have taxonomic importance (Adams *et al.*, 2004). For example, sample 52 was noted as having a three-lobed caudal fin, which differentiates it from all other samples in the study. However, this detail would have been overlooked in the caudal fin measurements gathered (caudal fork length and caudal-fin length). Ng and Bailey (2006) noted the variation in *Chiloglanis* caudal fin shape and its link to sexual dimorphism, with the fin usually shallowly or deeply forked in shape. Measurements which do not document this variation may ignore potential differentiating traits between species. Thus, future studies of *Chiloglanis* may benefit from methods that more comprehensively capture aspects of phenotypic variation, than linear measurements alone.

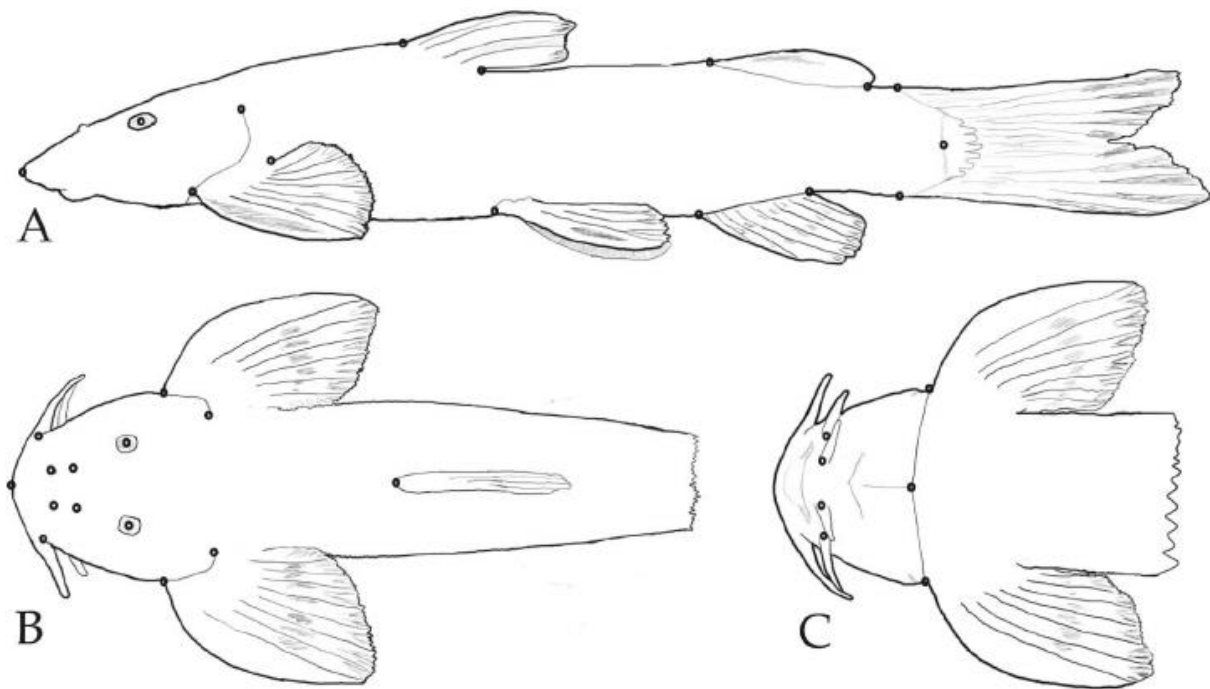


Fig 4.1. Diagram of a generalized *Amphilius* specimen, detailing the location of landmarks used in geometric analysis for (A) lateral, (B) dorsal and (C) ventral analysis. Landmark placement may also apply to *Chiloglanis* specimens in geometric analysis studies (from Schmidt and Pezold, 2011).

When external morphological measurement is unable to differentiate species, it may be possible that underlying differences in the skeletal and inner structures of fish could provide key systematically-relevant information. Computed tomography (CT) scans are now routinely used to investigate the otherwise unseen details of a specimen, allowing a 3D representation of their internal hard structures. Data can be gathered from these scans, including internal morphological measurement of bones. These internal structures are less likely to have been subjected to damage and degradation as a result of long preservation periods, in comparison to soft tissue which may be subject to shrinkage, for example of fish body mass and length (Santos *et al.*, 2009). Measurements of these structures are also less likely to be subject to human error when calculated digitally. There is considerable scope to study the internal morphology of *Chiloglanis* species, and this could offer valuable insight into species divergence.

4.3 Concluding remarks

A wide-scale project linking DNA barcoding with key morphological characters could lead to the development of reliable and accurate bio-identification of species in African rivers. It would allow for a greater appreciation of the true species richness of freshwater river systems, but also help to highlight cryptic species, such as those within *Chiloglanis* catfish. More reliable methods of identification would help in the mapping and monitoring of Africa's freshwater fishes, and barcode-informed eDNA surveys could prove useful in conservation initiatives. The results of further studies could be used to produce widely available, modern, and more reliable resources, offering alternatives to outdated sources of information that are widely used in biodiversity assessments within countries such as Tanzania.

References

- Adams, D., Rohlf, F. and Slice, D., 2004. Geometric morphometrics: Ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71(1), pp.5-16.
- Adams, D., West, M. and Collyer, M., 2007. Location-specific sympatric morphological divergence as a possible response to species interactions in West Virginia Plethodon salamander communities. *Journal of Animal Ecology*, 76(2), pp.289-295.
- Adeoba, M., Tesfamichael, S. and Yessoufou, K., 2019. Preserving the tree of life of the fish family Cyprinidae in Africa in the face of the ongoing extinction crisis. *Genome*, 62(3), pp.170-182.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990. Basic local alignment search tool. *Journal of molecular biology*, 215(3), pp.403-410.
- April, J., Mayden, R.L., Hanner, R.H. and Bernatchez, L., 2011. Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences*, 108(26), pp.10602-10607.
- Barluenga, M., Stölting, K., Salzburger, W., Muschick, M. and Meyer, A., 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, 439(7077), pp.719-723.
- Basiita, R.K., Zenger, K.R., Mwanja, M.T. and Jerry, D.R., 2018. Gene flow and genetic structure in Nile perch, *Lates niloticus*, from African freshwater rivers and lakes. *PloS ONE*, 13(7), p.e0200001.
- Becker, S., Hanner, R. and Steinke, D., 2011. Five years of FISH-BOL: brief status report. *Mitochondrial DNA*, 22(sup1), pp.3-9.
- Bhattacharya, M., Sharma, A.R., Patra, B.C., Sharma, G., Seo, E.M., Nam, J.S., Chakraborty, C. and Lee, S.S., 2016. DNA barcoding to fishes: current status and future directions. *Mitochondrial DNA Part A*, 27(4), pp.2744-2752.
- Bickford, D., Lohman, D., Sodhi, N., Ng, P., Meier, R., Winker, K., Ingram, K. and Das, I., 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), pp.148-155.
- Bingpeng, X., Heshan, L., Zhilan, Z., Chunguang, W., Yanguo, W. and Jianjun, W., 2018. DNA barcoding for identification of fish species in the Taiwan Strait. *PloS ONE*, 13(6), p.e0198109.
- Bradbeer, S., Harrington, J., Watson, H., Warraich, A., Shechonge, A., Smith, A., Tamatamah, R., Ngatunga, B., Turner, G. and Genner, M., 2018. Limited hybridization between introduced and Critically Endangered indigenous tilapia fishes in northern Tanzania. *Hydrobiologia*, 832(1), pp.257-268.
- Bwathondi, P.O.J. and Mwamsojo, G.U.J., The status of the fishery resource in the wetlands of Tanzania. *Wetlands of Tanzania*, p.49.

- Čandek, K. and Kuntner, M., 2014. DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources*, 15(2), pp.268-277.
- Cawthorn, D., Steinman, H. and Witthuhn, R., 2012. Evaluation of the 16S and 12S rRNA genes as universal markers for the identification of commercial fish species in South Africa. *Gene*, 491(1), pp.40-48.
- Chenuil, A., Cahill, A., Délémontey, N., Du Salliant du Luc, E. and Fanton, H., 2019. Problems and questions posed by cryptic species. A framework to guide future studies. *History, Philosophy and Theory of the Life Sciences*, pp.77-106.
- Collares-Pereira, M. and Cowx, I., 2004. The role of catchment scale environmental management in freshwater fish conservation. *Fisheries Management and Ecology*, 11(3-4), pp.303-312.
- Collins, R., Bakker, J., Wangensteen, O., Soto, A., Corrigan, L., Sims, D., Genner, M.J. and Mariani, S., 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, 10(11), pp.1985-2001.
- Cooke, G.M., Chao, N.L. and Beheregaray, L.B., 2012. Five cryptic species in the Amazonian catfish *Centromochlus existimatus* identified based on biogeographic predictions and genetic data. *PLoS ONE*, 7(11), p.e48800.
- Dai, Q., Gao, Q., Wu, C., Chesters, D., Zhu, C. and Zhang, A., 2012. Phylogenetic Reconstruction and DNA Barcoding for Closely Related Pine Moth Species (*Dendrolimus*) in China with Multiple Gene Markers. *PLoS ONE*, 7(4), p.e32544.
- Dallas, H.F. and Rivers-Moore, N., 2014. Ecological consequences of global climate change for freshwater ecosystems in South Africa. *South African Journal of Science*, 110(5-6), pp.01-11.
- Darwall, W., Smith, K., Lowe, T. and Vié, J.C., 2005. The status and distribution of freshwater biodiversity in Eastern Africa. *Occasional Paper of the IUCN Species Survival Commission*, 31.
- Darwall, W.R.T., Smith, K.G., Allen, D.J., Holland, R.A., Harrison, I.J., and Brooks, E.G.E. (eds.), 2011. *The Diversity of Life in African Freshwaters: Under Water, Under Threat. An analysis of the status and distribution of freshwater species throughout mainland Africa*. Cambridge and Gland: IUCN, pp16-25.
- Day, J., Bills, R. and Friel, J., 2009. Lacustrine radiations in African *Synodontis* catfish. *Journal of Evolutionary Biology*, 22(4), pp.805-817.
- Day, J., Peart, C., Brown, K., Friel, J., Bills, R. and Moritz, T., 2013. Continental Diversification of an African Catfish Radiation (Mochokidae: *Synodontis*). *Systematic Biology*, 62(3), pp.351-365.
- De Graaf, G. and Garibaldi, L., 2014. *The Value of African Fisheries (FAO Fisheries and Aquaculture Circular No. 1093)*. Rome: FAO, p.76.

- Decru, E., Moelants, T., De Gelas, K., Vreven, E., Verheyen, E. and Snoeks, J., 2015. Taxonomic challenges in freshwater fishes: a mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. *Molecular Ecology Resources*, 16(1), pp.342-352.
- Doody, K. and Hamerlynck O., 2003. *Biodiversity of Rufiji District—a summary. (Technical report no. 44, Rufiji Environment Management Project)*. Cambridge and Gland: IUCN.
- Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L. and Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological reviews*, 81(2), pp.163-182.
- Dye, B.J. and Hartmann, J., 2017. *The True Cost of Power: The Facts and Risks of Building the Stiegler's Gorge Hydropower Dam in the Selous Game Reserve, Tanzania*. Gland: World Wildlife Fund International.
- Eccles, D., 1992. *Field Guide To The Freshwater Fishes Of Tanzania*. Rome: FAO.
- Ferraris Jr., C. and Reis, R., 2005. Neotropical catfish diversity: an historical perspective. *Neotropical Ichthyology*, 3(4), pp.453-454.
- Ficke, A., Myrick, C. and Hansen, L., 2007. Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries*, 17(4), pp.581-613.
- Florio, A., Ingram, C., Rakotondravony, H., Louis, E. and Raxworthy, C., 2012. Detecting cryptic speciation in the widespread and morphologically conservative carpet chameleon (*Furcifer lateralis*) of Madagascar. *Journal of Evolutionary Biology*, 25(7), pp.1399-1414.
- Fox, J., Weisberg, S., 2019. *An R Companion to Applied Regression*, Third edition. Thousand Oaks CA: Sage publications.
<https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- Friel, J.P. and Vigliotta, T.R., 2008. *Atopodontus adriaensi*, a new genus and species of African suckermouth catfish from the Ogooué and Nyanga River systems of Gabon (Siluriformes: Mochokidae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, 157(1), pp.13-23.
- Friel, J.P. and Vigliotta, T.R., 2011. Three new species of African suckermouth catfishes, genus *Chiloglanis* (Siluriformes: Mochokidae), from the lower Malagarasi and Luiche rivers of western Tanzania. *Zootaxa*, 3063(1), pp.1-21.
- Froese, R. and Pauly, D., 2019. *FishBase*. World Wide Web electronic publication.
www.fishbase.org.
- Fry, F.E.J., 1971. The effect of environmental factors on the physiology of fish. *Fish physiology*, 6(1), pp.1-98.
- Hamerlynck, O., Duvail, S., Vandepitte, L., Kindinda, K., Nyingi, D., Paul, J., Yanda, P., Mwakalinga, A., Mgaya, Y. and Snoeks, J., 2011. To connect or not to connect?

- Floods, fisheries and livelihoods in the Lower Rufiji floodplain lakes, Tanzania. *Hydrological Sciences Journal*, 56(8), pp.1436-1451.
- Hammer, Ø., Harper, D.A. and Ryan, P.D., 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia electronica*, 4(1), p.9.
- Hebert, P., Cywinska, A., Ball, S. and deWaard, J., 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), pp.313-321.
- Hubert, N., Hanner, R., Holm, E., Mandrak, N., Taylor, E., Burrige, M., Watkinson, D., Dumont, P., Curry, A., Bentzen, P., Zhang, J., April, J. and Bernatchez, L., 2008. Identifying Canadian Freshwater Fishes through DNA Barcodes. *PLoS ONE*, 3(6), p.e2490.
- IUCN, 2019. *Technical Review of The Environmental Impact Assessment for the Rufiji Hydropower Project in Selous Game Reserve, Tanzania*. [online] Gland: IUCN. <https://www.iucn.org/resources/publications>.
- Iyiola, O., Nneji, L., Mustapha, M., Nzeh, C., Oladipo, S., Nneji, I., Okeyoyin, A., Nwani, C., Ugwumba, O., Ugwumba, A., Faturoti, E., Wang, Y., Chen, J., Wang, W. and Adeola, A., 2018. DNA barcoding of economically important freshwater fish species from north-central Nigeria uncovers cryptic diversity. *Ecology and Evolution*, 8(14), pp.6932-6951.
- Janzen, D., Burns, J., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert, P. and Grishin, N., 2017. Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology. *Proceedings of the National Academy of Sciences*, 114(31), pp.8313-8318.
- Jay, S., Jones, C., Slinn, P. and Wood, C., 2007. Environmental impact assessment: Retrospect and prospect. *Environmental Impact Assessment Review*, 27(4), pp.287-300.
- Jirsová, D., Štefka, J., Blažek, R., Malala, J., Lotuliakou, D., Mahmoud, Z. and Jirků, M., 2019. From taxonomic deflation to newly detected cryptic species: Hidden diversity in a widespread African squeaker catfish. *Scientific Reports*, 9(1), pp.1-13.
- Kadye, W. and Moyo, N., 2008. Stream fish assemblage and habitat structure in a tropical African river basin (Nyagui River, Zimbabwe). *African Journal of Ecology*, 46(3), pp.333-340.
- Kashaigili, J., McCartney, M. and Mahoo, H., 2007. Estimation of environmental flows in the Great Ruaha River Catchment, Tanzania. *Physics and Chemistry of the Earth, Parts A/B/C*, 32(15-18), pp.1007-1014.
- Kautt, A., Machado-Schiaffino, G., Torres-Dowdall, J. and Meyer, A., 2016. Incipient sympatric speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua due to differential use of the benthic and limnetic habitats?. *Ecology and Evolution*, 6(15), pp.5342-5357.

- Kautt, A., Machado-Schiaffino, G. and Meyer, A., 2018. Lessons from a natural experiment: Allopatric morphological divergence and sympatric diversification in the Midas cichlid species complex are largely influenced by ecology in a deterministic way. *Evolution Letters*, 2(4), pp.323-340.
- Knight, M.E. and Turner, G.F., 2004. Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1540), pp.675-680.
- Ko, H.L., Wang, Y.T., Chiu, T.S., Lee, M.A., Leu, M.Y., Chang, K.Z., Chen, W.Y. and Shao, K.T., 2013. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS ONE*, 8(1), p.e53451.
- Kress, W. and Erickson, D., 2008. DNA barcodes: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences*, 105(8), pp.2761-2762.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, 35(6), pp.1547-1549.
- Kusche, H., Elmer, K.R. and Meyer, A., 2015. Sympatric ecological divergence associated with a color polymorphism. *BMC biology*, 13(1), pp.1-12.
- Lajus, D., Sukhikh, N. and Alekseev, V., 2015. Cryptic or pseudocryptic: can morphological methods inform copepod taxonomy? An analysis of publications and a case study of the *Eurytemora affinis* species complex. *Ecology and Evolution*, 5(12), pp.2374-2385.
- Lalronunga, S., Lalnuntluanga, and Lalramliana, 2011. Diversity of catfish (Teleostei: Siluriformes) in rivers of Barak drainage of Mizoram, Northeastern India. *Advances in Environmental Chemistry*, pp.297-300.
- Lam, V., Cheung, W., Reygondeau, G. and Sumaila, U., 2016. Projected change in global fisheries revenues under climate change. *Scientific Reports*, 6(1), p.32607.
- Langerhans, R., 2008. Predictability of phenotypic differentiation across flow regimes in fishes. *Integrative and Comparative Biology*, 48(6), pp.750-768.
- Leadley, P., 2010. *Biodiversity Scenarios: Projections of 21st Century Change in Biodiversity, and Associated Ecosystem Services: A Technical Report for the Global Biodiversity Outlook 3* (No. 50). Montreal: UNEP/Earthprint.
- Lee, M.S. and Palci, A., 2015. Morphological phylogenetics in the genomic age. *Current Biology*, 25(19), pp.R922-R929.
- Lévêque, C., 1997. *Biodiversity Dynamics and Conservation*. New York: Cambridge University Press, p.17.
- Lévêque, C., and Paugy, D., 2017. *Freshwater Fishes in Africa*. Paris: EOLSS.

- Lostrom, S., Evans, J., Grierson, P., Collin, S., Davies, P. and Kelley, J., 2015. Linking stream ecology with morphological variability in a native freshwater fish from semi-arid Australia. *Ecology and Evolution*, 5(16), pp.3272-3287.
- Lv, J., Wu, S., Zhang, Y., Chen, Y., Feng, C., Yuan, X., Jia, G., Deng, J., Wang, C., Wang, Q., Mei, L. and Lin, X. (2014). Assessment of four DNA fragments (COI, 16S rDNA, ITS2, 12S rDNA) for species identification of the Ixodida (Acari: Ixodida). *Parasites & Vectors*, 7(1), p.93.
- Marshall, B., 2012. Does climate change really explain changes in the fisheries productivity of Lake Kariba (Zambia-Zimbabwe)?. *Transactions of the Royal Society of South Africa*, 67(1), pp.45-51.
- Meier, R., Zhang, G. and Ali, F., 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Systematic Biology*, 57(5), pp.809-813.
- Mennesson, M.I. and KeIth, P., 2017. Evidence of two species currently under the name of *Eleotris fusca* (Gobioidei: Eleotridae) in the Indian Ocean. *Cybium*, 41, pp.213-220.
- Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M. and Kawabata, Z., 2011. Surveillance of fish species composition using environmental DNA. *Limnology*, 13(2), pp.193-197.
- Mombo, F. M., Speelman, S., Van Huylbroeck, G., Hella, J., Pantaleo, M., & Moe, S., 2011. Ratification of the Ramsar convention and sustainable wetlands management: Situation analysis of the Kilombero Valley wetlands in Tanzania. *Journal of Agricultural Extension and Rural Development*, 3(9), pp.153–164.
- Morgenstern, R., 2008. "*Pseudocrenilabrus sp. 'ruaha'*". *Cichlid Room Companion*. Retrieved on April 02, 2020, from: <https://cichlidae.com/species.php?id=2041>.
- Morris, J., Ford, A., Ali, J., Peart, C., Bills, R. and Day, J., 2015. High levels of genetic structure and striking phenotypic variability in a sexually dimorphic suckermouth catfish from the African Highveld. *Biological Journal of the Linnean Society*, 117(3), pp.528-546.
- Msangameno, D.J. and Mangora, M.M., 2016. Aspects of seasonal and long-term trends in fisheries and livelihoods in the Kilombero River basin, Tanzania. *African Journal of Tropical Hydrobiology and Fisheries*, 14(1), pp.1-11.
- Muñoz-Mas, R., Sánchez-Hernández, J., Martínez-Capel, F., Tamatamah, R., Mohamedi, S., Massinde, R. and McClain, M., 2019. Microhabitat preferences of fish assemblages in the Udzungwa Mountains (Eastern Africa). *Ecology of Freshwater Fish*, 28(3), pp.473-484.
- Mwalyosi, R.B., 1988. Environmental impacts of the proposed Stiegler's Gorge hydropower project, Tanzania. *Environmental Conservation*, 15(3), pp.250-254.
- Mwalyosi, R.B., 1990. Resource potentials of the Rufiji River basin, Tanzania. *Ambio*, 29(1), pp.16-20.

- Nelson, J.S., 2006. *Fishes of the World*. 4th Edition. Hoboken: John Wiley & Sons.
- Ng, H.H. and Bailey, R.M., 2006. *Chiloglanis productus*, a new species of suckermouth catfish (Siluriformes: Mochokidae) from Zambia. Occasional Papers of the University of Michigan Museum of Zoology, 738, pp.1-13.
- Nyboer, E., Liang, C. and Chapman, L., 2019. Assessing the vulnerability of Africa's freshwater fishes to climate change: A continent-wide trait-based analysis. *Biological Conservation*, 236, pp.505-520.
- Otu, M., Ramlal, P., Wilkinson, P., Hall, R. and Hecky, R., 2011. Paleolimnological evidence of the effects of recent cultural eutrophication during the last 200 years in Lake Malawi, East Africa. *Journal of Great Lakes Research*, 37, pp.61-74.
- Palandačić, A., Naseka, A., Ramler, D. and Ahnelt, H., 2017. Contrasting morphology with molecular data: an approach to revision of species complexes based on the example of European Phoxinus (Cyprinidae). *BMC Evolutionary Biology*, 17(1), p.184.
- Plenderleith, M., Oosterhout, C.V., Robinson, R.L. and Turner, G.F., 2005. Female preference for conspecific males based on olfactory cues in a Lake Malawi cichlid fish. *Biology Letters*, 1(4), pp.411-414.
- Puillandre, N., Modica, M.V., Zhang, Y., Sirovich, L., Boisselier, M.C., Cruaud, C., Holford, M. and Samadi, S., 2012. Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology*, 21(11), pp.2671-2691.
- R Core Team, 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org/>.
- Recknagel, H., Elmer, K. and Meyer, A., 2014. Crater lake habitat predicts morphological diversity in adaptive radiations of cichlid fishes. *Evolution*, 68(7), pp.2145-2155.
- Ribbink, A. J., 2001. *Lake Malawi/Niassa/Nyassa Ecoregion: Biophysical Reconnaissance*. Harare, Zimbabwe: WWF Southern African Regional Programme Office.
- Roca, A., 2001. Genetic evidence for two species of elephant in Africa. *Science*, 293(5534), pp.1473-1477.
- Rosser, N., Freitas, A., Huertas, B., Joron, M., Lamas, G., Mérot, C., Simpson, F., Willmott, K., Mallet, J. and Dasmahapatra, K., 2018. Cryptic speciation associated with geographic and ecological divergence in two Amazonian *Heliconius* butterflies. *Zoological Journal of the Linnean Society*, 186(1), pp.233-249.
- Rytwinski, T., Algera, D.A., Taylor, J.J., Smokorowski, K.E., Bennett, J.R., Harrison, P.M. and Cooke, S.J., 2017. What are the consequences of fish entrainment and impingement associated with hydroelectric dams on fish productivity? A systematic review protocol. *Environmental Evidence*, 6(1), pp.1-9.
- Saitou, N. and Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406-425.

- Salzburger, W., 2018. Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics*, 19(11), pp. 705-717.
- Santos, J., Araújo, F. and Silva, D., 2009. Length correction for early-juvenile Brazilian herring *Sardinella janeiro* (Eigenmann, 1894) after preservation in formalin, ethanol and freezing. *Neotropical Ichthyology*, 7(1), pp.87-92.
- Sarkar, U., Pathak, A., Sinha, R., Sivakumar, K., Pandian, A., Pandey, A., Dubey, V. and Lakra, W., 2011. Freshwater fish biodiversity in the River Ganga (India): changing pattern, threats and conservation perspectives. *Reviews in Fish Biology and Fisheries*, 22(1), pp.251-272.
- Sayer, C. A., Máiz-Tomé, L., & Darwall, W. R. T., 2018. *Freshwater Biodiversity in the Lake Victoria Basin: Guidance for Species Conservation, Site Protection, Climate Resilience and Sustainable Livelihoods*. Cambridge and Gland: IUCN, pp.41-61.
- Sayer, C., Palmer-Newton, A. and Darwall, W., 2019. *Conservation Priorities for Freshwater Biodiversity in the Lake Malawi/Nyasa/Niassa Catchment*. Cambridge and Gland: IUCN, pp.46-69.
- Schmidt, R.C. and Pezold, F., 2011. Morphometric and molecular variation in mountain catfishes (Amphiliidae: *Amphilius*) in Guinea, West Africa. *Journal of Natural History*, 45(9-10), pp.521-552.
- Schmidt, R.C., Bart Jr, H.L. and Nyingi, W.D., 2015. Two new species of African suckermouth catfishes, genus *Chiloglanis* (Siluriformes: Mochokidae), from Kenya with remarks on other taxa from the area. *Zootaxa*, 4044(1), pp.45-64.
- Schmidt, R.C., Bart Jr, H.L. and Pezold, F., 2016. High levels of endemism in suckermouth catfishes (Mochokidae: *Chiloglanis*) from the Upper Guinean forests of West Africa. *Molecular Phylogenetics and Evolution*, 100, pp.199-205.
- Schmidt, R.C., Bart Jr, H.L. and Nyingi, W.D., 2018. Integrative taxonomy of the red-finned barb, *Enteromius apleurogramma* (Cyprininae: Smiliogastrini) from Kenya, supports recognition of *E. amboseli* as a valid species. *Zootaxa*, 4482(3), pp.566-578.
- Schmidt, R.C. and Barrientos, C., 2019. A new species of suckermouth catfish (Mochokidae: *Chiloglanis*) from the Rio Mongo in Equatorial Guinea. *Zootaxa*, 4652(3), pp.507-519.
- Schmieder, D.A., Benítez, H.A., Borissov, I.M. and Fruciano, C., 2015. Bat species comparisons based on external morphology: a test of traditional versus geometric morphometric approaches. *PloS ONE*, 10(5), p.e0127043.
- Scotland, R., Olmstead, R. and Bennett, J., 2003. Phylogeny reconstruction: the role of morphology. *Systematic Biology*, 52(4), pp.539-548.
- Seegers, L., 2008. *The Catfishes of Africa*. Rodgau, Germany: Aqualog Verlag.
- Seehausen, O. and Wagner, C., 2014. Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics*, 45(1), pp.621-651.

- Seeteram, N., Hyera, P., Kaaya, L., Lalika, M. and Anderson, E., 2019. Conserving rivers and their biodiversity in Tanzania. *Water*, 11(12), p.2612.
- Shechonge, A., Ngatunga, B.P., Tamatamah, R., Bradbeer, S.J., Harrington, J., Ford, A.G., Turner, G.F. and Genner, M.J., 2018. Losing cichlid fish biodiversity: genetic and morphological homogenization of tilapia following colonization by introduced species. *Conservation Genetics*, 19(5), pp.1199-1209.
- Shechonge, A., Ngatunga, B.P., Bradbeer, S.J., Day, J.J., Freer, J.J., Ford, A.G., Kihedu, J., Richmond, T., Mzighani, S., Smith, A.M. and Sweke, E.A., 2019. Widespread colonisation of Tanzanian catchments by introduced *Oreochromis tilapia* fishes: the legacy from decades of deliberate introduction. *Hydrobiologia*, 832(1), pp.235-253.
- Shuai, F., Yu, S., Lek, S. and Li, X., 2018. Habitat effects on intra-species variation in functional morphology: Evidence from freshwater fish. *Ecology and Evolution*, 8(22), pp.10902-10913.
- Skelton, P. and Swartz, E., 2011. Walking the tightrope: trends in African freshwater systematic ichthyology. *Journal of Fish Biology*, 79(6), pp.1413-1435.
- Smith, C.D.M., 2016. Environmental flows in Rufiji River basin assessed from the perspective of planned development in the Kilombero and Lower Rufiji sub-basins. *Report to the United States Agency for International Development*. 146p. Available for download at: <https://dec.usaid.gov/dec/content/search.aspx>, p.4.
- Soesbergen, A., Sassen, M., Kimsey, S. and Hill, S., 2019. Potential impacts of agricultural development on freshwater biodiversity in the Lake Victoria basin. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(7), pp.1052-1062.
- Souza-Shibatta, L., Tonini, J.F., Abrahão, V.P., Jarduli, L.R., Oliveira, C., Malabarba, L.R., Sofia, S.H. and Shibatta, O.A., 2018. Reappraisal of the systematics of *Microglanis cottoides* (Siluriformes, Pseudopimelodidae), a catfish from southern Brazil. *PLoS ONE*, 13(7), p.e0199963.
- Stacklies, W., Redestig, H., Scholz, M., Walther, D., Selbig, J., 2007. pcaMethods – a Bioconductor package providing PCA methods for incomplete data. *Bioinformatics*, 23, 1164–1167.
- Staffen, C., Staffen, M., Becker, M., Löfgren, S., Muniz, Y., de Freitas, R. and Marrero, A., 2017. DNA barcoding reveals the mislabeling of fish in a popular tourist destination in Brazil. *PeerJ*, 5, p.e4006.
- Stebbins, G. L., 1950. *Variation and evolution in plants*. New York: Columbia University Press, pp.643.
- Stepanović, S., Kosovac, A., Krstić, O., Jović, J. and Toševski, I., 2015. Morphology versus DNA barcoding: two sides of the same coin. A case study of *Ceutorhynchus erysimi* and *C. contractus* identification. *Insect Science*, 23(4), pp.638-648.
- Surget-Groba, Y., Johansson, H. and Thorpe, R., 2012. Synergy between allopatry and ecology in population differentiation and speciation. *International Journal of Ecology*, 2012, pp.1-10.

- Swartz, E.R., Mwale, M. and Hanner, R., 2008. A role for barcoding in the study of African fish diversity and conservation. *South African Journal of Science*, 104(7-8), pp.293-298.
- Takahara, T., Minamoto, T., Yamanaka, H., Doi, H. and Kawabata, Z., 2012. Estimation of Fish Biomass Using Environmental DNA. *PLoS ONE*, 7(4), p.e35868.
- Tanzanian Fisheries Sector, 2016. *The Tanzanian Fisheries Sector - Challenges and Opportunities*. United Republic of Tanzania Ministry of Agriculture, Livestock and Fisheries. <https://www.wiomsa.org/download/the-tanzanian-fisheries-sectors-challenges-and-opportunities/>
- Thompson, J., Higgins, D. and Gibson, T., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), pp.4673-4680.
- Thomsen, P. and Willerslev, E., 2015. Environmental DNA – an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, pp.4-18.
- Thorpe, A and Castillo, C.Z., 2018. The economic value of inland fisheries. Review of the state of the world fishery resources: inland fisheries. *FAO Fisheries and Aquaculture Circular*, 942(3), pp. 214-253.
- Tickner, D., Opperman, J., Abell, R., Acreman, M., Arthington, A., Bunn, S., Cooke, S., Dalton, J., Darwall, W., Edwards, G., Harrison, I., Hughes, K., Jones, T., Leclère, D., Lynch, A., Leonard, P., McClain, M., Muruven, D., Olden, J., Ormerod, S., Robinson, J., Tharme, R., Thieme, M., Tockner, K., Wright, M. and Young, L., 2020. Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. *BioScience*, 70(4), pp.330-342.
- Tweddle, D., Cambray, J., Kazembe, J., Marshall, B. & Moelants, T. 2019. *Brycinus lateralis*. *The IUCN Red List of Threatened Species* 2019: e.T60749A47213447. <https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T60749A47213447.en>.
- Utzing, J. and Charlwood, J.D., 1996. Fishery methods and fish diversity in the Kilombero River in south-eastern Tanzania. *African Journal of Tropical Hydrobiology and Fisheries*, 7(1), pp.55-64.
- Van der Knaap, M., 1994. *Status of Fish Stocks and Fisheries of Thirteen Medium-Sized African Reservoirs*. Rome: FAO.
- Van Ginkel, C.E., 2011. Eutrophication: present reality and future challenges for South Africa. *Water SA*, 37(5), pp.693-702.
- Vrtílek, M. and Reichard, M., 2016. Patterns of morphological variation among populations of the widespread annual killifish *Nothobranchius orthonotus* are independent of genetic divergence and biogeography. *Journal of Zoological Systematics and Evolutionary Research*, 54(4), pp.289-298.

- Vu, H. and Le, L., 2019. Bioinformatics analysis on DNA barcode sequences for species identification: A Review. *Annual Research & Review in Biology*, pp.1-12.
- Ward, R., Zemlak, T., Innes, B., Last, P. and Hebert, P., 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462), pp.1847-1857.
- Watson, H., 2020. 'Unrecognised endemic biodiversity within East African Chiloglanis catfish populations', Masters by Research Thesis, University of Bristol.
- Weyl, O., Ribbink, A. and Tweddle, D., 2010. Lake Malawi: fishes, fisheries, biodiversity, health and habitat. *Aquatic Ecosystem Health & Management*, 13(3), pp.241-254.
- Williams GC., 1966. *Adaptation and Natural Selection*. Princeton (NJ): Princeton University Press.
- Zheng, L., He, J., Lin, Y., Cao, W. and Zhang, W., 2014. 16S rRNA is a better choice than COI for DNA barcoding hydrozoans in the coastal waters of China. *Acta Oceanologica Sinica*, 33(4), pp.55-76.

Appendix

Appendix 2.1. Full DNA barcode database search results for the COI and 12S barcodes of specimens.

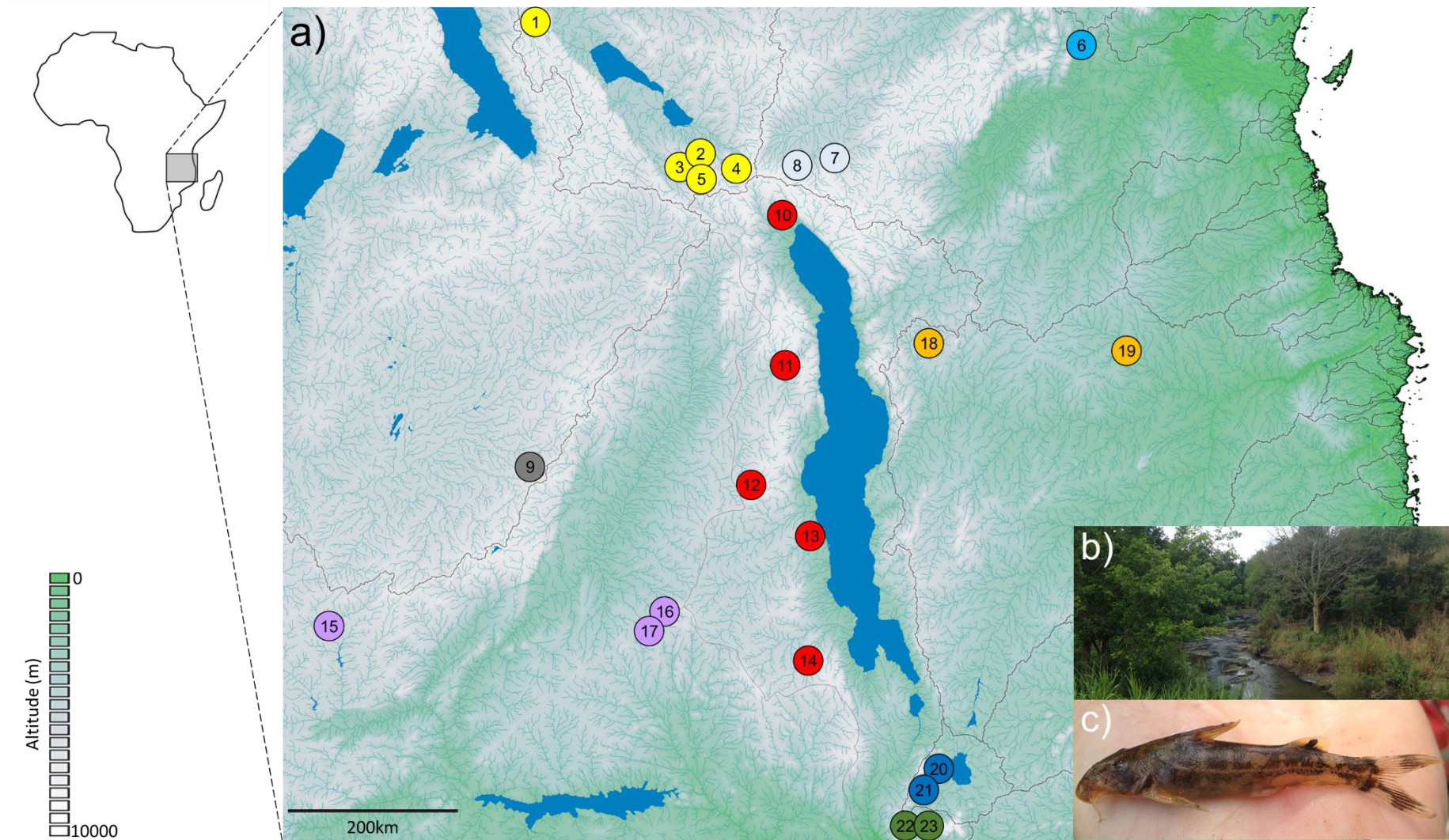
Sample	Phenotype ID	COI NCBI			COI BOLD		12S NCBI		
		Max Score	ID (%)	Species Name	Similarity (%)	Species Name	Max Score	ID (%)	Species Name
031-S8-09-12	<i>Enteromius cf. lineomaculatus "B"</i>						651	90.3	<i>Barbus sp. CToL 033</i>
034-S8-09-12	<i>Enteromius cf. lineomaculatus "A"</i>						706	92.5	<i>Barbus sp. CToL 033</i>
035-S8-09-12	<i>Enteromius cf. lineomaculatus "B"</i>	830	95.2	<i>Enteromius pallidus</i>	97.2	<i>Barbus sp. (early release)</i>			
036-S8-09-12	<i>Enteromius cf. lineomaculatus "B"</i>						675	89.8	<i>Barbus sp. CToL 033</i>
037-S8-09-12	<i>Enteromius cf. lineomaculatus "B"</i>	868	91.0	<i>Enteromius atkinsoni</i>	97.2	<i>Barbus sp. (early release)</i>	673	89.9	<i>Barbus sp. CToL 033</i>
038-S8-09-12	<i>Enteromius kerstenii</i>	1009	95.8	<i>Barbus sp. CToL 051</i>	96.1	<i>Enteromius kerstenii</i>	928	98.1	<i>Barbus sp. CToL 051</i>
039-S8-09-12	<i>Enteromius kerstenii</i>	1002	95.8	<i>Barbus sp. CToL 051</i>	96.1	<i>Enteromius kerstenii</i>	898	98.1	<i>Barbus sp. CToL 051</i>
040-S8-09-12	<i>Enteromius kerstenii</i>	1003	96.1	<i>Barbus sp. CToL 051</i>	96.2	<i>Barbus sp. CToL 051</i>			
041-S8-09-12	<i>Kneria ruaha</i>	807	89.5	<i>Parakneria cameronensis</i>	89.4	<i>Parakneria cameronensis</i>	944	98.9	<i>Parakneria cameronensis</i>
042-S8-09-12	<i>Kneria ruaha</i>	809	89.5	<i>Parakneria cameronensis</i>	89.4	<i>Parakneria cameronensis</i>	944	98.9	<i>Parakneria cameronensis</i>
043-S8-09-12	<i>Amphilius sp. "undetermined"</i>	723	87.3	<i>Amphilius zairensis</i>	100.0	<i>Amphilius pedunculus</i>	737	91.7	<i>Pangasianodon hypophthalmus</i>
044-S8-09-12	<i>Amphilius sp. "undetermined"</i>	731	87.3	<i>Amphilius zairensis</i>	100.0	<i>Amphilius pedunculus</i>	734	91.5	<i>Pangasianodon hypophthalmus</i>
045-S9-09-12	<i>Zaireichthys sp. "undetermined"</i>	707	92.0	<i>Zaireichthys monomotapa</i>	93.3	<i>Zaireichthys sp. (early release)</i>	708	92.0	<i>Amphilius sp. CD-2019</i>
046-S9-09-12	<i>Zaireichthys sp. "undetermined"</i>	707	92.0	<i>Zaireichthys monomotapa</i>	93.3	<i>Zaireichthys sp. (early release)</i>			
047-S9-09-12	<i>Zaireichthys sp. "undetermined"</i>	671	91.4	<i>Zaireichthys sp. 'slender'</i>	93.3	<i>Zaireichthys sp. (early release)</i>	678	89.7	<i>Bagarius yarrelli</i>
048-S9-09-12	<i>Amphilius sp. "undetermined"</i>	723	87.3	<i>Amphilius zairensis</i>	100.0	<i>Amphilius pedunculus</i>			
049-S9-09-12	<i>Amphilius sp. "undetermined"</i>	723	87.3	<i>Amphilius zairensis</i>	100.0	<i>Amphilius pedunculus</i>			
050-S9-09-12	<i>Kneria ruaha</i>	799	89.8	<i>Parakneria cameronensis</i>	89.7	<i>Parakneria cameronensis</i>			
051-S9-09-12	<i>Kneria ruaha</i>	811	89.4	<i>Parakneria cameronensis</i>	89.5	<i>Parakneria cameronensis</i>	928	98.9	<i>Parakneria cameronensis</i>
063-KIL01-05-19	<i>Labeo congoro</i>	1048	97.3	<i>Labeo lineatus</i>	97.4	<i>Labeo lineatus</i>	937	98.7	<i>Labeo altivelis</i>
064-KIL01-05-19	<i>Labeo congoro</i>	1043	97.1	<i>Labeo lineatus</i>	97.2	<i>Labeo lineatus</i>	883	99.0	<i>Labeo altivelis</i>
065-KIL01-05-19	<i>Labeobarbus macrolepis</i>	962	95.4	<i>Labeobarbus macrocephalus x Labeobarbus mawambiensis</i>	95.5	<i>Arabibarbus grypus</i>			
067-KIL01-05-19	<i>Distichodus petersii</i>						813	95.5	<i>Distichodus sexfasciatus</i>
068-KIL01-05-19	<i>Distichodus petersii</i>	1100	99.8	<i>Distichodus petersii</i>	99.8	<i>Distichodus petersii</i>	830	95.7	<i>Distichodus sexfasciatus</i>
069-KIL01-05-19	<i>Clarias gariepinus</i>	1079	99.0	<i>Clarias gariepinus</i>	100.0	<i>Clarias gariepinus</i>			
070-KIL01-05-19	<i>Clarias gariepinus</i>	1077	99.0	<i>Clarias gariepinus</i>	100.0	<i>Clarias gariepinus</i>			
071-KIL01-05-19	<i>Oreochromis urolepis</i>						976	100.0	<i>Oreochromis sp. 'red tilapia'</i>
072-KIL01-05-19	<i>Oreochromis urolepis</i>						983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
073-KIL01-05-19	<i>Hydrocynus vittatus</i>	973	95.1	<i>Hydrocynus vittatus</i>	95.4	<i>Hydrocynus vittatus</i>	656	90.6	<i>Hydrocynus goliath</i>
075-KIL01-05-19	<i>Synodontis rufigiensis</i>	950	95.0	<i>Synodontis sp. voucher BNF 110</i>	95.0	<i>Synodontis sp. BOLD:AAL5722</i>	845	98.5	<i>Synodontis nigromaculata</i>
076-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	955	95.0	<i>Synodontis sp. voucher BNF 110</i>	95.0	<i>Synodontis sp. BOLD:AAL5722</i>	830	98.5	<i>Synodontis nigromaculata</i>
077-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	991	94.6	<i>Synodontis thysi</i>	95.0	<i>Synodontis sp. BOLD:AAL5722</i>	889	98.4	<i>Synodontis nigromaculata</i>

Appendix 2.1. continued...

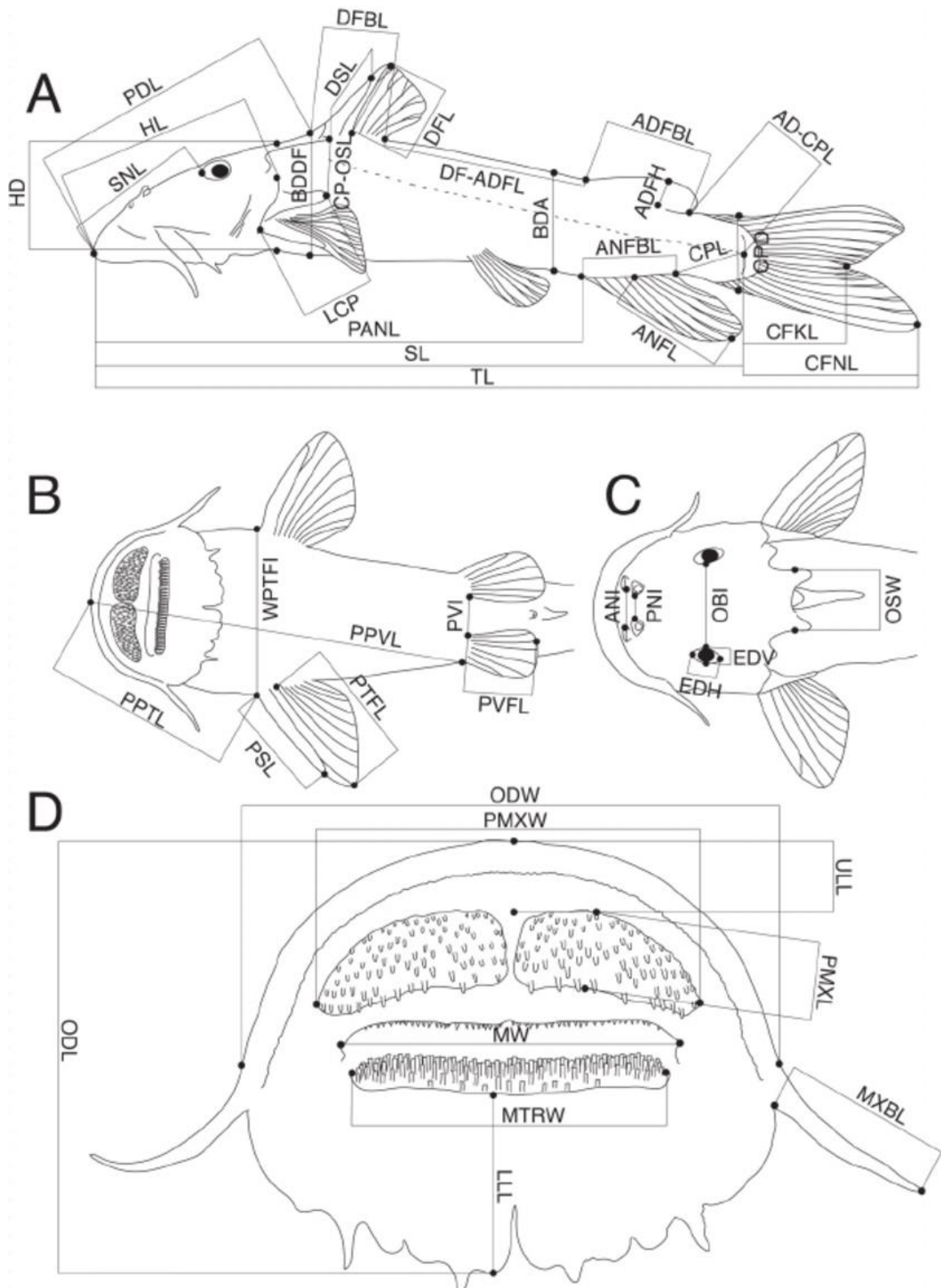
Sample	Phenotype ID	COI NCBI			COI BOLD		12S NCBI		
		Max Score	ID (%)	Species Name	Similarity (%)	Species Name	Max Score	ID (%)	Species Name
078-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	984	95.2	<i>Synodontis sp. voucher BNF 110</i>	95.0	<i>Synodontis sp. BOLD:AAL5722</i>			
079-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	1140	99.2	<i>Synodontis rukwaensis</i>	99.7	<i>Synodontis rukwaensis</i>	878	99.0	<i>Synodontis irsacae</i>
082-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	1101	98.7	<i>Synodontis rukwaensis</i>	99.1	<i>Synodontis rukwaensis</i>			
083-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	1098	99.2	<i>Synodontis rukwaensis</i>	99.7	<i>Synodontis rukwaensis</i>	911	99.0	<i>Synodontis irsacae</i>
084-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	1094	99.0	<i>Synodontis rukwaensis</i>	99.3	<i>Synodontis rukwaensis</i>	909	99.0	<i>Synodontis irsacae</i>
085-KIL01-05-19	<i>Synodontis sp. "undetermined"</i>	1110	99.1	<i>Synodontis rukwaensis</i>	99.5	<i>Synodontis rukwaensis</i>			
086-KIL01-05-19	<i>Synodontis rufigiensis</i>	973	95.1	<i>Synodontis sp. voucher BNF 110</i>	95.1	<i>Synodontis sp. BOLD:AAL5722</i>			
087-KIL01-05-19	<i>Synodontis rufigiensis</i>	976	95.1	<i>Synodontis sp. voucher BNF 110</i>	95.1	<i>Synodontis sp. BOLD:AAL5722</i>			
088-KIL01-05-19	<i>Synodontis rufigiensis</i>	970	95.0	<i>Synodontis thysi</i>	94.8	<i>Synodontis thysi</i>			
089-KIL02-05-19	<i>Citharinus congicus</i>	870	95.8	<i>Citharinus congicus</i>	94.8	<i>Citharinus congicus</i>	870	95.8	<i>Citharinus congicus</i>
090-KIL03-05-19	<i>Hydrocynus vittatus</i>	966	95.1	<i>Hydrocynus vittatus</i>	95.3	<i>Hydrocynus vittatus</i>	623	90.4	<i>Hydrocynus goliath</i>
091-KIL03-05-19	<i>Hydrocynus vittatus</i>	966	95.1	<i>Hydrocynus vittatus</i>	95.3	<i>Hydrocynus vittatus</i>	658	90.6	<i>Hydrocynus goliath</i>
092-KIL03-05-19	<i>Hydrocynus vittatus</i>	968	95.0	<i>Hydrocynus vittatus</i>	95.2	<i>Hydrocynus vittatus</i>			
094-KIL03-05-19	<i>Brycinus affinis</i>						739	94.2	<i>Brycinus imberi</i>
096-KIL02-05-19	<i>Hydrocynus vittatus</i>	978	95.1	<i>Hydrocynus vittatus</i>	95.4	<i>Hydrocynus vittatus</i>			
097-KIL02-05-19	<i>Brycinus imberi</i>	1009	96.4	<i>Brycinus imberi</i>	98.7	<i>Brycinus lateralis</i>	819	97.9	<i>Brycinus imberi</i>
098-KIL02-05-19	<i>Distichodus petersii</i>	1132	99.8	<i>Distichodus petersii</i>	99.8	<i>Distichodus petersii</i>	826	95.6	<i>Distichodus sexfasciatus</i>
099-KIL02-05-19	<i>Enteromius sp. "A"</i>	1131	100.0	<i>Enteromius apleurogramma</i>	98.9	<i>Barbus sp. (early release)</i>			
100-KIL02-05-19	<i>Astatotilapia sp. "undetermined"</i>	976	95.7	<i>Haplochromis burtoni</i>	95.7	<i>Julidochromis regani / Haplochromis burtoni</i>	953	99.1	<i>Tropheus duboisi</i>
101-KIL02-05-19	<i>Brycinus affinis</i>						734	94.2	<i>Brycinus imberi</i>
102-KIL02-05-19	<i>Oreochromis urolepis</i>						976	100.0	<i>Oreochromis sp. 'red tilapia'</i>
103-KIL02-05-19	<i>Oreochromis urolepis</i>						979	100.0	<i>Oreochromis sp. 'red tilapia'</i>
104-KIL02-05-19	<i>Astatotilapia sp. "undetermined"</i>	997	95.7	<i>Haplochromis burtoni</i>	95.6	<i>Julidochromis regani / Haplochromis burtoni</i>	911	99.0	<i>Tropheus duboisi</i>
105-KIL02-05-19	<i>Ctenopoma muriei</i>						669	89.7	<i>Stichaeus grigorjewi</i>
106-KIL07-05-19	<i>Oreochromis urolepis</i>						979	100.0	<i>Oreochromis sp. 'red tilapia'</i>
107-KIL07-05-19	<i>Oreochromis urolepis</i>						983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
108-KIL07-05-19	<i>Oreochromis urolepis</i>						970	99.8	<i>Oreochromis sp. 'red tilapia'</i>
109-KIL07-05-19	<i>Oreochromis urolepis</i>						948	100.0	<i>Oreochromis sp. 'red tilapia'</i>
110-KIL07-05-19	<i>Oreochromis urolepis</i>						976	100.0	<i>Oreochromis sp. 'red tilapia'</i>
111-KIL07-05-19	<i>Oreochromis urolepis</i>						983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
112-KIL07-05-19	<i>Oreochromis urolepis</i>						983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
113-KIL07-05-19	<i>Oreochromis urolepis</i>						948	100.0	<i>Oreochromis sp. 'red tilapia'</i>
114-KIL07-05-19	<i>Oreochromis urolepis</i>	1134	99.4	<i>Oreochromis sp. 'red tilapia'</i>	99.5	<i>Oreochromis urolepis</i>	983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
115-KIL07-05-19	<i>Oreochromis urolepis</i>						983	100.0	<i>Oreochromis sp. 'red tilapia'</i>
116-KIL07-05-19	<i>Astatotilapia sp. "undetermined"</i>	984	95.7	<i>Julidochromis regani</i>	95.7	<i>Julidochromis regani / Haplochromis burtoni</i>	946	99.2	<i>Tropheus duboisi</i>
117-KIL07-05-19	<i>Brycinus imberi</i>	1019	96.3	<i>Brycinus imberi</i>	98.5	<i>Brycinus lateralis</i>	819	97.9	<i>Brycinus imberi</i>
118-KIL07-05-19	<i>Brycinus imberi</i>	1020	96.5	<i>Brycinus imberi</i>	98.7	<i>Brycinus lateralis</i>			

Appendix 2.1. continued...

Sample	Phenotype ID	COI NCBI			COI BOLD		12S NCBI		
		Max Score	ID (%)	Species Name	Similarity (%)	Species Name	Max Score	ID (%)	Species Name
119-KIL07-05-19	<i>Clarias gariepinus</i>	1110	99.1	<i>Clarias gariepinus</i>	100.0	<i>Clarias gariepinus</i>			
120-KIL11-05-19	<i>Alestes stuhlmannii</i>	872	91.5	<i>Alestes baremoze</i>	91.5	<i>Alestes baremoze</i>			
121-MG10-05-19	<i>Clarias theodora</i>	965	94.1	<i>Clarias angolensis</i>	95.8	<i>Clarias sp. 2DO</i>	821	95.9	<i>Clarias gariepinus</i>
122-KIL08-05-19	<i>Nannaethiops sp. "upper Rufiji"</i>	805	88.5	<i>Distichodus hypostomatus</i>	88.7	<i>Distichodus hypostomatus</i>			
123-KIL08-05-19	<i>Nannaethiops sp. "upper Rufiji"</i>	728	87.6	<i>Distichodus mossambicus</i>	88.5	<i>Distichodus hypostomatus</i>	806	95.6	<i>Distichodus fasciolatus</i>
124-KIL08-05-19	<i>Nannaethiops sp. "upper Rufiji"</i>	721	87.6	<i>Distichodus mossambicus</i>	88.4	<i>Distichodus hypostomatus</i>	758	95.0	<i>Distichodus fasciolatus</i>
125-KIL09-05-19	<i>Nannaethiops sp. "upper Rufiji"</i>						780	95.5	<i>Distichodus fasciolatus</i>
126-KIL09-05-19	<i>Nannaethiops sp. "upper Rufiji"</i>	699	87.4	<i>Distichodus mossambicus</i>	88.6	<i>Distichodus hypostomatus</i>	763	95.2	<i>Distichodus fasciolatus</i>
127-S22-08-13	<i>Synodontis matthesi</i>	1127	99.4	<i>Synodontis rukwaensis</i>	99.6	<i>Synodontis rukwaensis</i>			
128-S22-08-13	<i>Eleotris klunzingerii</i>	1146	99.8	<i>Eleotris klunzingerii</i>	99.8	<i>Eleotris klunzingerii</i>	955	99.1	<i>Eleotris klunzingerii</i>
131-S22-08-13	<i>Citharinus congicus</i>	983	95.3	<i>Citharinus congicus</i>	95.2	<i>Citharinus congicus</i>	850	96.0	<i>Citharinus congicus</i>
132-S22-08-13	<i>Citharinus congicus</i>	989	95.3	<i>Citharinus congicus</i>	95.1	<i>Citharinus congicus</i>	845	95.8	<i>Citharinus congicus</i>
133-S22-08-13	<i>Brycinus affinis</i>						741	94.2	<i>Brycinus imberi</i>
134-S22-08-13	<i>Brycinus affinis</i>						739	94.2	<i>Brycinus imberi</i>
135-S22-08-13	<i>Pareutropius longifilis</i>	870	91.7	<i>Schilbe intermedius</i>	91.7	<i>Schilbe intermedius</i>	806	96.3	<i>Schilbe intermedius</i>
136-S22-08-13	<i>Pareutropius longifilis</i>	865	91.6	<i>Schilbe intermedius</i>	91.5	<i>Schilbe intermedius</i>	806	96.3	<i>Schilbe intermedius</i>
137-S22-08-13	<i>Labeo cylindricus</i>	959	94.0	<i>Labeo parvus</i>	94.6	<i>Labeo lunatus</i>	878	98.0	<i>Labeo parvus</i>
138-S22-08-13	<i>Labeo cylindricus</i>	959	94.0	<i>Labeo parvus</i>	94.6	<i>Labeo lunatus</i>	878	98.0	<i>Labeo parvus</i>
139-S22-08-13	<i>Alestes stuhlmannii</i>	878	91.8	<i>Alestes baremoze</i>	91.8	<i>Alestes baremoze</i>	634	89.4	<i>Micralestes sp. NM-2010</i>
140-S22-08-13	<i>Alestes stuhlmannii</i>	880	91.7	<i>Alestes baremoze</i>	91.8	<i>Alestes baremoze</i>	634	89.4	<i>Micralestes sp. NM-2010</i>
141-S22-08-13	<i>Synodontis rukwaensis</i>	1120	98.7	<i>Synodontis rukwaensis</i>	99.1	<i>Synodontis rukwaensis</i>	863	99.0	<i>Synodontis irsacae</i>
142-S22-08-13	<i>Glossogobius giuris</i>	1123	100.0	<i>Glossogobius giuris</i>	100.0	<i>Glossogobius giuris</i>			
143-S22-08-13	<i>Glossogobius giuris</i>	1134	100.0	<i>Glossogobius giuris</i>	100.0	<i>Glossogobius giuris</i>	793	94.1	<i>Glossogobius giuris</i>
144-S22-08-13	<i>Synodontis sp. "Utete"</i>	1018	96.4	<i>Synodontis aff. Ilebrevis</i>	96.7	<i>Synodontis granulosus</i>	911	99.2	<i>Synodontis irsacae</i>
145-S22-08-13	<i>Hydrocynus vittatus</i>	961	95.2	<i>Hydrocynus forskahlii</i>	95.3	<i>Hydrocynus vittatus</i>	667	90.7	<i>Hydrocynus goliath</i>
146-KIL08-19	<i>Pseudocrenilabrus sp. "Ruaha"</i>	1050	97.3	<i>Pseudocrenilabrus multicolor</i>	97.3	<i>Astatotilapia sp. BOLD:AAF6307</i>	898	97.5	<i>Tropheus duboisi</i>
147-KIL08-19	<i>Pseudocrenilabrus sp. "Ruaha"</i>	1050	97.3	<i>Pseudocrenilabrus multicolor</i>	97.3	<i>Astatotilapia sp. BOLD:AAF6307</i>	904	97.7	<i>Tropheus duboisi</i>
148-KIL08-19	<i>Oreochromis urolepis</i>						928	100.0	<i>Oreochromis sp. 'red tilapia'</i>
149-KIL08-19	<i>Enteromius sp. "A"</i>	1171	100.0	<i>Enteromius apleurogramma</i>	98.9	<i>Barbus sp. (early release)</i>			
150-KIL08-19	<i>Micropanchax sp. "undetermined"</i>						641	89.4	<i>Lacustricola pumilus</i>
151-KIL08-19	<i>Enteromius macrotaenia</i>	856	91.5	<i>Barbus innocens</i>	94.7	<i>Enteromius cf. macrotaenia</i>	704	91.9	<i>Systemus sarana sarana</i>
152-KIL08-19	<i>Enteromius sp. "A"</i>	1151	99.8	<i>Enteromius apleurogramma</i>	98.7	<i>Barbus sp. (early release)</i>			
153-KIL08-19	<i>Hemigrammopetersius barnadi</i>						785	93.4	<i>Micralestes sp. NM-2010</i>
154-KIL08-19	<i>Hemigrammopetersius barnadi</i>						732	93.0	<i>Micralestes sp. NM-2010</i>
155-KIL08-19	<i>Enteromius apleurogramma</i>	750	88.4	<i>Enteromius lineomaculatus</i>	99.0	<i>Barbus sp. (early release)</i>			
156-KIL08-19	<i>Enteromius apleurogramma</i>	723	88.4	<i>Barbus sp. 'Congo'</i>	99.0	<i>Barbus sp. (early release)</i>	627	89.5	<i>Barbus sp. CToL 033</i>
157-KIL08-19	<i>Brycinus lateralis</i>	990	95.8	<i>Brycinus lateralis</i>	96.4	<i>Brycinus sp. C03</i>	758	94.0	<i>Brycinus imberi</i>
158-KIL10-19	<i>Mastacembelus frenatus</i>						774	94.5	<i>Mastacembelus cunningtoni</i>
159-KIL10-19	<i>Labeo cylindricus</i>	1118	99.7	<i>Labeo fuelleborni</i>	99.8	<i>Labeo cylindricus</i>	928	99.2	<i>Labeo parvus</i>
160-KIL10-19	<i>Enteromius kerstenii</i>	974	94.9	<i>Enteromius sp. A7-09-847</i>	96.9	<i>Enteromius kerstenii</i>	869	97.5	<i>Barbus sp. CToL 051</i>



Appendix 3.1. a) Collection sites for *Chiloglanis* specimens considered in this study. Colours indicate the sampling catchment. b) An example of typical *Chiloglanis* habitat (at Site 6 in the Rufiji system). c) *Chiloglanis* sp. (from Site 6 in the Rufiji system). From Watson (2020).



Appendix 3.2. Illustration depicting the linear morphometric measurements, Lateral (A), ventral (B), dorsal (C) and oral disc (D) views of a representative chiloglanidin. Illustration from Friel and Vigliotta (2008).