The taxonomy of the topminnows (Teleostei: Cyprinodontiformes) (Procatopodidae: Procatopodinae) of southern Africa.

A thesis submitted in fulfilment of
the requirements of the degree of

Master of Science in Zoology

at

Rhodes University

by

Ryan Matthew van Zeeventer

BSc (University of Johannesburg 2012); BSc Hons (Zoology) (Rhodes 2013)

November 2019

Abstract

Taxonomy is a tool that plays a crucial role in the implementation of effective management assessments, the ability to identify different fish species and their life stages, and create effective conservation strategies. Recent years have seen an increase in the discovery of cryptic species which could have an effect on how we implement conservation strategies, increasing the importance of accurate taxonomic assessments.

The southern African topminnows of the genera *Micropanchax* and *Lacustricola* (Family Procatopodidae) include several species of egg-laying fishes that live in the marginal areas of rivers and lakes throughout the river systems of southern Africa in separated and connected populations. It was recently discovered that these populations showed subtle differences in their morphology and colouration, which led to the belief that new and cryptic species may be present.

Previous studies of the Procatopodidae relied heavily on phylogenetic analyses that drew on morphological traits without reliance on molecular methods of analysis. A few of the taxonomic issues were resolved through these limited studies but ultimately the taxonomic status of the Procatopodidae remained poorly understood. Furthermore, these studies had not been updated in recent years and hence the benefits of new technology had not been brought to bear on the issues.

A multi-gene analysis using standard phylogenetic methods and five molecular markers (mitochondrial cytochrome C oxidase (COI), mitochondrial cytochrome b (cyt*b*), mitochondrial 16S ribosomal RNA (16S rRNA), nuclear 28S ribosomal RNA (28S rRNA) and mitochondrial Tyrosine Kinase (X-src) (TyrK)) allowed phylogenetic trees to be made which demonstrated clear relationships within *Micropanchax* and *Lacustricola* of southern Africa and the out-groups used in the analysis. The inference trees showed that *Micropanchax johnstoni* was represented by three clades and was shown to be polyphyletic while *Micropanchax hutereaui*, which was represented by two clades, and *Micropanchax katangae,* which was represented by one clade, were monophyletic. The two species of *Lacustricola*, *L. myaposae* and *L. macrurus* showed very close relationships to *Micropanchax* and it is suggested that these two species be transferred to *Micropanchax*, leaving *Lacustricola* to only be found in central Africa.

The morphology of the species was looked at in the form of a multivariate truss-network analysis to determine distinct morphological identifiers for the different populations and species. The results of the analysis showed that there was very little variation between *Micropanchax* and *Lacustricola* species of southern Africa. The analysis of the morphology between populations in a species showed that some variation, although small occurred between populations in rivers and lakes. Live colouration was shown to play a role in identifications of species but could not be analysed as specimens stored for a longer period of time had lost their colouration.

The phylogenetic and morphometric analyses showed the presence of new and cryptic species throughout the different populations and the evidence gathered suggests that *L. myaposae* and *L. macrurus* should be transferred to *Micropanchax*.

Table of Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Figures

Chapter 2

Chapter 3

Chapter 4

Acknowledgements

I am extremely grateful to everyone for the help and support throughout the entire process of completing my thesis. I have been privileged to have the support and supervision of Prof. Martin Villet and Roger Bills throughout the process of completing this thesis. I will be eternally grateful to Prof. Villet for the many hours he invested in the success of this study, answering my copious questions, helping me gain greater insight for my research, and the time he spent reading through my work.

My sincere thanks go to Roger Bills for the assistance and support he provided throughout the study. His help in the laboratory and guidance in the processes for collecting and storing samples were extremely valuable, and I very much appreciate the time he took to travel to northern Namibia with me to collect samples in different river systems.

This study would not have been possible without the research funding provided by Prof. Martin Villet at Rhodes University and Roger Bills at the South African Institute for Aquatic Biodiversity (SAIAB).

I am extremely grateful to the staff at SAIAB Collection Laboratory and Molecular Laboratory, particularly Taryn Bodill from the Molecular Lab for helping me with my DNA extractions and amplifications and for providing insight when problems occurred, Nkosinathi Mazungula, Mzwandile Dwani, Bafo Konqobe and Zandi Adams for helping with cataloguing specimens and retrieving specimens from the collection.

I would like to thank James Maclaine and Oliver Crimmen from the British Museum of Natural History, and John Sparks, Barbara Brown and Radford Arrindell from the American Museum of Natural History for providing photographs of type specimens for the focus species of this study.

A huge thank you to Dr Shelley Edwards for all the advice given on the analysis of my phylogenetic work and for taking the time to teach me how to set up my sequences and complete the Bayesian analyses.

Last but by no means least, thank you to all my family and friends for their continued support throughout the study. Special thank you to Colleen Adler for proof reading my thesis. Without the ongoing support, love and confidence of my parents it would not have been possible for me to reach this point, thank you for all that you have done for me.

I truly appreciate you all.

1

An introduction to the topminnows (Teleostei: Cyprinodontiformes) of southern Africa.

1.1. Introduction

Topminnows, also known as lampeyes in English, are currently placed in the family Procatopodidae Fowler 1916 on accumulated molecular evidence (Braganca *et al.* 2018). They include species that occur in several river systems in southern Africa (Skelton 2001; Marshall 2011). Recently it has been found that populations of certain southern African species, specifically *Micropanchax johnstoni* (Günther 1894), *Micropanchax katangae* (Boulenger 1912) and *Micropanchax hutereaui* (Boulenger 1913), show subtle geographical differences in colouration and morphology that were not reported in previous studies. These subtle differences between populations may indicate the existence of cryptic species. Cryptic species may arise when conspecific populations have been geographically isolated for long periods of time, ultimately becoming independent lineages (Bickford *et al.* 2006). Currently, little research has been done using newer techniques such as multi-gene analysis and Box-truss morphometric analyses on these topminnows. This may be because of the lack of interest in the taxonomic field or the underestimated commercial and ecological value they could and do provide (Steenkamp *et al.* 2001).

The family Procatopodidae includes several genera of egg-bearing African topminnows of the subfamilies Aplocheilichthyinae and Procatopodinae (Skelton 2001, 2002; Lucinda 2003; Marshall 2011; Braganca *et al.* 2018). They are generally less than 50 mm long (Skelton 2001). The Aplocheilichthyinae includes one found in southern African genus, *Aplocheilichthys* Bleeker 1863 (Table 1.1). The Procatopodinae consist of substantially more genera, with the important genera in southern Africa being *Lacustricola* Myers 1924 and *Micropanchax* Myers 1924 (Table 1.1). The genus *Micropanchax* has species that show slender and deep body shapes, for example *M. johnstoni*, which grows to approximately 50 mm in length with a slender or shallow body depth while *M. katangae* shows a deeper body depth at the same length. The two important *Lacustricola* species in southern Africa, *L. myaposae* and *L. macrurus*, display body shapes similar to that of *M. johnstoni* (Skelton 2001; Marshall 2011).

The name ‘topminnow’ is derived from their characteristic habitat, which is generally near the surface of shallow water with weedy vegetation, and their mode of feeding which usually consists of small invertebrates such as mosquitoes on the water surface. Mosquito larvae and pupae may form an important component of their diet (Skelton 2001; Marshall 2011). Topminnows are normally found in small shoals where invertebrate prey is easily obtained. Breeding may occur all year (Marshall 2011) when food is abundant for topminnows, with eggs laid on or near vegetation.

Topminnows form an important resource and provide important ecosystem services for rural communities living along the edges of rivers and lakes such as the Okavango Delta (Steenkamp *et al.* 2001; Mmopelwa *et al.* 2009). Feeding on invertebrates such as mosquito larva is important for the control of mosquitoes and could further limit the spread of malaria (Steenkamp *et al.* 2001). They play an important role in local fisheries as bait fish for larger predatory fish, such as *Hydrocynus vittatus* (tiger fish) and *Clarias gariepinus* (African sharptooth catfish) (Skelton 2001; Marshall 2011), that are sold at markets and for export. Species such as *M. johnstoni* and *M. hutereaui* are popular in the aquarium trade, further adding to their economic value (Mmopelwa *et al.* 2009; Siziba *et al.* 2011). Topminnows have also been shown to be good indicators of environmental health in rivers and lakes, making them candidates for environmental toxicology studies (Steenkamp *et al.* 2001). Accurate taxonomic assessment of these fishes would be crucial for all of these studies.

The importance of taxonomic assessments for conservation and management strategies has increased in recent years due to the presence of cryptic species (Bickford *et al.* 2006). The ability to identify fish species and their life stages accurately has benefitted conservation efforts and allowed effective management policies to be implemented (Ward *et al.* 2005). Resolving taxonomic controversies has often been specific to populations or species that have been thought to be more valuable for the ecosystem or to industry because of certain characteristics that these species may possess or display (Vrijenhoek 1998). Their perceived diversity, evolutionary uniqueness or charismatic value may increase the priority of research done on these species and may lead to other species being neglected (Vrijenhoek 1998).

The Procatopodidae is diverse in species but there are gaps in its taxonomy, with various name changes and transfers of species into new or existing genera (Ghedotti 2000; Braganca & Costa 2019). The older taxonomic studies were primarily focused on phylogenetic assessments that drew on the use of morphological traits and did not take molecular evidence into consideration. A newer study of the African topminnows (Braganca *et al.* 2018) looked at family-level taxonomy using molecular techniques but did not discuss species-level taxonomy. Species-level taxonomy of the subfamilies Aplocheilichthyinae and Procatopodinae has not been updated and is poorly described (Skelton 2001; Steenkamp *et al.* 2001). The southern African procatopodines are ripe for taxonomic study at the species level.

Traditional taxonomic methods have been found to be limiting in highly diverse groups, especially in freshwater ichthyofauna (Teletchea 2009). For example, the haplochromine species that inhabit the African lakes have shown to be highly diverse, with morphological traits diverging in parallel and making cladistic approaches to identifying species incredibly difficult (Vrijenhoek 1998; Koblmüller *et al.* 2010). New methods that emerged over the past 20 years have been successful at resolving evolutionary lineages within hyperdiverse lake ichthyofaunas (Koblmüller *et al.* 2010). These methods, which include molecular studies using mitochondrial DNA (mtDNA) and nuclear DNA markers, have allowed more in-depth studies of the relationships of species (Packer *et al.* 2009; Pereira *et al.* 2013). New morphological methods have also allowed more in-depth analysis of the body shape of freshwater fishes. These methods have been successful in revealing cryptic species and helped to implement new conservation and management strategies for highly diverse fish species (Vrijenhoek 1998; Packer *et al.* 2009; Pereira *et al.* 2013). This thesis aims to apply some of these new techniques to a taxonomic assessment of the southern African procatopodines at the species level.

1.2. The southern African Procatopodidae

The species of focus for this study, *Micropanchax johnstoni* (Günther 1894), *M. katangae* (Boulenger 1912), *M. hutereaui* (Boulenger 1913), *L. myaposae* and *L. macrurus*; have undergone individual taxonomic revisions since their description.

Haplochilus johnstoni **Günther, 1894** was described from near Fort Johnston, Malawi. Its placement in *Haplochilus* Agassis, 1846 was not supported for long, and was transferred to the genus *Aplocheilichthys* Bleeker, 1863 (Wilderkamp 1986; Skelton 1993; Lazara 2001; Skelton 2001; Snoeks 2004). In 1999, Huber re-diagnosed the taxon as *Lacustricola johnstoni* but this re-diagnosis was not supported and Ghedotti (2000) suggested that the genus should be *Micropanchax* Myers, 1924. This was supported and *Micropanchax johnstoni* retained the current status (Ghedotti 2000; Marshall 2011; Fricke *et al.* 2018).

Micropanchax johnstoni includes the synonym *Haplochilus carlislei* van der Horst, 1934 (Fricke *et al.* 2018).

Of the southern African procatopodids, *Micropanchax johnstoni* occurs most widely, from Angola to Malawi and South Africa. It is found in Lake Rukwa and Lake Malawi, the Cunene, Okavango, Zambezi, Pungwe, Buzi and Limpopo River systems, with the Congo River system in northern Zambia (Skelton 2001; Mmopelwa *et al.* 2009; Marshall 2011). It usually inhabits well-vegetated, shallow areas, often along the banks of rivers, river backwaters, floodplains and swamps (Wildekamp 1995; Skelton 2001).

***Haplochilus katangae* Boulenger, 1912** was described from the Lubumbashi River near Elisabethville, Katanga, now in the Democratic Republic of Congo. The genus was transferred to *Lacustricola* Myers 1924 (Huber 1999) and then to *Aplocheilichthys* (Wildekamp 1986; Skelton 1993; Lazara 2001; Skelton 2001). Previous taxonomic work by Huber (1982) and Ghedotti (2000) places all but one species from *Aplocheilichthys* in the newly promoted *Micropanchax,* which is still supported and *Micropanchax katangae* remains the valid current status (Marshall 2011; Fricke *et al.* 2018). *Micropanchax katangae* includes the synonym *Aplocheilus luluae* Fowler, 1930 (Fricke *et al.* 2018).

Micropanchax katangae can be found across the top half of southern Africa in the Cunene, Okavango and Zambezi River systems. Its distribution stretches into the Mozambique coastal plain and south into KwaZulu-Natal, near Richards Bay, South Africa (Wildekamp 1995; Skelton 2001; Marshall 2011). It is usually found in dense vegetation of the marginal areas of rivers and is uncommon in floodplains (Skelton 2001).

***Haplochilus hutereaui* Boulenger, 1913** was described from the upper Uelé River, Dungu, Democratic Republic of Congo. *Haplochilus* was later transferred to *Aplocheilichthys* (Wildekamp *et al.* 1986; Lévèque *et al.* 1991; Skelton 1993). Huber (1999) proposed that *A. hutereaui* should be placed in *Lacustricola* but this was not widely accepted and it remained *A. hutereaui* (Lazara 2001; Skelton 2001). The work done by Ghedotti (2000) was in agreement with work done by Huber (1982) and places *A. hutereaui* in *Micropanchax,* with *M. johnstoni* and *M. katangae*. *Micropanchax hutereaui* remains the current status for this fish (Marshall 2011; Fricke *et al.* 2018). *Micropanchax hutereaui* includes the following synonyms: *Micropanchax baudoni* Myers, 1924, *Aplocheilus chobensis* Fowler, 1935, and *Aplocheilichthys schalleri* Scheel & Radda, 1974 (Fricke *et al.* 2018).

Micropanchax hutereaui is distributed mainly in the Okavango and upper Zambezi systems. This distribution stretches into the lower Shire and the Pungwe-Buzi river systems. It may also be found in the northern Zambia and may occur in the northern reaches of the Congo River (Wildekamp 1995; Skelton 2001; Marshall 2011). It is almost exclusively found in floodplains and marshes surrounding rivers (Huber 1996; Skelton 2001; Moelants 2015).

***Haplochilus myaposae* Boulenger, 1908** was described from the “Myaposa” River (= Mposa River), Zululand, in the KwaZulu-Natal province of South Africa. The genus was later transferred to *Aplocheilichthys* (Wildekamp *et al.* 1986; Lazara 2001; Skelton 2001). Huber (1999) placed the species in the genus *Lacustricola* and the species was moved to the subfamily Procatopodinae. *Lacustricola myaposae* remains the current status for this fish, without synonyms (Fricke *et al.* 2018).

Lacustricola myaposae is more narrowly distributed, occurring in coastal lakes and rivers in Maputaland and KwaZulu-Natal, South Africa (Skelton 2001; Mmopelwa *et al.* 2009 Marshall 2011). It inhabits well-vegetated inshore areas of lakes and streams (Skelton 2001).

***Haplochilus macrurus* Boulenger, 1904** was described from Lake Sarmento, Marimba, Angola. The genus was later transferred to the genus *Aplocheilichthys* (Wildekamp *et al.* 1986; Lazara 2001; Skelton 2001). Paugy *et al.* (1990) suggested that A. macrurus was a synonym of *Aplocheilichthys schioetzi* Scheel 1968. Huber (1999) transferred *A. macrurus* to the genus *Lacustricola* and the subfamily Procatopodinae. *Lacustricola macrurus* has remained the species current status, without synonyms (Fricke *et al.* 2018).

Lacustricola macrurus occurs in central and southern Africa. The central African distribution includes the upper Kasai River drainage and the Congo River basin. In southern Africa, *L. macrurus* is found in the Kunene, Kwanza and Okavango river systems in Angola and Namibia (Huber 1996). It is found in small rivers, streams, brooks and shallow marginal areas of lakes and other larger water bodies (Huber 1996).

1.3. Taxonomic history and relationships of the African Aplocheilichthyinae and Procatopodinae

In a taxonomic revision that involves phylogenetic analysis, such as this thesis, it is important to identify a suitable out-group. This entails understanding assessing the likely close relatives of the target taxa, which are *Micropanchax* and *Lacustricola* in this case.

Unfortunately, the taxonomy of the Aplocheilichthyinae and Procatopodinae has been poorly understood until recently. Regan (1911) was the first to suggest a close relationship between fishes now placed under these subfamilies. This led to the expansion of the genus *Aplocheilichthys* and the reassignment of many species from *Aplocheilus* McClelland, 1839 to the newly expanded *Aplocheilichthys* (Ghedotti 2000). Regan further suggested a close relationship among the genera *Aplocheilichthys*, *Procatopus* Boulenger, 1904 and *Lamprichthys* Regan, 1911 (Ghedotti 2000). Fowler (1916) created two new subfamilies, the Procatopinae and Lamprichthyinae. The creation of Procatopinae was contested by Myers (1924), who argued that *Procatopus* was not distinctive enough to be given the rank of subfamily. Later Myers (1928) described the tribe Aplocheilichthyini to include the genera *Aplocheilichthys, Procatopus, Plataplochilus* Ahl, 1928, *Hypsopanchax* Myers, 1924, *Platypanchax* Ahl, 1928 and *Cynopanchax* Ahl, 1928. He then placed *Lamprichthys* within the subfamily Lamprichthyinae and in 1955 described the Procatopodinae to replace Aplocheilichthyini without giving an adequate explanation or reasoning behind this name change (Ghedotti 2000).

The Aplocheilichthyinae was stated to be one of the well-corroborated monophyletic groups of fishes by Parenti (1981), and the genus *Aplocheilichthys* at that time contained over 60 nominal species. The type species of *Aplocheilichthys* was originally named *Poecilia spilauchena* Duméril 1861, currently *A. spilauchen* Duméril, 1861 (Parenti 1981). *Micropanchax* Myers 1924 (based on the species *Haplochilus schoelleri* Boulenger 1904 [= *Micropanchax loati* (Boulenger 1901)] from Egypt) was demoted to a subgenus of *Aplocheilichthys* for many years (Huber 1982; Ghedotti 2000). *Aplocheilichthys* was later recognized as a monotypic genus containing only *A. spilauchen*, which led to the other species then placed under *Aplocheilichthys* being moved to the resurrected genus *Micropanchax* (Huber 1982; Ghedotti 2000).

Lacustricola Myers 1924 was described as a subgenus of *Micropanchax*, based on the species *Haplochilus pumilus* Boulenger 1906 from Lake Tanganyika, but is currently "[r]etained for now as a valid genus since a number of species are treated as valid in *Lacustricola*" (Fricke *et al.* 2018). Similarly, *Poropanchax* Clausen 1967, based on the species *Aplocheilichthys macrophthalmus* Meinken 1932 (= *Fundulopanchax luxophthalmus* Brüning, 1929), has also been treated as a subgenus of *Micropanchax*, but is currently treated as a valid genus (van der Zee et al. 2007).

Systematic and taxonomic work on the Aplocheilichthyinae and Procatopodinae halted after the revisions of Parenti (1981) and Huber (1982), until a small study by Costa (1996) placed the Neotropical genus *Fluviphylax* Whitely, 1965 within the Aplocheilichthyinae and described two tribes in the subfamily. *Fluviphylax* fell under the tribe Fluviphylacini and all other aplocheilichthyines were placed in the tribe Aplocheilichthyini (Costa 1996; Ghedotti 2000). The subfamily names were both subsequently recognized as separate subfamilial-rank taxa, with *A. spilauchen* remaining the sole member of the genus *Aplocheilichthys* and the subfamily Aplocheilichthyinae, and all other species falling under Procatopodinae (Fowler 1916; Ghedotti 2000).

Ghedotti (2000) further recognized the genus *Micropanchax*, in Procatopodinae, with three subgenera: *Poropanchax* Clausen, 1967, *Lacustricola* and *Micropanchax*. This revision is similar to that previously proposed by Huber (Parenti 1981; Huber 1982; Ghedotti 2000). Two clades, roughly characterized by their overall body shape, form the Procatopodinae. The first clade consists of approximately seven genera, which typically display a body form that is laterally compressed and deep-bodied (Ghedotti 2000). The second clade houses the primitively terete proctopodines under *Micropanchax*. Each clade is known to have at least one species that has evolved pelagic habits independently within the African Great Lakes (Ghedotti 2000). Ghedotti (2000) also recognized two tribes within the Procatopodinae: Fluviphylacini, which contained *Fluviphylax*, and Procatopodini, which consisted of the genera *Micropanchax*, *Platypanchax* Ahl, 1928, *Lamprichthys* Regan, 1911, *Pantanodon* Myers, 1955, *Hypsopanchax* Poll, 1965, *Procatopus* Boulenger, 1904, *Cynopanchax* Ahl, 1928 and *Plataplochilus* Ahl, 1928.

Braganca *et al.* (2018) presented a study to resolve the phylogenetic position of *Pantanodon* and provide a clearer picture of the taxonomy at family level. The study found that the families Cyprinodontidae and Poeciliidae, as then understood, were not monophyletic due to the intrusion of other families and thus new classifications at the family level were proposed. The genus *Pantanodon* was supported as a sister group to the clade of the Cyprinodontoids and was placed in its own family Pantanodontidae Myers 1955 (Braganca *et al.* 2018). The New World livebearers, the Poeciliines were found to be the sister group of the Anablepidae and not of the African species. The Old World genera *Aphanius* and *Valencia* showed to be closer to the African species, which led the African species and the American species to be placed in separate families (Braganca *et al.* 2018). The African species were placed in the family Procatopodidae Fowler 1916, which now houses the subfamilies

Aplocheilichthyinae and Procatopodinae, whereas the Poeciliidae will now consist only of the American livebearers (Braganca *et al.* 2018).

The last revelation of the study by Braganca *et al.* (2018) was that the South American genus *Fluviphylax* was not closely related to the African species as previous morphometric studies had indicated and was placed into its own family, Fluviphylacidae Roberts 1970. This change removes *Fluviphylax* from the African clade and helps to clear up some of the taxonomic confusion of Cyprinodontoidei (Braganca *et al.* 2018). The phylogenetic tree created by Braganca *et al.* (2018) showed the phylogenetic position of the families relative to one another and suggested new relationships and taxonomy.

Braganca & Costa (2019) completed a more in-depth analysis of the family Procatopodidae. The study agreed with previously mentioned studies that the Procatopodidae formed a monophylic assemblage and confirmed a closer sister relationship to the families Aphaniidae and Valenciidae. The study also confirmed that *Fluviphylax* is a distinct lineage which belonged to its own family, Fluviphylacidae (Roberts 1970), and sister clade to the Poeciliidae and Anablepidae (Braganca & Costa 2019). When the study compared the internal relationships of the Procatopodidae, *Plataplochilus* was found to be the most basal lineage while *Aplocheilichthys spilauchen* formed the sole member of a sister clade to all other Proacatopodidae with the exception of *Plataplochilus*. These results showed a different story of the lineages of the Procatopodidae than previously thought to be the case as shown by Huber (1999) and Ghedotti (2000) using morphologically based analyses.

The study revealed that *Lacustricola Micropanchax* and *Poropanchax* were paraphyletic genera. Two species from *Lacustricola*, *L. pumilus* and *L. lualabaensis* Poll, 1938 were shown to group within the *Congopanchax* (Poll 1971) and *Lamprichthys* clade while the rest of the *Lacustricola* species placed in the same clade as *Micropanchax, Poropanchax* and *Rhexipanchax* Huber, 1999 (Braganca & Costa 2019). *Micropanchax scheeli* (Roman 1971) was shown to be more closely related to *Poropanchax* while *Poropanchax normani* Ahl, 1928 was shown to be the sole member of a sister clade to *Poropanchax, Micropanchax* and *Rhexipanchax* (Braganca & Costa 2019). The paraphyletic nature of *Lacustricola* and *Micropanchax* would possibly lead to new generic names being introduced from the work completed in the study at a later date (Braganca & Costa 2019; Braganca *et al.* in review). The phylogenetic analysis provided by Braganca & Costa (2019) has provided clarity on many of the relationships of the species and genera within the

Procatopodidae and updated much of the information from previous studies such as those done by Parenti (1981), Huber (1999) and Ghedotti (2000).

1.4. Scope and aims

The scope of the study was limited to species of *Micropanchax* and *Lacustricola* occurring in the southern African region, including Angola, Namibia, Botswana, Zambia, Malawi, Mozambique and South Africa. This study aims to review the taxonomy of the southern African topminnows with special attention to *M. johnstoni*, *M. katangae* and *M. hutereaui*. Preliminary data were collected by staff of the South African Institute for Aquatic Biodiversity (SAIAB) for these species and gaps in this information need to be identified and tested using genetic methods. DNA will be used to assess the relationships of the different populations, and a morphometric study to assist the identification of the species. Finally, new taxa will be diagnosed and described if the need arises.

The main aim of this thesis is to investigate genetic and morphometric variation among and within populations of the three *Micropanchax* and two *Lacustricola* species from different rivers systems across southern Africa to further the understanding of their status and relationships. The secondary aims are therefore to:

- Analyse data from three mitochondrial and one nuclear gene from specimens of each population of *Micropanchax johnstoni*, *M. katangae* and *M. hutereaui* to describe their taxonomic and biogeographical relationships.
- Use morphometrics to identify phenotypic markers for genetic clades and genetically distinct populations.
- Where necessary, revise the taxonomy of the populations of the three species to reflect the findings of the first three aims, and to describe any new species that may arise or to reinstate previous synonymized names.

Allowances were made for the discovery and description of new species, and the possible identification of populations that may require independent management plans.

1.5. References

AHL, E. 1928. Beiträge zur Systematik der afrikanische Zahnkarpfen. *Zoologischer Anzeiger* **79 (3-4)**: 113-123.

AHL, E. 1928. Descriptions of two new cyprinodont fishes from Nigeria. *Journal of Natural History* **2(12)**: 600-602.

AGASSIZ, L. 1846. *Nomenclatoris zoologici index universalis: continens nomina systematica animalium classium, ordinum, familiarum et generum animalium omnium, tam viventium quam fossilium, secundum ordinem alphabeticum unicum disposita, adjectis homonymiis plantarum*. Sumptibus et typis Jent et Gassmann.

BICKFORD, D., LOHMAN, D. J., SODHI, N. S., NG, P. K. L., MEIER, R., WINKER, K., INGRAM, K. K. & DAS, I. 2006. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22**: 148-155.

BLEEKER, P. 1863. *Mémoire sur les poisons de le côte de Guinée.* (Vol. 1). Loosjes.

BOULENGER, G.A. 1904. Descriptions of new West-African freshwater fishes. *Annals and Magazine of Natural History* **14**: 16-20.

BOULENGER, G.A. 1906. Fourth contribution to the ichthyology of Lake Tanganyika1. – Report on the collection of fishes made by Dr. WA Cunnington during the third Tanganyika expedition, 1904-1905. *The Transactions of the Zoological Society of London* **17(5)**: 537-600.

BOULENGER, G.A. 1908. On a collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum* **1**: 219-235.

BOULENGER, G.A. 1912. Description d'un poisson nouveau du genre *Haplochilus* provenant du Katanga. *Revue de Zoologie Africaine* **2**: 47-48.

BOULENGER, G.A. 1913. Sur une petite collection de poisons recueillis dans l'Uelé par la mission dirigée par *M, hutereaui*. *Revue de Zoologie Africaine* **2**: 155-161.

BOULENGER, G.A. 1913. LXXII. Description of a new cyprinodont fish from the Sobat River. *Journal of Natural History* **12 (72)**: 566-566.

BRAGANCA, P.H.N., AMORIM, P.F. & COSTA, W.J.E.M. 2018 . Pantanodontidae (Teleostei, Cyprinodontiformes), the sister group to all other cyprinodontoid killifishes as inferred by molecular data. *Zoosystematics and Evolution* **94 (1)**: 137-145.

BRAGANCA, P.H.N. & COSTA, W.J.E.M. 2019. Multigene fossil-calibrated analysis of the African lampeyes (Cyprinodontoidei: Procatopodidae) reveals an early Oligocene origin and Neogene diversification driven by palaeogeographic and palaeoclimatic events. *Organisms Diversity and Evolution* **19(2)**: 303-320.

BRÜNING, C. 1929. Der grüne Leuchtaugenfisch und der afrikanische Glanzkärpfling. *Wochenschrift für Aquarien- und Terrarienkunde* **26 (23)**: 356-356.

CLAUSEN, H.S. 1967. Tropical Old World cyprinodonts: reflections on the taxonomy of tropical Old World cyprinodonts, with remarks on their biology and distribution. *Akademisk Forlag:* 1967.

COSTA, W.J.E.M. 1996. Relationships, monophyly and three new species of the Neotropical miniature poeciliid genus *Fluviphylax* (Cyprinodontiformes: Cyprinodontoidei). *Ichthyological Exploration of Freshwaters* **7**: 111-130.

DUMÉRIL, A.H.A. 1861. Reptiles et poisons de l'Afrique occidentale. Tome 10. Etude précedée de considérations générals sur leur distribution géographique. *Archives du Muséum d'Histoire Naturelle* **10**: 258.

DUMÉRIL, A.H.A. 1928. Poissons de la côte occidentale d'Afrique. *Archives du Muséum d'Histoire Naturelle, Paris* **10**: 241-268.

FOWLER, H. W. 1916. Notes on the fishes of the orders Haplomi and Microcyprini. *Proceedings of the Academy of Natural Sciences, Philadelphia* **68**: 415-439.

FOWLER, H.W. 1930. The fresh-water fishes obtained by the Gray African Expedition – 1929. With notes on other species in the academy collection. *Proceedings of the Academy of Natural Sciences of Philadelphia* **82**: 27-83.

FOWLER, H.W. 1935. Scientific results of the Vernay-Lang Kalahari expedition, March to September, 1930. *Annals of the Transvaal Museum* **16**: 251-293.

FRICKE, R., ESCHMEYER, W.N. & VAN DER LAAN, R. (eds). 2018. Catalog of fishes: genera, species, references (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Accessed 6 December 2018.

GHEDOTTI, M.J. 2000. Phylogenetic analysis and taxonomy of the poecilioid fishes (Teleostei: Cyprinodontiformes). *Zoological Journal of the Linnean Society* **130**: 1-53.

GÜNTHER, A. 1894. Second report on the reptiles, batrachians, and fishes transmitted by Mr. H. H. Johnston, C. B., from British Central Africa. *Proceedings of the Zoological Society of London 1893* (pt 4): 616-628, Pls. 53-57.

HRBEK, T., SECKINGER, J. & MEYER, A. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution* **43**: 986-998.

HUBER, J. H. 1982. *A review of the cypronodont fauna of the coastal plain in Rio Muni, Gabon, Congo, Cabinda, and Zaire, with taxonomic shifts in* Aphyosemion, Epiplatys and *West African Procatopodins.* British Killifish Association, London.

HUBER, J.H. 1996. Killi-Data 1996. Updated checklist of taxonomic names, collecting localities and bibliographic references of oviparous cyprinodont fishes (Atherinomorpha, Pisces). *Société Française d'Ichtyologie, Muséum National d'Histoire Naturelle*, Paris, France, 399 p.

HUBER, J.H. 1999. Updates to the phylogeny and systematics of the African aplocheilichthyins, the lampeye schooling cyprinodonts (Cyprinodontiformes: Aplocheilichthyinae). *Cybium* **23**: 53-77.

KOBLMÜLLER, S., EGGER, B., STURMBAUER, C. & SEFC, K.M. 2010. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution* **55**: 318-334.

LAZARA, K.J., 2001. The killifishes, an annotated checklist, synonymy, and bibliography of recent oviparous cyprinodontiform fishes. *The killifish master index*, **4**.

LÉVÈQUE, C., PAUGY, D. & TEUGELS, G. G. 1991. Annotated check-list of the freshwater fishes of the Nilo-Sudan river basins, in Africa. *Revue d'Hydrobiologie Tropicale* **24**: 131-154.

LUCINDA, P.H. 2003. Family Poeciliidae. *Check list of the freshwater fishes of south and Central America*, pp. 555-581.

MARSHALL, B. 2011. *The fishes of Zimbabwe and their biology*. Smithiana Monographs. **3**: i-vi + 1-290.

MCCLELLAND, J. Indian Cyprinidae. *Journal of Asiatic Researchers* **19 (2)**: 217-471.

MCCLELLAND, J. 1938. Observations on six new species of Cyprinidae, with an outline of a new classification of the family. *Journal of the Asiatic Society of Bengal* **7**: 941-948.

MMOPELWA, G., MOSEPELE, K., MOSEPELE, B., MOLEELE, N. & NQWENYA, B. 2009. Environmental variability and the fishery dynamics of the Okavango Delta, Botswana: the case of subsistence fishing. *African Journal of Ecology* **47**: 119-127.

MOELANTS, T. 2015. Diversity and ecology of the ichthyofauna of the middle and Upper Congo basin: a case-study in the region of the Wagenia falls (Democratic Republic of Congo). *Katholieke Universiteit Leuven, Faculty of Science, Leuven (Belgium)*, 320p.

MYERS, G.S. 1924. A new poeciliid fish of the genus *Micropanchax*, from Ubangi. *American Museum Novitates* **122**: 1-3.

MYERS, G.S. 1924. New genera of African poeciliid fishes. *Copeia* **129**: 41-45.

MYERS, G. S. 1928. The systematic position of the phallostethid fishes, with diagnosis of a new genus from Siam. *American Museum Novitates* **295**: 1-12.

MYERS, G.S. 1955. Notes on the classification and names of cyprinodont fishes. *Tropical Fish Magazine*. March 1955: 7-7.

PACKER, L., GIBBS, J., SHEFFIELD, C. & HANNER, R. 2009. DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources* **9**: 42-50.

PARENTI, L. R. 1981. A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History* **168**: 335-557.

PEREIRA, L.H.G., HANNER, R., FORESTI, F. & OLIVEIRA, C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics* **2013**: 14-20.

POLL, M. 1938. Poissons du Katanga (basin du Congo) récoltés par le professeur Paul Brien. *Revue de Zoologie et de Botanique Africaines* **30(4)**: 389-423.

POLL, M. & LAMBERT, J.G. 1965. Contribution ah l'etude systématique et zoogéographique des Procatopodinae de l’Afrique centrale (Pices, Cyprinodontidae). *Bulletin des Seances. Académie Royale des Sciences d’Outre-mer* **1965**: 615-631.

POLL, M. 1971. Un genre nouveau et une sepèce nouvelle de Cyprinodontidae congolais. *Revue de Zoologie et de Botanique Africaines* **83(3-4)**: 302-308.

PAUGY, D., LÉVÊQUE, C., TEUGELS, G.G., BIGORNE, R. & ROMAND, R. 1990. Freshwater fishes of Sierra Leone and Liberia. Annotated checklist and distribution. *Revue d’Hydrobiologie Tropicale* **23 (4)**: 329-350.

REGAN, C.T. 1911. The osteology and classification of teleostean fishes of the order Microcyprini. *Annals and Magazine of Natural History* **8**: 320-327.

ROMAN, B. 1971. Nuevas especies de peces de Río Muni (Guinea Ecuatorial). *Publicaciones del Instituto de Biología Aplicado Barcelona* **49**: 5-23.

SCHEEL, J.J. 1968. *Rivulins of the old world*. TFH Publications, Jersey City.

SCHEEL, J.J. & RADDA, A.C. 1974. Beschreibung zweier neuer Cyprinodontidae aus dem tropischen Afrika. *Aquaria: Vivaristische Fachzeitschrift für die Schweiz und Oesterreich* **21 (11)**: 157-162.

SIZIBA, N., CHIMBARI, M.J., MOSEPELE, K. & MASUNDIRE, H. 2011. Spatial and temporal variations in densities of small fishes across different temporary floodplain types of the lower Okavango Delta, Botswana. *African Journal of Aquatic Science* **36**: 309-320.

SKELTON, P. H. 1993. *A complete guide to the freshwater fishes of southern Africa.* Southern Book Publishers, Johannesburg.

SKELTON, P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town.

SKELTON, P.H. 2002. Changes to the scientific and common names of southern African freshwater fishes. *African Journal of Aquatic Science* **27**: 171-174.

SNOEKS, J. 2004. The non-cichlid fishes of the Lake Malawi system: a compilation. p. 20-26. In Snoeks, J. (ed.) *The cichlid diversity of Lake Malawi/Nyasa/Niassa: identification, distribution and taxonomy*. Cichlid Press, El Paso, USA.

STEENKAMP, M.K.J., ENGELBRECHT, G.D. & MULDER, P.F.S. 2001. Allozyme variation in a Johnston's topminnow, *Aplocheilichthys johnstoni*, population from the Zambezi River system. *Water SA* **27**: 53-55.

TELETCHEA, F. 2009. Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries* **19**: 265-293.

TWEDDLE, D. 2010. Overview of the Zambezi River System: Its history, fish fauna, fisheries, and conservation. *Aquatic Ecosystems Health and Management* **13**: 224-240.

VAN DER HORST, C.J. 1934. Two new fishes from Transvaal. *Annals of the Transvaal Museum* **15 (3)**: 281-282.

VRIJENHOEK, R.C. 1998. Conservation genetics of freshwater fish. *Journal of Fish Biology* **53**: 394-412.

WARD, R. D., ZEMLAK, T. S., INNES, B. H., LAST, P. R. & HEBERT, P. D. N. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society* **360**: 1847-1857.

WHITLEY, G.P. 1965. Some fish genera scrutinized. *Proceedings of the Royal Zoological Society of New South Wales* **1964-65**: 25-26.

WILDEKAMP, R.H., ROMAND, R. and SCHEEL, J.J. 1986. Cyprinodontidae. *Check-list of the freshwater fishes of Africa. CLOFFA: ISNB, Brussels, MRAC.*165-276.

WILDEKAMP, R.H. 1995. A world of killies. Atlas of the oviparous cyprinodontiform fishes of the world. Volume II. The genera *Aplocheilichthys*, *Aplocheilus*, *Austrofundulus*, *Campellolebias*, *Crenichthys*, *Cualac*, *Cubanichthys*. *Cynolebias*, *Cynopoecilus* and *Cyprinodon*. *The American Killifish Association, Inc*. 384 p.

Table 1.1. Summary of the currently accepted taxonomy of the species of Aplocheilichthyinae and Procatopodinae examined in this thesis.

Valid status	Original Description	Type locality	Habitat
Aplocheilichthyinae			
Aplocheilichthys spilauchen	*Poecilia spilauchena* Duméril 1861	Ogooué (Gabon) River	Freshwater/brackish
Procatopodinae			
Micropanchax johnstoni	*Haplochilus johnstoni* Günther 1894	Fort Johnston, Malawi	Freshwater
	= *Haplochilus carlislei* van der Horst 1934	Aaipes River, South Africa	Freshwater
Micropanchax hutereaui	*Haplochilus hutereaui* Boulenger 1913	Dungu, Upper Uelé River, DRC	Freshwater
	= *Micropanchax baudoni* Myers 1924	Fort Crampel, Gribingi River, CAR	Freshwater
	= *Aplocheilus chobensis* Fowler 1935	Chobe River, Botswana	Freshwater
	= *Aplocheilichthys schalleri* Scheel & Radda 1974	180km south of Beira, Mozambique	Freshwater
Micropanchax katangae	*Haplochilus katangae* Boulenger 1912	Lubumbashi River, Katanga, DRC	Freshwater
	= *Aplocheilus luluae* Fowler 1930	Lulua River, DRC	Freshwater
Micropanchax kingii	*Haplochilus kingii* Boulenger 1913	Khor on Sobat River, eastern Africa	Freshwater
Hypsopanchax jubbi	*Hypsopanchax jubbi* Poll & Lambert 1965	Upper Zambezi River, Mwinilunga District, Zambia	Freshwater
Lacustricola myaposae	*Haplochilus myaposae* Boulenger 1908	Myaposae River, Zululand, South Africa	Freshwater
Lacustricola macrurus	*Haplochilus macrurus* Boulenger 1904	Marimbo, Lake Sarmento, Angola	Freshwater
Poropanchax normani	*Aplocheilichthys normani* Ahl 1928	Kiyawa River, Katagum, Nigeria	Freshwater
Poropanchax luxophthalmus	*Fundulopanchax luxophthalmus* Brüning 1929	Unknown	Freshwater
	= *Aplocheilichthys macrophthalmus*	Near Lagos, Nigeria	

2

Species delimitation in Southern African topminnows (Cyprinodontiformes: Cyprinodontoidei).

2.1. Introduction

The topminnows (Procatopodidae: Procatopodinae) are a distinctive and speciose element of the African ichthyofauna, and apparently five species (*Micropanchax johnstoni* (Günther, 1894), *Micropanchax katangae* (Boulenger 1912), *Micropanchax hutereaui* (Boulenger 1913), *Lacustricola myaposae* (Boulenger 1908) and *Lacustricola macrurus* (Boulenger 1904)) occur in several river systems and lakes in southern Africa (Skelton 2001; Marshall 2011). Topminnows acquired their name from living in shallow waters near the margins of lakes and rivers and by feeding on small invertebrates on or near the water surface. The mouth of the topminnow opens upward to allow easier capture of invertebrates in the meniscus (Skelton 2001). Topminnows are important as part of the diet of larger predatory fish and in the biological control of mosquitoes and malaria (Skelton 2001; Steenkamp *et al.* 2001). Two southern African species, *M. johnstoni* and *M. hutereaui*, were popular in the aquarium trade, which has further increased the economic value of these fishes (Mmopelwa *et al.* 2009; Siziba *et al.* 2011). Since *Micropanchax* (Myers 1924) contained at least 14 species (Fricke *et al.* 2018; Froese *et al.* 2018), accurate taxonomic evaluation of the fishes would be crucial for further studies and management plans for these fishes.

The traditional taxonomy of the African Procatopodidae was based on their external morphological features, dentition, frontal sensory systems, and colouration, but their small size, interspecifically conservative but intraspecifically variable morphology and colouration that is affected by sex, allometry and geography, and high potential mobility, make their taxonomic assessment difficult (Huber 1982). Older research has focussed on integrating osteological data with morphology, with qualified success (Parenti 1981; Ghedotti 2000). The species-level taxonomy of the African subfamilies Aplocheilichthyinae and Procatopodinae has not been updated in the last 20 years, and was poorly described even then (Ghedotti 2000; Skelton 2001; Steenkamp *et al.* 2001; Lucinda & Reis 2005). Although morphology forms the basis of traditional taxonomy, it is unclear which populations differ sufficiently to indicate distinct species.

Parenti (1981) used osteological features to produce a cladistic classification of the cyprinodontiforms into five families: Cyprinodontidae, Anablepidae, Goodeidae, Jenynsiidae and Poeciliidae. The Poeciliidae included the African genera *Aplocheilichthys* (Bleeker, 1863) and *Micropanchax* (Myers, 1924); although at the time *Micropanchax* was considered to be a subgenus of *Aplocheilichthys*. Parenti (1981) described *Aplocheilichthys* as a polyphyletic group with four subgenera: *Micropanchax*, *Lacustricola* (Myers, 1924), *Poropanchax* (Clausen, 1967) and *Congopanchax* (Poll, 1971). Further revision of the genus was recommended.

Based on a cladistic analysis of morphological and osteological characters, Ghedotti (2000) partitioned Poeciliidae into three monophyletic subfamilies: Aplocheilichthyinae, Procatopodinae and Poeciliinae (Ghedotti 2000; Lucinda & Reis 2005). The Aplocheilichthyinae contained only *Aplocheilichthys spilauchen* (Duméril, 1861), and was sister to the other subfamilies; the Procatopodinae had two tribes, Fluviphylacini, containing the Neotropical genus *Fluviphylax* Whitley, 1965 and Procatopodini, containing several African genera of lamp-eyed killifishes, including *Micropanchax* with three subgenera: *Micropanchax* Myers, 1924, *Lacustricola* Myers, 1924 and *Poropanchax* Clausen, 1967 (Ghedotti 2000)[*]. All other species previously deemed to be in *Aplocheilichthys* were moved to *Micropanchax* including *M. johnstoni*, *M. hutereaui* and *M. katangae* (Boulenger 1912) (Huber 1982; Ghedotti 2000).

Subsequently, *Lacustricola* and *Poropanchax* were recognised as genera independent of *Micropanchax* by Huber (2005) and van der Zee *et al*. (2007), respectively. Van der Laan *et al*. (2014) listed Lamprichthyinae Fowler, 1916, Pantanodontinae Myers, 1955, Fluviphylacinae Roberts, 1970 and Micropanchina Huber, 2000 as synonyms of Procatopodinae Fowler, 1916, but the on-line Catalogue of Fishes (Fricke *et al.* 2018) currently lists the families Fluviphylacidae Roberts, 1970, Poeciliidae Bonaparte, 1831 and Procatopodidae Fowler 1916, without explanation. The explanation may lie in a multi-gene phylogenetic analysis that showed that Poeciliidae is paraphyletic, that Poeciliinae is related to the Anablepidae Bonaparte, 1831, and that the Valenciidae Parenti, 1981 (containing the single species *Valencia hispanica* (Valenciennes, 1846)) is the sister-group to the African Procatopodinae (Pohl *et al*. 2015; Reznick *et al*. 2017). The study by Braganca *et al.* (2018)

---

[*] Procatopodinae was deemed to include *Micropanchax* (~35 species), ‘Micropanchax’ (unnamed genus; 4 species), *Platypanchax* (2 species), *Lamprichthys* (1 species), *Pantanodon* (2 species), *Hypsopanchax* (5 species), *Pocatopus* (3 species), *Cynopanchax* (1 species) and *Plataplocheilus* (4 species) (Ghedotti 2000; cf. Huber 1982).

confirmed that the Poeciliines were found to be the sister group of the Anablepidae, while the Old World genera *Aphanius* and *Valencia* demonstrated closer relationship to the African species. This result led to the placement of the Poeciliines and the African sub-families in separate families, the African subfamilies, Aplocheilichthyinae and Procatopodinae, were placed in the family Procatopodidae Fowler 1916. The second family, Poeciliidae, consist of the only the New World livebearers (Braganca *et al.* 2018).

A newer study by Braganca and Costa (2019) provided a more in-depth analysis of the relationships in the family Procatopodidae. The phylogenetic analyses performed by Braganca and Costa (2019) confirmed a closer sister relationship between the Procatopodidae and the families Aphaniidae and Valenciidae. The study agrees with previous papers that the genus *Fluviphylax* belonged to its own family, Fluviphylacidae which had a closer sister relationship to the families Poeciliidae and Anablepidae. The Baraganca and Costa (2019) study took a deeper look into the relationships with the family Procatopodidae and found that while the family Procatopodidae had a monophylic assemblage, some of the genera within the family were shown to be paraphyletic. *Lacustricola* and *Micropanchax* were found to be paraphyletic. *Lacustricola pumilus* and *L. lualabaensis* were found to clade within the *Congopanchax* and *Lamprichthys* clade, with the remainder of the *Lacustricola* species grouped in the same clade as *Micropanchax, Poropanchax* and *Rhexipanchax*. *Micropanchax scheeli* showed a closer relationship to *Poropanchax* than *Micropanchax* and *P. normani* was found to be the sole member of a sister clade to *Micropanchax, Poropanchax* and *Rhexipanchax*. Finally *Plataplochilus* was discovered to be the most basal lineage of the Procatopodidae and *A. spilauchen* formed the sole member of a clade sister to all other Procatopodidae, with *Plataplochilus* being the only exception (Braganca & Costa 2019). The southern African lineages were poorly represented in these molecular analyses, which therefore warrant extension.

In recent years the use of DNA markers has acquired an important role in the identification of fish species and other organisms. The use of DNA markers has helped to distinguish populations or species of fishes that are morphologically similar but after closer inspection showed genetic variation and were found to be evolutionarily independent of one another (Steinke *et al.* 2009). The use of DNA in identification has become ever more popular due to the discovery of cryptic species and natural hybridization events that make detecting these taxa through morphology increasingly difficult (Hebert *et al.* 2003; Krück *et al.* 2013; van Zeeventer *et al.* in prep).

The mitochondrial protein-coding gene cytochrome c oxidase 1 (COI) was suggested to be ideal for identifying fish and other animal species because of two major advantages: COI has robustly universal primers that allow the retrieval of a sequence from most animal phyla (Hebert *et al.* 2003); and the phylogenetic signal range of COI was greater than that of most other mitochondrial genes (Hebert *et al.* 2003). The COI gene is relatively easy to amplify and sequence and most organisms produce many mitochondria within cells, including anucleate red blood cells (Hebert *et al.* 2003). The relatively rapid rate of evolution of COI allows the discrimination of closely-related species and populations of the same species from different geographic locations. The technique of using COI has been referred to as *DNA barcoding* (Hebert *et al.* 2003). These DNA barcode sequences are stored on the online servers of various global projects that allow submissions from independent laboratories around the world. An example of such a project is The Fish Barcode of Life Initiative (FISH-BOL, http://www.fishbol.org) (Hebert *et al.* 2003; Krück *et al.* 2013).

COI barcoding is cost-effective, but has been criticized for over-simplification, specifically in the use of only one mitochondrial gene when considering the wide spectrum of differentiation among species and hybrids already recognized (Krück *et al.* 2013), giving rise to a multi-gene barcoding approach that uses both mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) markers (Krück *et al.* 2013). This approach has yielded greater genetic resolution as the discovery of cryptic species and natural hybridization events grows, and additionally delivers more robust estimates of phylogenetic relationships. The use of multiple genetic markers in establishing relationships is advantageous because different genes may evolve at different rates and reflect different aspects of evolutionary histories. This allows studies to provide an accurate or more in-depth interpretation of evolutionary history and relationships within and between species (Rubinoff 2006; Krück *et al.* 2013; Pohl *et al.* 2015; Bragança *et al.* 2018). The need for increased awareness of these events is important for fisheries management and conservation efforts (Krück *et al.* 2013), and goes beyond mere identification.

The combination of markers that may be best-suited to answering particular question can be highly variable and choosing markers that suit the project is important. The different evolutionary rates of genes can determine if they suite a project and choosing genes that have a spectrum of rates of evolution can help the analysis so that misleading (unsuitably fast) or uninformative (unsuitably slow) information is ameliorated (Rubinoff 2006). For this reason, this study has made use of a multi-gene analysis using 28S ribosomal RNA (28S rRNA) (a nuclear ribosomal gene), 16S ribosomal RNA (16S rRNA) (a mitochondrial ribosomal gene), COI (a mitochondrial protein-coding gene), cytochrome b (cyt*b*) (another mitochondrial protein-coding gene), and the mitochondrial Tyrosine Kinase (X-src) (TyrK) to identify relationships between the focal species of *Micropanchax*: *M. johnstoni*, *M. katangae* and *M. hutereaui*. The COI marker was chosen for its previous success at identifying fish species, for the role it plays in barcoding programs such as The Fish Barcode of Life Initiative (Hebert *et al.* 2003), and because its relatively rapid evolutionary rate helps to reconstruct close phylogenetic relationships between species or populations (Brown et al 1979; Ward et al. 2005; Rubinoff 2006).

The only nuclear marker, 28S rRNA, was chosen for its slow evolutionary rate to investigate evolutionary and divergence events that occurred earlier in the history of the fishes. It is a large subunit nuclear rRNA gene and shows more variation in rates of evolution than its small subunit counterparts (Hillis & Dixon 1991). The markers 16S rRNA and cyt*b* were chosen because their intermediate evolutionary rates help to bridge the mutation rates of the COI and 28S rRNA markers.

The study aims to use a multi-gene approach to assess the diversity and relationships among the southern African topminnows of the subfamily Procatopodinae, with specific interest in *M. johnstoni*, *M. katangae*, *M. hutereaui*, *L. myaposae* and *L. macrurus*. While this study is not aimed at exploring the phylogeny of *Micropanchax* and *Lacustricola*, it would be useful to include as many African topminnow taxa as possible to test the monophyly of, and distinctions between, species. The available taxa include *Micropanchax kingii* (Boulenger 1913), *Hypsopanchax platysternus* (Nichols & Griscom 1917), *Poropanchax normani* (Ahl 1928), *P. luxophthalmus* (Brüning 1929), *Procatopus nototaenia* (Boulenger 1904), *P. aberrans* (Ahl 1927), *Lamprichthys tanganicanus* (Boulenger 1898) and *A. spilauchen* (Duméril 1861).

2.2. Methods

2.2.1. Specimens and tissue samples

Specimens and tissue samples were obtained from the South African Institute for Aquatic Biodiversity (SAIAB) collection, Makhanda, South Africa. They were collected, mostly by SAIAB staff under permit, from across the region, and include specimens collected in the field by the author (Table 2.1). Colections made by the author were done under permission of Kamutjonga Inland Fisheries Institute (KIFI), Namibia and a permit housed by them. Ethical clearance was obtained from Rhodes University for the collection and euthanasia of specimens in the field. Ethical clearance number: RU-LAD-15-03-0003. Specimens from the type locality for both *M. katangae* and *M. hutereaui* could not be obtained for the analysis but topotypical specimens were available for *M. johnstoni, L. myaposae* and *L. macrurus*.

In the field, fish selected for genetic analysis were anaesthetized using clove oil. A photograph of the left side of each fish was taken to record its live colouration. Specimens that did not have a tissue sample removed from them in the field were placed into containers of ethanol to preserve the specimen without destroying its DNA. A waterproof label was placed into each of the containers recording the collection date, sample number, location (including GPS coordinates), details of capture and preservation, and the specimen number. A data sheet was also completed with the same information and the live colouration photographs' numbers.

Tissue samples were removed from specimens in either the field or the laboratory. Muscle tissue from the end point of the dorsal fin to the beginning of the caudal fin was removed carefully and placed into 95% ethanol in separate microcentrifuge tubes with a recorded tissue sample number. Fin clippings were taken from specimens too small to provide sufficient muscle tissue and placed into 95% ethanol. The specimen was then placed into 10% formalin for short-term storage and later transferred to 75% ethanol for long-term storage. The DNA and whole specimens (Table 2.1) were stored at the South African Institute for Aquatic Biodiversity (SAIAB), Makhanda.

Fish collected in the field were identified by R. M. van Zeeventer and I. R. Bills and sorted into storage jars at the SAIAB collection laboratory. Other SAIAB material (Table 2.1) was identified by past SAIAB curators.

2.2.2. DNA data

Tissue samples were completely dried prior to DNA extraction and placed into separate new microcentrifuge tubes for the extraction protocol. DNA was extracted using the GeneJet® Genomic DNA Purification kit (ThermoFisher Scientific) and the NucleoSpin® Tissue kit (Machery-Nagel Gmbh & Co. KG) following the manufacturer's standard protocol for animal tissue isolation. The tissue was incubated for approximately one hour until complete tissue digestion occurred. All other standard steps in the protocol were followed and no volumes were altered.

The concentration and purity of each DNA sample was determined using a NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific). The values for DNA concentration, A260, A280, 260/280, and 260/230 were recorded.

For the COI polymerase chain reaction amplification (PCR) the primers used were LCO1490 and HCO2198 (Folmer *et al.* 1994); for the 28S rRNA PCR, 28S-2F and 28S-2R (Unpublished data); for the 16S rRNA PCR, 16S-AR and 16S-BR (Pohl *et al.* 2015); and for the cyt*b* PCR, GluDGL and H16460 (Perdices *et al.* 2012) (Table 2.2).

For the amplification of 28S rRNA, reactions were conducted in 20 µl reaction volume and included 1 µl of DNA (100 ng), 0.5 µl forward primer (10 pmol), 0.5 µl reverse primer (10 pmol), 10 µl Master Mix (2x) (Kapa Biosystems) and 8 µl dH_2O. For the amplification of the COI, 16S and genes, reactions included 1-3 µl $MgCl_2$ (25 mM) (Kapa Biosystems). Amplifications were conducted in an Applied Biosystems Veriti 96-well thermocycler PCR machine. The primers 28S-2F and 28S-2R on average amplify a 800 bp fragment of the 28S gene. The conditions for the PCR included an initial denaturation temperature of 94°C for 5 min for a single cycle, followed by an annealing of 94°C for 45 s, 58°C for 45 s and 72°C for 2 min; this was repeated for 35 cycles. The final elongation stage was run for 5 min at 72°C.

For 16S rRNA, the primers 16S-AR and 16S-BR amplified a 529 bp fragment of the gene for *Procatopus nototaenia* (subfamily: Procatopodinae) (Pohl *et al.* 2015). The conditions for the PCR amplification of 16S were an initial denaturation of 95°C for 3 min, annealing of 95°C for 50 s, 50°C for 30 s and 72°C for 50 s for 35 cycles. The final elongation stage was run for 10 min at 72°C.

For COI, the primers LCO1490 and HCO2198 amplify an average of 658 bp in a wide range of different vertebrate species (Hebert *et al.* 2003). The conditions for the PCR were initial denaturation of 95°C for 5 min, annealing of 95°C for 1 min, 43-45°C for 1 min and 72°C for 2 min, repeated for 35 cycles. The final elongation was run for 5 min at 72°C.

For cyt*b*, the primers GluDGL and H16460 amplified 1140 bp of the gene for certain fish species (Perdices *et al.* 2012; Bagley *et al.* 2013). The conditions for the PCR were an initial denaturation of 94°C for 4 min, annealing of 94°C for 1 min, 48-50°C for 1 min and 72°C for 3 min for 35 cycles. The final elongation stage was run for 10 min at 72°C.

The PCR products were electrophoresed on a 1% agarose gel at 100 V for 30 min. The PCR products were sent to Macrogen Inc for capillary electrophoresis sequencing using both forward and reverse primers for all four markers. The final nucleotide sequences were edited in BioEdit sequence alignment software (Hall 2004). Consensus sequences were generated in BioEdit by aligning the forward and reverse sequences using a ClustalW multiple sequence alignment (Larkin *et al.* 2007). All of the sequences for a marker were aligned using BioEdit and ClustalX (Larkin *et al.* 2007) and finally saved in Nexus format (*.nex files).

2.2.3. Phylogenetic analysis

The phylogenetic conclusions of morphological studies (Parenti 1981; Huber 1982, 1999; Ghedotti 2000) have been substantially altered by recent molecular studies (van der Laan *et al*. 2014; Fricke *et al*. 2017; Bragança *et al*. 2018; Braganca & Costa 2019) that removed many African genera from the Poeciliidae and created the Procatopodidae. These newer studies did not include sufficient African topminnow genera to make the choice of an out-group obvious for the current study, so several out-groups were chosen from within the cyprinodontiform families (Table 2.3). Each out-group species (Table 2.3) was chosen based on how complete a set of target genes (COI, cyt*b*, 16S, and 28S) were available on GenBank (http://www.ncbi.nlm.nih.gov/genbank/) for it. In addition, sequences for TyrK were downloaded from GenBank for potential out-group taxa (Table 2.3) because of their utility in recent molecular phylogenetic analyses of the out-groups (van der Laan *et al*. 2014; Fricke *et al*. 2017; Bragança *et al*. 2018; Braganca & Costa 2019).

A preliminary maximum likelihood tree was created using the tyrosine kinase gene to determine which species may work best as out-group taxa (Table 2.3). The top candidates, for the use of outgroup were *Aplocheilichthys* and *Valencia* (Fig. 2.2). *Aplocheilichthys* was know to be the sister shown to be a sister clade to the species of Procatopodidae and Valenciidae was found to be more closely related to the Procatopodidae than Poeciliidae (Braganca & Costa 2019). The maximum likelihood tree was created using Tyrosine kinase marker (TyrK) to determine the best outgroups for the phylogenetic tree and to determine the position of the root in cases where the same out-group could not be used. The tree was created in Mega6 and was run for a total of a 1000 bootstrap replications (Fig. 2.2).

Exceedingly few sequences for the target genes were available for any other species of *Micropanchax* and *Lacustricola* despite their being the largest genera in the subfamily. COI was retrieved from GenBank for *M. hutereaui*, and COI and 16S for *Micropanchax kingii* (Boulenger, 1913) (Table 2.3). Sequences for *L. macrurus* were downloaded from the Barcode of Life Data System (BOLD) (Ratnasingham & Hebert 2007). The sequences from each genetic marker were aligned with their respective GenBank sequences using BioEdit (Hall 1999; Larkin *et al.* 2007) and saved as FASTA files (*.fas).

For all of the genetic markers, saturation plots (transition *versus* transversion graphs) were created in DAMBE (Data Analysis in Molecular Biology and Evolution) (Xia 2017). The saturation plots and the substitution saturation test (Xia *et al.* 2003; Xia & Lemey 2009) were used to determine if mutational saturation had been reached and which of the genetic marker data sets could be used for confident phylogenetic analysis (Table 2.4) (Fig. 2.1).

The model of best-fit for the genetic markers was determined using JModelTest2 (Darriba *et al.* 2012) and PartitionFinder (Lanfear *et al.* 2016) using Bayesian Information Criterion (BIC) weightings (Schwarz 1978). JModelTest2 was run on the CIPRES Science Gateway (Miller *et al.* 2010) using a substitution scheme equal to three to produce MrBayes (Ronquist & Huelsenbeck 2003) acceptable models. PartitionFinder was also run on the CIPRES Science Gateway using the BIC model selection criteria, greedy search choice parameter and model choice being set to all. A Bayesian inference tree was then created using the model of best-fit for each marker in MrBayes on the CIPRES Science Gateway. The Bayesian inference tree used a total of 20 000 000 generations and the Markov chain Monte Carlo (MCMC) was run for four chains, and had two independent runs. The nexus files for the Bayesian analysis were set up following a partitioned approach with each gene forming its own partion and model parameters were unlinked for each partition. The protein coding genes COI and cyt*b* were further partitioned by codon position. The trace files of the Bayesian analysis showed that a burn-in value of 25% should be removed and that convergence of the trees in the analysis had been reached with the trace file ESS values all exceeding 200. The trace files were evaluated in Tracer (Rambaut *et al.* 2014). The final consensus trees were created in FigTree version 1.4 (Rambaut 2012) and TreeView (Page 1996).

A maximum likelihood analysis was completed for all the genetic markers and the total molecular data using GARLI in the CIPRES Sciences Gateway (Miller et al. 2010). The maximum likelihood for the total molecular data included partitions for each gene and codon partitioning for the two protein-coding genes. The maximum likelihood analysis included the same models employed in the Bayesian inference analysis. The maximum likelihood tree was created using a total of 1000 bootstrap replications. The output file, allboottrees.tre, containing all the trees created in GARLI for the total molecular analysis was uploaded to the program Mesquite (Maddison & Maddison 2018) to create a consensus tree. Lastly the tree was edited on the International Tree of Life website (https://itol.embl.de/: Ciccarelli *et al.* 2006).

To assess ambiguities behind the phylogeny estimated using Bayesian inference, neighbor-joining networks were created using Neighbor-Net analyses of uncorrected P-distances in the Splitstree package (Huson & Bryant 2006) and the aligned files for each genetic marker, except 28S rRNA. The neighbor-joining network was not created for the gene fragment due to the low resolution of 28S rRNA in the Bayesian inference tree; it was believed that due to its low resolution, the neighbor network would be too compressed to read.

2.2.4 Species delimitation

A Bayesian General Mixed Yule-Coalescent Model (bGMYC) of the total molecular dataset was run in the statistical package R version 1.2.1335 (RStudio Team 2018). Four phylogenetic trees were created using Beast version 1.10.4 (Suchard *et al.* 2018) for the bGMYC. The input files for Beast were created in Beauti version 1.7.5 (Drummond *et al.* 2012) and consisted of two models split between the four trees. The Yule and coalescent models were employed and set to use a relaxed and a strict clock for each model, thus creating two trees for each model. The input files were then run through Beast on the CIPRES Science Gateway to create the four phylogenetic trees. Once the trees were completed, their posterior distributions were summarised using TreeAnnotator version 1.7.5 (Drummond *et al.* 2012) and uploaded to R.

In R, the trees were converted to the ultrametric format using the package APE and the analysis was completed using the bGMYC package. The analysis was run for all four trees (MCMC = 80 000, Burnin = 40 000, thinning = 100) with the parameters allowing 400 tree samples to be retained in the analysis. The Yule strict clock tree was chosen as the best-fit dataset for the bGMYC species delimitation as it had the lowest BIC score of the four trees.

Bayesian Poisson tree processes (bPTP) species delimitation was completed using the Bayesian inference tree for the total molecular data on the The Exelixis Lab Species Delimitation Server (https://species.h-its.org/ptp/) (Zhang *et al.* 2013). The analysis was shown to reach convergence in the MCMC iterations plot (Fig 2.15). The bPTP analysis was rooted on *Valencia hispanica* and run at 200000 MCMC generations, thinning of 100 and burnin of 0.1. In both the bGMYC and bPTP, out-groups were removed from the analysis in an attempt to keep the analysis as accurate as possible.

2.3. Results

2.3.1. Total molecular evidence

A total length of 3607 base pairs made up the concatenated sequences for 82 samples included in a total molecular evidence analysis. The data was concatenated in the gene order of COI, cyt*b*, 16S rRNA, 28S rRNA and TyrK. Tyrosine kinase was used to help bring greater resolution to the out-groups and therefore only concatenated to the out-group sequences where possible.

The Bayesian inference for the total molecular evidence was rooted on *V. hispanica*. The out-group species branched out earlier than the in-group demonstrating a more distant relationships to the in-group species, with *A. spilauchen* appearing on its own clade within the out-groups and showed a relatively distant relationship to *Micropanchax* (branch posterior probability = 1.00). *Micropanchax* and *Lacustricola* together formed ambiguously monophyletic group with *M. kingii* forming a clade basal to all of the in-group species (branch posterior probability = 1.00) (Fig. 2.4).

Micropanchax hutereaui was straightforward to interpret. It formed a sister clade to a clade comprising all of the other Southern African taxa (branch posterior probability = 0.99). The first, Clade A, contained a specimen from the Shire River, Malawi and specimens from the Zambezi and Okavango Rivers in Namibia. The second clade, Clade B, showed two subclades between the three specimens (branch posterior probability = 1.00) of *M. hutereaui* (branch posterior probability = 1.00). Specimens from the Okavango River appeared only in Clade A while specimens from the Zambezi River were found in both clades (Fig. 2.4).

Micropanchax johnstoni was shown to be polyphyletic and specimens were split across three clades that were each more closely related to other species. Clade C was sister to a clade of *L. myaposae* and contained specimens from the type locality of *M. johnstoni* and specimens identified as *L. myaposae*. Clade D was the sister clade to a single, distinct specimen that could not be identified morphologically, and was comprised of specimens identified as *M. johnstoni* and *L. macrurus*. Clade E was sister to the *M. katangae* clade. The three clades of *M. johnstoni* were not geographically isolated, with specimens from the same river systems found in each of the clades (Fig. 2.4).

Micropanchax katangae was monophyletic (branch posterior probability = 1.00), with two genetically distinctive specimens from the Shinengeni Stream and Lake Curumbo. The *M. katangae* clade showed a close relationship to Clade E from three river systems, the Kwando and Zambezi Rivers in Namibia and the Okavango River system crossing Angola, Namibia and Botswana.

Lacustricola macrurus from Olushandja Dam found in Clade D showed no close relationship to *L. macrurus* collected directly from the Kunene River, which appeared in a separate clade that showed a closer relationship to Clade C (branch posterior probability = 1.00).

Lacustricola myaposae showed a similar situation to *L. macrurus* with specimens from the Lake Sibaya system forming a clade sister to Clade C (branch posterior probability = 0.80), while specimens of *L. myaposae* from the Kosi Bay system further north lay within Clade C (Fig. 2.4).

The Maximum likelihood analysis for the total molecular data was rooted on *V. hispanica*, with the only difference in the relationships of the out-groups being that *M. kingii* diverged before *A. spilauchen* (branch posterior probability = 0.91), which now showed a closer relationship to the in-group specimens than in the Bayesian inference tree but with weaker support (branch posterior probability = 0.72) (Fig. 2.5). The in-group species showed the same topology as the Bayesian inference analysis with *M. hutereaui* still forming basal to the other groups (branch posterior probability = 0.75). The overall support for the analysis was good and comparable to that of the Bayesian inference analysis with the only two major splits showing weak support. The split between *M. katangae* and Clade E with *L. macrurus*, *L. myaposae*, Clade C and Clade D had weak branch posterior probability value of 0.68. The split between Clade D and *L. myaposae*, *L. macrurus* and *M. johnstoni* (Clade C) showed weaker branch posterior probability value of 0.65 (Fig. 2.5).

The Neighbor-net analysis showed all of the out-groups to diverge on one side of the network and showed a more distant relationship to the in-group species of *Micropanchax*. The in-group species showed similar relationships to those of the Bayesian inference tree, with one clade for *M. katangae*, two clades for *M. hutereaui* (Clades A and B), and three clades for *M. johnstoni* (Clades C, D and E). Clade C contained specimens from the type locality of *M. johnstoni* and showed a close relationship to *L. macrurus* from the Kunene River and *L. myaposae* from the Kosi Bay and Lake Sibaya systems. The specimens of *L. myaposae* from the Lake Sibaya system formed another clade distinct from *L. macrurus* from the Kunene River and Clade C (Fig. 2.6).

Clade D showed similar grouping of specimens to the Bayesian inference tree with *L. macrurus* from Olushandja Dam having a close relationship to the specimens of *M. johnstoni* despite being geographically isolated from them. *Lacustricola macrurus* from the Kunene River once again showed no close relationships to *L. macrurus* from Olushandja Dam, the latter grouping in Clade D. Clade D showed a close relationship to the single unidentified specimen from the Kwanza River but also a potentially earlier divergence than the Kwanza River specimen (Fig. 2.6). Clade E again showed a close relationship to *M. katangae* but also a fairly long branch, as five of the specimens were plotted along the length of this branch (Fig. 2.6). The *M. katangae* clade plotted with different specimens showed individual splits from one another but all formed closely together. The two specimens from the Shinengeni Stream and Lake Curumbo on the Bayesian inference tree branched separately on the neighbor-net analysis but showed the relationship may not be as close as the Bayesian inference estimate suggested (Fig. 2.6).

Micropanchax hutereaui formed two clades, as in the Bayesian inference analysis. Clade A contained the bulk of the specimens, mostly from rivers in the Caprivi Strip, Namibia, and Clade B contained three specimens, two from Namibia and one from Zambia(Fig 2.6).

2.3.2. 28S rRNA

A total of 62 sequences were produced or downloaded from GenBank for 28S (Tables 2.1, 2.3), and only *P. reticulata* was available as an out-group. They were aligned with difficulty using the BioEdit software package (Hall 1999; Larkin *et al.* 2007) and T-Coffee web portal (Notredame *et al.* 2000), and produced an alignment with a total of 743 bp. The transition vs transversion scatterplot indicated some mutational saturation but the saturation test (Xia 2003; Xia and Lemey 2009) showed that the sequences were not saturated (Table 2.4) (Fig. 2.1).

The model of best-fit for the Bayesian inference indicated by PartitionFinder (Guindon *et al.* 2010; Lanfear *et al.* 2016) for 28S rRNA was TRN+G. The Bayesian inference analysis (Fig. 2.3) showed very little phylogenetic resolution, with formation of only two clades in the in-group. The tree showed a lot of mixing between the specimens but did place the bulk of *M. johnstoni* specimens with *L. myaposae*, and *L. macrurus*, while *M. katangae* grouping in the same clade. The bulk of *M. hutereaui* grouped together in its own clade on the tree. The 28S gene therefore contributed to the resolution of in-group relationships, even though it had practically no value for rooting the tree.

Due to the low phylogenetic resolution of 28S, the Neighbor-Net analysis did not produce any significant insight into the tree and was therefore not presented.

2.3.3. COI

A total of 661 base pairs of COI were sequenced for 82 samples (Tables 2.1, 2.3) that were easily aligned to one another including out-groups acquired from GenBank. The scatterplot for COI showed no mutational saturation (Fig. 2.1), which was confirmed by the saturation test by Xia (2003) for each of the three codon positions (Table 2.4).

The model of best-fit for the Bayesian inference indicated by PartitionFinder (Guindon *et al.* 2010; Lanfear *et al.* 2016) for the COI gene was the same for all three codon positions, HKY+I+G. The Bayesian inference tree placed most of the candidate out-groups outside the in-group (Fig. 2.7). The rest of the specimens in the COI tree showed the same pattern as the total molecular evidence tree, with *M. hutereaui* forming Clades A and B with the same member specimens. *Micropanchax johnstoni* was once again not monophyletic and formed the same clades as the total molecular evidence tree with Clades C, D and E in the same topology (Fig. 2.7). *Lacustricola macrurus* and *Lacustricola myaposae* showed the same relationships on the COI tree and specimens were placed in the same clades shown in Figure 2.4.

The neighbor-net analysis (Fig. 2.8) shared a similar topology to the tree generated using Bayesian inference analysis. The out-group specimens formed on the side of the neighbor network and showed similar groupings to the Bayesian inference tree (Fig. 2.8). *Aplocheilichthys spilauchen* clustered with the out-groups and showed a distant relationship to the study taxa (Fig. 2.8).

The in-group species on the neighbor-net analysis showed similar grouping of specimens to the Bayesian inference tree with Clade C including *L. myaposae* from the Kosi Bay system and specimens of *M. johnstoni* from the type locality. *Lacustricola myaposae* from the Lake Sibaya system showed a close relationship to this group and *L. macrurus* collected from the Kunene River (Fig 2.8). The lone unidentified specimen from the Kwanza River once again lay isolated but closest to Clade D, which contained *L. macrurus* collected from Olushandja Dam, and showed that these specimens were separated from *L. macrurus* from the Kunene River (Fig. 2.8).

Micropanchax katangae formed its own clade with the two specimens from the Shinengeni Stream and Lake Curumbo from Zambia and Angola, respectively, placed on their own. Clade E branched near to *M. katangae* and did not show a close relationship with Clades C or D (Fig. 2.8). *Micropanchax hutereaui* showed the same clade grouping as the Bayesian inference analysis (Fig. 2.8).

2.3.4. Cytb

A total of 1076 base pairs were sequenced for 33 samples (Tables 2.1, 2.3) and were easily aligned to one another including out-groups acquired from GenBank. The scatterplot for cyt*b* showed no mutational saturation and the test for saturation by Xia (2003) showed no saturation at each of the codon positions (Table 2.4, Fig. 2.1).

The model of best-fit for the Bayesian inference indicated by PartitionFinder (Guindon *et al.* 2010; Lanfear *et al.* 2016) for the cyt*b* gene was split between the three codon positions, with the first two codon positions models being HKY+I+G and the third codon position model TRN+G. The Bayesian inference tree was rooted on *F. cingulatus* and demonstrated a clear split between the out-groups and the in-group. The Bayesian inference analysis showed a clear split between *A. spilauchen* and *Micropanchax* with good support (branch posterior probability = 0.99) (Fig. 2.9). Although the cyt*b* tree had fewer specimens than the COI tree, the same clades emerged for *M. hutereaui* and *M. johnstoni*, with the only difference in these species being a closer relationship between *M. hutereaui* and Clades C and D (branch posterior probability = 0.87) (Fig. 2.8). One specimen of *M. katangae* from Chingueia Tributary, Angola, showed a close relationship to *M. hutereaui* and not to the rest of the *M. katangae* specimens.

Clade E followed the overall pattern of the previous trees and showed a close relationship to *M. katangae*. *Micropanchax katangae* showed a slight difference from the previous trees, with a split within the *M. katangae* clade. The smaller group contained specimens from the Zambezi River only, while the larger contained specimens from throughout the range of *M. katangae* (Fig. 2.9). The final two specimens of *M. katangae* from the Shinengeni Stream and Lake Curumbo showed genetic distinction from the rest of the *M. katangae* specimens.

The Neighbor-net analysis (Fig. 2.10) showed all of the out-group species to diverge on their own branches on one side of the network, *A. spilauchen* branching near the middle of the neighbor-net and showing a more distant relationship with *Micropanchax*, which mirrored the observed pattern in the total molecular and COI trees. *Micropanchax johnstoni* again formed three clades, and once again Clades C and D showed the closest relationship to one another while Clade E showed a closer relationship to *M. katangae*, which only formed one clade on the neighbor-net analysis, unlike the two clades in the Bayesian inference tree (Fig. 2.10). The specimens from Clades A and B of *M. hutereaui* seemed to form one clade on the neighbor-net analysis with a split between the two clades having occurred near the end of clade branch.

2.3.5. 16S rRNA

A total of 629 bp was sequenced for 16S rRNA, and 44 samples (Tables 2.1, 2.3) including out-groups were aligned using BioEdit (Hall 1999; Larkin *et al.* 2007). The in-group samples were easily aligned to one another, but aligned to out-groups only with some difficulty. The scatterplot for 16S rRNA showed no mutational saturation and the saturation test by Xia (2003) indicated no saturation (Table 2.4, Fig. 2.1).

The model of best-fit for the Bayesian inference indicated by PartitionFinder (Guindon *et al.* 2010; Lanfear *et al.* 2016) for the 16S gene was TVMEF+G. The Bayesian inference tree, rooted on *Poropanchax luxophthalmus*, showed separation between the out-group and in-group species (branch probability = 0.57) (Fig. 2.11). *Micropanchax kingii* was shown here to be more basal than *A. spiluachen,* which was a similar result in the Maximum likelihood analysis but was not well supported. The Bayesian inference tree showed the same pattern and clades as the total molecular data and COI trees with Clades A and B of *M. hutereaui* in the same position and showed a more basal relationship to the rest of the in-group species. *Micropanchax johnstoni* formed the same three clades as previously, with Clade C showing a close relationship to *L. myaposae* from Lake Sibaya system, Clade D including *L. macrurus* from Olushandja Dam and Clade E splitting from *M. katangae*. The clades within the in-group specimens showed good support values throughout the tree (Fig. 2.11).

The neighbor-net analysis showed similar groupings to that of the Bayesian inference tree, where the out-groups radiated away from the in-group specimens, and in-group specimens showed similar splits to the Bayesian inference analysis. *Micropanchax hutereaui* showed one specimen branching before the others and a slight split between specimens of the A and B Clades (Fig 2.12). *Micropanchax johnstoni* was shown to split into three clades. Clade C included specimens from different rivers and had a close relationship to *L. myaposae* from the Mseleni River, Lake Sibaya system. This small group also contained one unidentified specimen assigned to *Micropanchax*. A similar grouping was shown in the Bayesian inference tree. Clade D showed a mixture of specimens identified as *M. johnstoni* (from the Kafue, Okavango and Kwando Rivers), *L. macrurus* (from Olushandja Dam) (Fig. 2.12). Clade E of remained close to *M. katangae* as seen in the other molecular markers and Bayesian inference trees. Two specimens from the Kwanza River showed genetic distinctness and grouped together on both the Bayesian tree and the neighbor-net analysis and could be seen by one specimen in the total molecular and COI trees.

2.3.6. Species Delimitation

The results of the Bayesian Generalized Mixed Yule-Coalescent (bGMYC) method indicated that there could be five possible species within the inference tree. The bGMYC showed that five groups were significantly different from each other. *Micropanchax katangae* clade, Clade D and Clade E all formed their own candidate species in the analysis while both clades of *M. hutereaui* formed one candidate species and *L. macrurus*, *L. myaposae* and *M. johnstoni* Clade C formed one candidate species. The analysis showed that within the Clade C, *L. macrurus* and *L. myaposae*, and *M. hutereaui* groups, the difference within the two groups was not significant enough to separate them as their own species although the Bayesian inference showed them forming their own clades (Fig. 2.13).

The Bayesian Poisson Tree Processes analysis showed more presumed species than the bGMYC with a total of 13 candidate species. The first candidate new species from the bPTP is the unidentified *Micropanchax* sp from the Kwanza River with a Bayesian support value of 1.00. Clade D of *M. johnstoni* and the *L. macrurus* was shown to form one species group by the bPTP (Bayesian support = 0.74), this result agreed with the bGMYC (Fig. 2.14). The two specimens of *M. katangae* from lake Curumbo, Angola, and Shinengeni Stream, Zambia, that were isolated in the Bayesian and maximum likelihood analysis were shown here to each be candidate species (Bayesian support = 1.00 and 086 respectively). *Micropanchax hutereaui* Clade A (Bayesian support = 0.83), *M. johnstoni* Clade C (Bayesian support = 0.68), *M. johnstoni* Clade E (Bayesian support = 0.86), *L. macrurus* (Bayesian support = 0.60) and *M. katangae* (Bayesian support = 0.86) were shown here to be presumed species. *Micropanchax hutereaui* Clade B showed a split between the three specimens, with the two specimens from Namibia forming one species (Bayesian support = 0.80) and the specimen from Zambia its own species, but the single specimen showed weak independence (Bayesian support = 0.47) (Fig. 2.14). The final clade of *L. myaposae* showed a split into two candidate species with two specimens from Mseleni River (Bayesian support = 0.95) forming one group, and the other containing three specimens, one from Lake Mgobozeleni and two originally unidentified specimens from Hluhluwe River (Bayesian support = 0.70).

2.4. Discussion

A clear perspective on the southern African populations of *Micropanchax* emerged from the analysis. Consistent relationship patterns were provided by the neighbor-net and Bayesian analyses of the total evidence and individual genes, maximum likelihood and, the species delimitation analyses which supported the overall monophyly of the Procatopodidae of southern Africa with the exception of *M. johnstoni*, which appear to represent three phylogenetically independent lineages.

As expected of a slowly-evolving gene (Mindell & Honeycutt 1990), the nuclear marker 28S showed little to no species-specific variation, but still provided some information to resolve in-group relationships. It did not contribute to rooting because data were not available for most out-group taxa, but this role was filled by TyrK. The mtDNA markers COI, and 16S reflected relatively rapid rates of substitution without saturation (at least within the Procatopodidae). This allowed mtDNA to be particularly useful for species-level questions when close phylogenetic relationships for specific taxa were unknown (Rubinoff 2006). The Cytochrome c oxidase I gene had an evolutionary rate that was approximately three times faster than that of 16S and allowed discrimination between the closely-related species and groups within single species (Hebert *et al.* 2003). The faster substitution rate of COI and the different substitution rates of the other molecular markers allowed for testing at different historical levels (Hebert *et al.* 2003; Rubinoff 2006).

2.4.1. Micropanchax

Micropanchax hutereaui (Boulenger 1913) and *M. katangae* (Boulenger 1912) showed low genetic variation within their respective clades (Fig. 2.10). *Micropanchax hutereaui* showed low genetic variation between the Okavango, Kwando and Zambezi systems within Namibia’s Caprivi Strip and may be the result of inter-basin transfer or the proximity of the rivers that could connect via extensive flooding in the wet season (Schlettwein *et al.* 1990). The Kwando River terminates in the Linyanti River. The Linyanti River flowed into Lake Liambezi in northern Bostwana and the the Chobe River leaves the lake and finally terminates in the Zambezi River, connecting the two river systems and allowing the migration of fish (Schlettwein *et al.* 1990; Cronberg *et al.* 1995; Joyce *et al.* 2005; Kurugundla *et al.* 2010). The rivers were diverted in the Pleistocene to flow into a large drainage basin which formed the Lake palaeo-Makgadikgadi (Haddon & McCarthy 2005; Joyce *et al.* 2005). This lake may have allowed gene flow to occur between separate populations and when the lake emptied in the Holocene, the Cuando/Kwando, upper Zambezi and Linyanti Rivers were captured by the middle Zambezi River (Joyce *et al.* 2005).

Micropanchax katangae showed low genetic variation over a wider portion of its distribution in southern Africa. Most of the specimens come from geographically neighbouring rivers that may be connected to one another directly or indirectly. The two specimens from South Africa showed low genetic variation from the other specimens with no connectivity between the populations and large distances between the geographic areas.

The phylogenetic analysis of *M. johnstoni* (Günther 1894) revealed three distinct clades, which suggests the presence of cryptic species. The use of molecular methods is an effective first approach in provisionally identifying cases of cryptic species (Steinke *et al.* 2009). Convergent evolution of morphological characteristics among the three clades of *M. johnstoni* could cause the cryptic species to arise, where they would be identified in the field as *M. johnstoni* but show substantial genetic variation. The three different clades were collected from the same rivers and sampling sites and from the same river systems across their distribution. The species were all caught from similar habitats with similar vegetation structure and in most cases the same predation pressures from the same predators, such as tiger fish (*Hydrocynus vittatus*). The selective pressures may cause the animals to develop similar phenotypic traits and share morphology that makes them difficult to distinguish without DNA analysis (Endler 1982; Doolittle 1994; Tobler *et al.* 2011).

Convergent evolution is described as the adaption of similar phenotypic characteristics in two or more populations that are unrelated or distantly related (Endler 1982; Doolittle 1994) but the term parallelism or parallel evolution, which describes organisms that may be closely related so that the similarities between organisms are homologous, is more applicable in this situation (Reznick *et al.* 1996).

Micropanchax johnstoni was polyphyletic, branching closely with *L. macrurus* and *L. myaposae*. Clade E showed a strong relationship to *M. katangae* but shown in the species delimitation analyses to most likely be separate species. This could suggest that the specimens of Clade E were closely related to *M. katangae* and to be significantly different from *M. johnstoni* and this need to be investigated further as they showed distinctive genetic variation from the topotypes of *M. johnstoni*.

2.4.2. Lacustricola

Lacustricola was classified as a subgenus of *Micropanchax* (Ghedotti 2000) but the type species for *Lacustricola* was *Haplochilus pumilus* (Boulenger 1906) from Lake Tanganyika and showed a distribution primarily based in central Africa (Wildekamp 1995). The close relationship seen between the two *Lacustricola* species and Clades C and D suggest that *L. myaposae* and *L. macrurus* should be transferred to the genus *Micropanchax*. This elevation to the genus *Micropanchax* would leave *Lacustricola* to be primarily distributed in central and eastern Africa (Wildekamp 1995, Froese & Pauly 2018).

Lacustricola macrurus (Boulenger 1904) was described from Lake Sarmento, Northern Angola and reaches as far south as the Kunene River, Northern Namibia (Skelton 2001). This study showed two populations of *L. macrurus*, both found in the Kunene system, one in the Kunene River itself and the other in Olushandja Dam, which is connected to the river via a series of canals. One possible explanation for the genetic separation of the two groups may be due to human disturbances, such as the creation of canals. Canals can create new lacustrine environments that could contain more favourable habitats and could alter the flow regime of rivers or streams, allowing fish to disperse (Moyle & Marchetti 2006; Roman & Darling 2007). The canals that connect the dam to the river may provide the fish with a passage to travel to the dam, despite dam walls generally preventing upstream movement (Pelicice *et al.* 2015). The canal actually enters Olushandja Dam above the dam’s surface, which would act as a barrier similar to a waterfall preventing fish from traveling in one direction (Crass 1969), genetically isolating a population. This allowed the population from the dam to undertake a separate evolutionary path. Each population was found on a separate branch with strong support suggesting that the Olushandja Dam population may be a new species. A second explanation may be that the group of *L. macrurus* from Olushandja Dam showed a close relationship to one clade of *M. johnstoni*. This could suggest that both groups diverged from a common ancestor and may not have evolved from the *L. macrurus* population in the Kunene River.

Lacustricola myaposae (Boulenger 1908) was described from the Myaposa River (currently known as the Mposa River), KwaZulu-Natal, South Africa and has a narrow distribution within KwaZulu-Natal and Maputaland (Skelton 2001; Mmopelwa *et al.* 2009; Marshall 2011), normally occurring in coastal lakes and rivers. Two unidentified specimens collected from the Hluhluwe River, South Africa, grouped on the same branch as three specimens of *L. myaposae* from the Lake Sibaya System in KwaZulu-Natal, suggesting that they are also *L. myaposae*. Three specimens from the Nkanini and Mpungwune Rivers further north in KwaZulu-Natal and identified in the field as *L. myaposae* grouped with *M. johnstoni* showing a close relationship to topotypical specimens of *M. johnstoni*, and suggesting that the three specimens in question should be reidentified as *M. johnstoni* and not *L. myaposae*. If these results are correct then the known distribution of *M. johnstoni* should be adjusted, since the distribution of *M. johnstoni* is currently understood to extend only as far south as the Limpopo River System (Skelton 2001; Mmopelwa *et al.* 2009; Marshall 2011), while these results suggest it may enter northern KwaZulu-Natal near the border with Mozambique.

A number of specimens that were not identifiable in the field were included in the Bayesian and neighbor-net analyses, most of which grouped with identified specimens, revealing their identity; for example two specimens from the Hluhluwe River, South Africa, grouped with *L. myaposae*. The Bayesian and neighbor-net analyses showed two specimens for 16S and one for COI that were unidentified and appeared on their own branches. The two specimens from the Kwanza River, Angola, may be sufficiently distinctive to be recognised as new species showing strong branch support with neighbouring clades showing similar branch support. The Kwanza River specimens were shown in the bGMYC to belong to the same species as Clade D but was most likely not a member of *M. johnstoni* and was shown in the bPTP to be its own candidate species.

2.4.3. Aplocheilichthys

Although *Aplocheilichthys* is currently a monotypic genus, the specimens of *A. spilauchen* formed two very distinct lineages with a short common origin between the out-groups and showed no direct relationship to *Micropanchax*. This result agreed with Huber's (1982) and Ghedotti's (2000) promotion of the subgenus *Micropanchax* to the rank of genus and the reassignment of all species under *Aplocheilichthys* to the genus *Micropanchax*, leaving *A. spilauchen* the sole member of *Aplocheilichthys* and the only genus in the subfamily Aplocheilichthyinae. Huber (1982) and Ghedotti (2000) were correct to promote *Micropanchax* from subgenus to genus level and leave *A. spilauchen* as the sole member of *Aplocheilichthys*.

2.5. Conclusion

A number of cryptic species were found within *M. johnstoni*, primarily from the Okavango and Zambezi River Systems. Several reasons could cause the evolution of similar morphologies as they were all found in similar habitats with similar faunal and floral communities. *Lacustricola myaposae* and *L. macrurus* showed close relationships to *M. johnstoni*, with the only the bPTP splitting them into separate species. The morphological analysis of these species may shed further light on its current status.

A number of unidentified specimens were tested and the species delimitation analyses showed that a number of possible new species may have been uncovered. This may be confirmed with morphological analysis. Assessing the morphology of the populations is the goal for Chapter 3.

2.6. References

AHL, E., 1927. Neue afrikanische Fische der Familien Anabantidae und Cyprinodontidae. Sber. Ges. naturf. Freunde Berlin, **1972(4/7)**: 76-81.

AHL, E. 1928. LXXVII. Descriptions of two new Cyprinodont fishes from Nigeria. *Journal of Natural History*, ***2*(12)**: 600-602.

BAGLEY, J.C., SANDEL, M., TRAVIS, J., DE LOURDES LOZANO-VILANO, M. & JOHNSON, J.B. 2013. Paleoclimatic modelling and phylogeography of least killifish, *Heterandria formosa*: insights into Pleistocene expansion-contraction dynamics and evolutionary history of North American Coastal Plain freshwater biota. *BMC Evolutionary Biology* **13**: 1471-2148.

BAGLEY, J.C., ALDA, F., BREITMAN, M.F., BERMINGHAM, E., VAN DEN BERGHE, E.P. & JOHNSON, J.B. 2015. Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae). *PLOS One* **10(4)**: p.e0121139.

BOULENGER, G.A. 1898. Report on the fishes recently obtained by Mr. JES Moore in Lake Tanganyika. *Proceedings of the Zoological Society of London* **3**: 494-497.

BOULENGER, G.A. 1904. Descriptions of new West-African freshwater fishes. *Annals and Magazine of Natural History* **14**: 16-20.

BOULENGER. G.A. 1906. Fourth contribution to the ichthyology of Lake Tanganyika. Report on the collection of fishes made by Dr. W. A. Cunnington during the third Tanganyika Expedition, 1904-1905. *Transactions of the Zoological Society of London* **17(6, 1)**: 537-601.

BOULENGER, G.A. 1908. On a collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum* **1**: 219-235.

BOULENGER, G.A. 1912. Description d’un poisson nouveau du genre *Haplochilus* provenant du Katanga. *Revue de Zoologie Africaine* **2**: 47-48.

BOULENGER, G.A. 1913. Sur une petite collection de poisons recueillis dans l’Uelé par la mission dirigée par *M. Hutereau*. *Revue de Zoologie Africaine* **2**: 155-161.

BOULENGER, G.A., 1913. LXXII. Description of a new Cyprinodont fish from the Sobat River. *Journal of Natural History* **12(72)**: 566-566.

BRAGANÇA, P.H.N., AMORIM, P.F. & Costa, W.J.E.M. 2018. Pantanodontidae (Teleostei, Cyprinodontiformes), the sister group to all other cyprinodontoid killifishes as inferred by molecular data. *Zoosystematics and Evolution* **94(1)**: 137-145.

BRAGANCA, P.H.N. & COSTA, W.J.E.M. 2019. Multigene fossil-calibrated analysis of the African lampeyes (Cyprinodontoidei: Procatopodidae) reveals an early Oligocene origin and Neogene diversification driven by palaeogeographic and palaeoclimatic events. *Organisms Diversity and Evolution* **19(2)**: 303-320.

BROWN, W.M., GEORGE, M., WILSON, J.R. & WILSON, A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences* **76**: 1967-1971.

BRÜNING, C. 1929. Der grüne Leuchtaugenfisch und der afrikanische Glanzkärpfling. *Wochenschrift für Aquarien- und Terrarienkunde* **26(23)**: 356-356.

CHEN, W.J., BONILLO, C. & LECOINTRE, G. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular Phylogenetics and Evolution* **26(2)**: 262-288.

CHEN, X., GOU, S., GONG, Y., YUAN, F., XU, Y. & YANG, J. 2016. 16S rRNA and COI gene barcodes applied for identification of ornamental fishes. *Fujian Journal of Agricultural Sciences* **31(12)**: 1267-1272.

CICCARELLI, F.D., DOERKS, T., VON MERING, C., CREEVEY, C.J., SNEL, B. & BORK, P. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* **311(5765)**: 1283-1287.

CRASS, R.S. 1969. The effects of land use on freshwater fish in South Africa, with particular reference to Natal. *Hydrobiologia* **34(1)**: 38-56.

CRONBERG, G., GIESKE, A., MARTINS, E., PRINCE NENGU, J. & STENSTROM, I.M. 1995. Hydrobiological Studies of the Okavango Delta and Kwando/Linyanti/Chobe River, Botswana: I surface water quality analysis. *Botswana Notes and Records* **27**: 151-226.

CUVIER, G. & VALENCIENNES, A. 1846. Histoire naturelle des poisons. Tome dix-huitième. Cyprinoïdes. Livre dix-neuvième. *Des Èsoces ou lucioïdes* **18**: i-xix + 2pp + 1-505 + 2pp. 520-553.

DARRIBA, D., TABOADA, G.L., DOALLO, R. & POSADA, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.

DECRU, E., MOELANTS, T., DE GELAS, K., VREVEN, E., VERHEYEN, E. & SNOEKS, J. 2016. 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)**: 342-352.

DOOLITTLE, R.F. 1994. Convergent evolution: the need to be explicit. *Trends in Biochemical Sciences* **19**: 15-18.

DRUMMOND, A.J., SUCHARD, M.A., XIE, D. & RAMBAUT, A. 2012. Bayesian phylogenetics with Beauti and the Beast 1.7. *Molecular Biology and Evolution* **29(8)**: 1969-1973.

DUMÉRIL, A.H.A., 1861. Poissons de la côte occidentale d'Afrique. *Archives du Muséum d'Histoire Naturelle* **10**: 241-268.

ENDLER, J.A. 1982. Convergent and divergent effects of natural selection on colour patterns in two fish faunas. *Evolution* **36**: 178-188.

FOLMER, O., BLACK, M., HOEH, W., LUTZ, R. & VRIJENHOEK, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**: 294-299.

FREYHOF, J., KAERST, H. & Geiger, M.F. 2014. *Valencia robertae,* a new killifish from southern Greece (Cyprinodontiformes: Valenciidae). *Ichthylogical Exploration of Freshwaters* **24(4)**: 289-298.

FRICKE, R., ESCHMEYER, W.N. & VAN DER LAAN, R. (eds). 2018. Catalog of fishes: genera, species, references (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 11 November 2019.

FROESE, R. & PAULY, D. EDS. 2018. World Wide Web electronic publication (http://www.fishbase.org), version (02/2018). Electronic version accessed 11 November 2019.

GHEDOTTI, M.J. 2000. Phylogenetic analysis and taxonomy of the poecilioid fishes (Teleostei: Cyprinodontiformes). *Zoological Journal of the Linnean Society* **130**: 1-53.

GUINDON, S., DUFAYARD, J.F., LEFORT, V., ANISIMOVA, M., HORDIJK, W. & GASCUEL, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance pf PhyML 3.0. *Systematic Biology* **59(3)**: 307-321.

GÜNTHER, A. 1984. Second report on the reptiles, batrachians, and fishes transmitted by Mr. H. H. Johnston, C. B., from British Central Africa. *Proceedings of the Zoological Society of London 1893* (pt 4): 616-628, Pls. 53-57.

HADDON, I.G. & MCCARTHY, T.S. 2005. The Mesozoic-Cenozoic interior sag basins of Central Africa: The Late-Cretaceous-Cenozoic Kalahari and Okavango basins. *Journal of African Earth Sciences* **43**: 316-333.

HALL, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98.

HEBERT, P.D.N., CYWINSKA, A., BALL, S.L. & DEWAARD, J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* **270**: 313-321.

HECKEL, J.J. 1848. Eine neue gattung von Poecilien mit rochenartigem anklmmerungs-organe. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. *Mathematisch-Naturwissenschaftliche Classe* **1(1-5)**: 289-303.

HILLIS, D.M. & DIXON, M.T. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* **66**: 411-453.

HRBEK, T., SECKINGER, J. & MEYER, A. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution* **43(3)**: 986-998.

HUBER, J.H. 1982. *A review of the cypronodont fauna of the coastal plain in Rio Muni, Gabon, Congo, Cabinda, and Zaire, with taxonomic shifts in* Aphyosemion, Epiplatys and *West African Procatopodins.* British Killifish Association, London.

HUBER, J.H. 1999. Updates to the phylogeny and systematics of the African aplocheilichthyins, the lampeye schooling cyprinodonts (Cyprinodontiformes: Aplocheilichthyinae). *Cybium* **23**: 53-77.

HUSON, D.H. & BRYANT, D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254-267.

JOYCE, D.A., LUNT, D.H., BILLS, R., TURNER, G.F., KATONGO, C., DUFTNER, N., STURMBAUER, C. & SEEHAUSEN, O. 2005. An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature* **435**: 90-95.

KRÜCK, N.C., TIBBETTS, I.R., WARD, R.D., JOHNSON, J.W., LOH, W.K.W. & OVENDEN, J.R. 2013. Multi-gene barcoding to discriminate sibling species within a morphologically difficult fish genus (*Sillago*). *Fisheries Research* **143**: 39-46.

KURUGUNDLA, C.N., DIKGOLA, K., KALAOTE, K. & MPHO, M. 2010. Restoration and rehabilitation of Zibadianja lagoon in Kwando-Linyanti river system in Botswana. *Botswana Notes and Records* **42**: 79-89.

LANFEAR, R., FRANDSEN, P.B., WRIGHT, A.M., SENFELD, T. & CALCOTT, B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34(3)**: 772-773.

LARKIN, M.A., BLACKSHIELDS, G., BROWN, N.P., CHENNA, R., MCGETTIGAN, P.A., MCWILLIAM, H., VALENTIN, F., WALLACE, I.M., WILM, A., LOPEZ, R., THOMPSON, J.D., GIBSON, T.J. & HIGGINS, D.G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

LUCINDA, P.H.F. & REIS, R.E. 2005. Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. *Neotropical Ichthyology* **3**: 1-60.

MADDISON, W.P. & MADDISON, D.R. 2018. Mesquite: a modular system for evolutionary analysis. Version 3.51. Available at: http://www.mesquiteproject.org.

MARSHALL, B. 2011. *The fishes of Zimbabwe and their biology*. Smithiana Monographs. **3**: i-vi + 1-290.

MEREDITH, R.W., PIRES, M.N., REZNICK, D.N. & SPRINGER, M.S. 2010. Molecular phylogenetic relationships and the evolution of the placenta in *Poecilia* (*Micropoecilia*) (Poeciliidae: Cyprinodontiformes). *Molecular Phylogenetics and Evolution* **55(2)**: 631-639.

MEYER, A. & LYDEARD, C. 1993. The evolution of copulatory organs, internal fertilization, placentae and viviparity in killifishes (Cyprinodontiformes) inferred from a DNA phylogeny of the tyrosine kinase gene X-src. *Proceedings of the Royal Society of London* **254(1340)**: 153-162.

MEYER, A., MORRISSEY, J.M. & SCHARTL, M. 1994. Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. *Nature* **368(6471)**: 539-542.

MINDELL, D.P. & HONEYCUTT, M.R. 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Annual Review of Ecology and Systematics* **21**: 541-566.

MILLER, M.A., PFEIFFER, W. & SCHWARTZ, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE),* 14 Nov 2010, New Orleans, LA: 1-8.

MMOPELWA, G., MOSEPELE, K., MOSEPELE, B., MOLEELE, N. & NQWENYA, B. 2009. Environmental variability and the fishery dynamics of the Okavango Delta, Botswana: the case of subsistence fishing. *African Journal of Ecology* **47**: 119-127.

MOYLE, P.B. & MARCHETTI, M.P. 2006. Predicting invasion success: freshwater fishes in California as a model. *BioScience* **56(6)**: 515-524.

NICHOLS, J.T., GRISCOM, L., LANG, H., CHAPIN, J.P. & BOULENGER, G.A. 1917. Fresh-water fishes of the Congo Basin obtained by the American Museum Congo Expedition, 1909-1915. *Bulletin of the American Museum of Natural History*; 37: article 25.

NOTREDAME, C., HIGGINS, D.G. & HERINGA, J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**:

205-217. (http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee) Electronic version accessed 15 September 2018.

OUFIERO, C.E., MEREDITH, R.W., JUGO, K.N., TRAN, P., CHAPPELL, M.A., SPRINGER, M.S., REZNICK, D.N. & GARLAND, T.Jr. 2014. The evolution of the sexually selected sword in *Xiphophorus* does not compromise aerobic locomotor performance. *Evolution* **68(6)**: 1806-1823.

PAGE, R.D.M. 1996. Treeview: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**: 357-358.

PARENTI, L.R. 1981. A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History* **168**: 335-557.

PELCICE, F.M., POMPEU, P.S. & AGOSTINHO, A.A. 2015. Large reservoirs as ecological barriers to downstream movements of Neotropical migratory fish. *Fish and Fisheries* **16**: 697-715.

PERDICES, A., VASIL’EV, V. & VASIL’EVA, E. 2012. Molecular phylogeny and intraspecific structure of loaches (genera *Cobitis* and *Misgurnus*) from the Far East region of Russia and some conclusions on their systematics. *Ichthyological Research* **59**: 113-123.

PEREIRA, L.H., HANNER, R., FORESTI, F. & OLIVEIRA, C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics* **14(1)**: 20.

POHL, M., MILVERTZ, F.C., MEYER, A. & VENCES, M. 2015. Multigene phylogeny of cyprinodontiform fishes suggests continental radiations and a rogue taxon position of *Pantanodon*. *Vertebrate Zoology* **65**: 37-44.

POLLUX, B.J.A., MEREDITH, R.W., SPRINGER, M.S. & REZNICK, D.N. 2014. The evolution of the placenta drives a shift in sexual selection in livebearing fish. *Nature* **513(7517)**: 233.

RAMBAUT, A. 2012. FigTree version 1.4. Available from: http://tree.bio.ed.ac.uk/software/figtree/.

RAMBAUT, A., SUCHARD, M.A., XIE, D. & DRUMMOND, A.J. 2014. Tracer v1.6. Available from: http://beast.bio.ed.ac.uk/Tracer.

RATNASINGHAM, S. & HEBERT, P.D.N. 2007. BOLD: The barcode of life data system (www.barcodinglife.org). Available from: www.boldsystems.org.

REZNICK, D.N., RODD, F.H. & CARDENAS, M. 1996. Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. *The American Naturalist* **147**: 319-338.

REZNICK DN, FURNESS AI, MEREDITH RW, SPRINGER MS. 2017. The origin and biogeographic diversification of fishes in the family Poeciliidae. *PLoS ONE* **12(3)**: e0172546. https://doi.org/10.1371/journal.pone.0172546.

ROMAN, J. & DARLING, J.A. 2007. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution* **22(9)**: 454-464.

RONQUIST, F. & HUELSENBECK, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.

RStudio Team. 2018. RStudio: Integrated Development for R. *RStudio, Inc., Boston, Massachusetts.* Available at: URL http://www.rstudio.com/.

RUBINOFF, D. 2006. Utility of mitochondrial DNA barcodes in species conservation. *Conservation Biology* **20**: 1026-1033.

SCHLETTWEIN, C.H.G., SIMMONS, R.E., MACDONALD, A. & GROBLER, H.J.W. 1991. Flora, fauna and conservation of East Caprivi wetlands. *Madoqua* **17(2)**: 67-76.

SCHWARZ, G. 1978. Estimating the dimension of a model. *The Annals of Statistics* **6**: 461-464.

SKELTON, P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town.

STEENKAMP, M.K.J., ENGELBRECHT, G.D. & MULDER, P.F.S. 2001. Allozyme variation in a Johnston's topminnow, *Aplocheilichthys johnstoni*, population from the Zambezi River system. *Water SA* **27**: 53-55.

STEINKE, D., ZEMLAK, T.S., BOUTILLIER, J.A. & HEBERT, P.D.N. 2009. DNA barcoding of Pacific Canada's fishes. *Marine Biology* **156**: 2641-2647.

STOECK, M. & WANG, J. 2015. Direct submission. *National Center for Biotechnology Information*. Avalable from: https://www.ncbi.nlm.nih.gov/nuccore/KT175512.1/. Accessed: 15 September 2018.

SUCHARD, M.A., LEMEY, P., BAELE, G., AYRES, D.L., DRUMMOND, A.J. & RAMBAUT, A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4(1)**: vey016, https://doi.org/10.1093/ve/vey016.

SUZUKI-MATSUBARA, M., MURASE, Y. & MORIYAMA, A. n/d. Higashiyama Zoo DNA barcode project. Unpublished.

TOBLER, M., PALACIOS, M., CHAPMAN, L.J., MITROFANOV, I., BIERBACH, D., PLATH, M., ARIAS-RODRIGUEZ, L., GARCÍA DE LEÓN, F.J. & MATEOS, M. 2011. Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. *Evolution* **65**: 2213-2228.

VAN DER LAAN, R., ESCHMEYER, W. N. & FRICKE, R. 2014. Family-group names of Recent fishes. *Zootaxa Monograph* **3882**: 1-230. DOI 10.11646/zootaxa.3882.1.1.

VILA, I., MORALES, P., SCOTT, S., POULIN, E., VELIZ, D., HARROD, C. & MENDEZ, M.A. 2013. Phylogenetic and phylogeographic analysis of the genus *Orestias* (Teleostei: Cyprinodontiformes) in the southern Chilean Altiplano: the relevance of ancient and recent divergence processes in speciation. *Journal of Fish Biology* **82(3)**: 927-943.

WARD, R.D., ZEMLAK, T.S., INNES, B.H., LAST, P.R. & HEBERT, P.D. N. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society* **360**: 1847-1857.

WEBB, SA., GRAVES, J.A., MACIAS-GARCIA, C., MAGUURAN, A.E., FOIGHIL, D.O. & RITCHIE, M.G. 2004. Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular Phylogenetica and Evolution* **30(3)**: 527-544.

WILDEKAMP, R.H. 1995. A world of killies. Atlas of the oviparous cyprinodontiform fishes of the world.Volume II. The genera *Aplocheilichthys*, *Aplocheilus*, *Austrofundulus*, *Campellolebias*, *Crenichthys*, *Cualac*, *Cubanichthys*, *Cynolebias*, *Cynopoecilus* and *Cyprinodon*. *The American Kilifish Association, Inc.* 384 p.

XIA, X. 2017. DAMBE6: New tools for microbial genomics, phylogenetics and molecular evolution. *Journal of Heredity* **108**: 431-437.

XIA, X., XIE, Z., SALEMI, M., CHEN, L. & WANG, Y. 2003. An index substitution saturation and its application. *Molecular Phylogenetics and Evolution* **26**: 1-7.

XIA, X. & LEMEY, P. 2009. Assessing substitution saturation with DAMBE. Pp. 615-630 in Philippe Lemey, Marco Salemi and Anne-Mieke Vandamme, eds. *The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny*. 2nd edition Cambridge University Press.

ZHANG, J., KAPLI, P., PAVLIDIS, P. & STAMATAKIS, A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29(22): 2869-2876.

Table 2.1. Samples used in the study, with SAIAB catalogue and locality information listed in alphabetical order from genus names.

Full Name	System	River	Geocoordinates	Catalogue Number	Tube Number
Aplocheilichthys Bleeker, 1863					
spilauchen	Kwanza	Lucala	9.52°S 14.39°E	84696	ES07-F133
spilauchen	Kwanza	Kwanza	9.69°S 14.42°E	85367	ES07-D007
					ES07-D008
Lacustricola Myers, 1924					
macrurus	Kunene	Kunene	17.00°S 13.25°E	SAFW007-06	A 066
				SAFW008-06	A 067
				78465	A068
			17.41°S 14.22°E	SAFW052-06	A 341
				SAFW053-06	A 342
				78760	A344
					A345
		Olushandja Dam	17.43°S 14.64°E	200762	RB15-A148
					RB15-A152
					RB15-A153
					RB15-A154
					RB15-A156
myaposae	-	Nkanini	26.95°S 32.78°E	88639	PM09-A100
	-	Lake Mgobozeleni	27.54°S 32.66°E	88653	PM09-A142
	Kosi	Mpungwune	26.96°S 32.83°E	88645	PM09-A118
	Lake Sibaya	Mseleni	27.36°S 32.53°E	88658	PM09A164
					PM09A166
Micropanchax Myers, 1924					

Full Name	System	River	Geocoordinates	Catalogue Number	Tube Number
johnstoni	Okavango	Cuito	15.16°S 19.2°E	186866	Rb12-A164
			15.17°S 19.19°E	186834	RB12-A305
		Kwando	17.57°S 23.26°E	101120	RB13-A350
		Thamalakane	20.01°S 23.43°E	97403	RC10-C174
			19.94°S 23.49°E	98251	RC10-C167
				187053	RB12-A430
		Okavango	18.76°S 22.36°E	75438	D074
	Shire	Ruo	16.07°S 35.67°E	119062	TM11-116
		Shire	15.06°S 35.22°E	118776	TM11-264
	Zambezi	Zambezi	16.19°S 33.62°E	97044	RC10-C046
			16.13°S 33.56°E	97183	RC10-C011
			15.75°S 31.61°E	81254	204
			15.60°S 32.68°E	81389	398
	Cuando-Linyanti-Chobe	Kwando	17.88°S 23.32°E	200808	RB15-A305
					RB15-A324
					RB15-A420
			17.79°S 23.34°E	200812	RB15-A309
	Okavango	Okavango	18.12°S 21.58°E	85360	RB08ZaNam361
			18.22°S 21.75°E	200764	RB15-A127
					RB15-A128
			17.97°S 20.47°E	200766	RB15-A132
					RB15-A137
				200079	RB15-A143
	Zambezi	Zambezi	17.56°S 24.46°E	200802	RB15-A248
					RB15-A300
				200813	RB15-A159
	Limpopo	Mokolo	24.11°S 27.8°E	SAFW096-08	IRB313
	Zambezi	Chisolo	12.19°S 25.21°E	97970	JWH10-A007

Full Name	**System**	**River**	**Geocoordinates**	**Catalogue Number**	**Tube Number**
		Kafue	15.76°S 28.09°E	83820	RB08ZaNam051
					RB08ZaNam076
					RB08ZaNam080
					RB08ZaNam089
		Luampa	14.96°S 24.34°E	72276	294
johnstoni	Congo	Lulimala	12.29°S 30.62°E	99047	MI11-196
	Luapula	Mulembo	12.54°S 30.39°E	76546	514
hutereaui	Okavango	Boa-fé flood plain	17.47°S 23.08°E	101089	RB13-A400
		Boa-fé Lagoon	17.33°S 20.13°E	101066	RB13-A275
		Cuito	16.62°S 19.05°E	101178	RB13-A430
	Okavango	Okavango Delta	19.2°S 23.46°E	68657	176
	Cuando-Linyanti-Chobe	Kwando	17.88°S 23.32°E	200091	RB15-A225
					RB15-A290
					RB15-A335
				200810	RB15-A213
	Okavango	Okavango	18.22°S 21.75°E	200765	RB15-A119
					RB15-A122
					RB15-A123
				200085	RB15-A358
	Zambezi	Zambezi	-	63528	U9517
			17.56°S 24.46°E	200803	RB15-A345
					RB15-A391
				Nam15-13	RB15-A249
					RB15-A315
				200814	RB15-A188
	Zambezi	Zambezi	16.15°S 23.29°E	71199	198

Full Name	**System**	**River**	**Geocoordinates**	**Catalogue Number**	**Tube Number**
		Musangezhi	12.35°S 25.11°E	96124	RB10-B149
hutereaui	Congo	Lulimala	12.29°S 30.62°E	99046	ACEP09-1502
katangae	Okavango	Boa-fé Lagoon	17.33°S 20.13°E	101067	RB13-A281
		Chingueia tributary	13.69°S 17.06°E	186731	RB12-A105
		Cuito	17.51°S 20.07°E	101097	RB13-A166
			16.62°S 19.05°E	101179	RB13-A402
	Okavango	Thamalakane	19.94°S 23.49°E	98252	RC10-C166
	Cuando-Linyanti-Chobe	Kwando	17.88°S 23.32°E	200809	RB15-A304
					RB15-A411
	Okavango	Okavango	18.21°S 21.75°E	200087	RB15-A114
				200089	RB15-A321
					RB15-A415
	Zambezi	Zambezi	17.56°S 24.46°E	200804	RB15-A179
					RB15-A187
					RB15-A291
		Nkanini	26.95°S 32.78°E	88640	PM09A105
	Lake Sibaya	Mseleni	27.36°S 32.53°E	88657	PM09A162
	Zambezi	Musanda	12.04°S 25.42°E	97994	JWH10-A138
		Lwalaba	12.36°S 25.36°E	140103	RB11-D002
		Shinengeni stream	12.36°S 25.33°E	140155	RB11-D023
		Zambezi	15.57°S 23.28°E	200059	DT10 124
					DT10 176
katangae	Congo	Lake Carumbo	7.79°S 19.95°E	118846	ES11-BT086

Full Name	**System**	**River**	**Geocoordinates**	**Catalogue Number**	**Tube Number**
sp.	Kwanza	Kwanza	9.73°S 15.86°E	85468	ES07-D080
					ES07-D089
	Okavango	Kwando	17.53°S 23.19°E	101072	RB13-A467
					RB13-A482
	St. Lucia	Hluhluwe	28.15°S 32.28°E	190720	RB13-B251
					RB13-B283

Table 2.2. Primer names and sequences used in the phylogenetic study.

Primer Name	**Primer Sequence**	**Reference**
COI		
LCO1490	5´ GGT CAA CAA ATC ATA AAG ATA TTG G3´	Folmer *et al.* 1994
HCO2198	5´ TAA ACT TCA GGG TGA CCA AAA AAT CA 3´	Folmer *et al.* 1994
28S		
28S-2F	5´ CTT GAA ACA CGG ACC AAG GA 3´	Unpublished data
28S-2R	5´ GTA TAG GCA CGA CGC TTC AG 3´	Unpublished data
16S		
16S-AR	5´ CGC CTG TTT ATC AAA AAC AT 3´	Pohl *et al.* 2015
16S-BR	5´ CCG GTC TGA ACT CAG ATC ACG T 3´	Pohl *et al.* 2015
cyt*b*		
GluDGL	5´ TGA CTT GAA RAA CCA YCG TTG 3´	Perdices *et al.* 2012
H16460	5´ CGA YCT TCG GAT TAC AAG ACC G 3´	Perdices *et al.* 2012

Table 2.3. Additional sequences downloaded from GenBank to seek a root for the in-group.

Species	**Accession Number**					**Sources**
	COI		**16S**	**28S**	**TyrK**	
Procatopodinae						
Micropanchax hutereaui	AY356594					Web *et al.* 2004
Micropanchax kingii	KJ844737		KJ844801			Pohl *et al.* 2015
Lamprichthys tanganicanus			KJ844799			Pohl *et al.* 2015
	LC1537741					Suzuki-Matsubara *et al.* n/d
	KJ8447351					Pohl *et al.* 2015
Hypsopanchax platysternus	KT1930111					Decru *et al.* 2016
	KT1930101					Decru *et al.* 2016
	KT1930091					Decru *et al.* 2016
	KT1930081					Decru *et al.* 2016
Plataplochilus miltotaenia	KJ844757		KJ844821			Pohl *et al.* 2015
Poropanchax luxophthalmus	KJ844759		KJ844823			Pohl *et al.* 2015
Poropanchax normani	KJ844760					Pohl *et al.* 2015
		KJ696787			KJ697575	Pollux *et al.* 2014
Procatopus nototaenia	KJ844761		KJ844824			Pohl *et al.* 2015
					KJ697663	Pollux *et al.* 2014
Procatopus aberrans	LC1548071					Suzuki-Matsubara *et al.* n/d
Aplocheilichthyinae						
Aplocheilichthys spilauchen		KJ6967881			KJ697576	Pollux *et al.* 2014
	LC1548161					Suzuki-Matsubara *et al.* n/d
					U02344	Meyer & Lydeard 1993
Valenciidae						
Valencia hispanica	KF767523					Freyhof *et al.* 2014
	KF767525					Freyhof *et al.* 2014
	KF767528					Freyhof *et al.* 2014
					KJ697669	Pollux *et al.* 2014

Poeciliidae						
Gambusia holbrooki		KJ696803				Pollux *et al.* 2014
					KJ697591	Pollux *et al.* 2014
Gambusia luma		KJ696804				Pollux *et al.* 2014
					KJ697592	Pollux *et al.* 2014
Poecilia gillii	KP7005091					Bagley *et al.* 2015
	KP700466					Bagley *et al.* 2015
Poecilia mexicana		KP699802			KP700732	Bagley *et al.* 2015
		KT175512				Stoeck & Wang 2015
Poecilia reticulata	KP7005211	KP700413				Bagley *et al.* 2015
					GU179162	Meredith *et al.* 2010
				AH011837		Chen *et al.* 2003
			KX816038			Chen *et al.* 2016
	JN9891411					Pereira *et al.* 2013
Poecilia sphenops	KP700445	KP699970				Bagley *et al.* 2015
		KJ696840			KJ697638	Pollux *et al.* 2014
Tomeurus gracilis					KJ697668	Pollux *et al.* 2014
Xiphophorus maculatus		EF017551				Hrbek *et al.* 2007
					U06609	Meyer *et al.* 1994
Xiphophorus hellerii		EF017548				Hrbek *et al.* 2007
					KJ525899	Oufiero *et al.* 2014
Xenodexia ctenolepis		EF017557				Hrbek *et al.* 2007
					KJ697670	Pollux *et al.* 2014
Fluviphylacidae						
Fluviphylax pygmaeus		EF017511				Hrbek *et al.* 2007
					U02350	Meyer & Lydeard 1993
Fluviphylax simplex		KJ696800			KJ697587	Pollux *et al.* 2014
Anablepidae						
Anableps dowei					KJ697573	Pollux *et al.* 2014
Anableps anableps	LC154806					Suzuki-Matsubara *et al.* n/d

	KJ696784	KJ697572	Pollux *et al*. 2014
	EF017508		Hrbek *et al*. 2007
Jenynsia lineata	KJ696808	KJ697598	Pollux *et al*. 2014
	EF017509		Hrbek *et al*. 2007
Jenynsia multidentata	KJ696809	KJ697599	Pollux *et al*. 2014
Fundulidae			
Lucania goodei	KJ696822	KJ697614	Pollux *et al*. 2014
Lucania parva	KJ696823	KJ697615	Pollux *et al*. 2014
Fundulus cingulatus	KJ696801	KJ697589	Pollux *et al*. 2014
Fundulus lineolatus	KJ696802	KJ697590	Pollux *et al*. 2014
Cyprinodontidae			
Orestias agassizii	JX092150		Vila *et al*. 2013
		KJ697620	Pollux *et al*. 2014

Table 2.4. Scores from the saturation test applied according to Xia *et al.* (2003) in DAMBE. The table showed the scores for each molecular marker used in the analysis. The Iss score being lower than the Iss.c score with a significant P-value indicates low saturation in sequences and that sequences are useful in phylogenetic studies.

COI Codon position 1						**CYTB Codon position 1**						**28S**					
NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P
4	0,121	0,777	26,589	219	0	4	0,175	0,785	27,42	358	0	4	0,192	0,809	30,107	742	0
8	0,114	0,733	24,843	219	0	8	0,161	0,738	25,903	358	0	8	0,21	0,772	20,02	742	0
16	0,111	0,651	22,221	219	0	16	0,157	0,693	24,566	358	0	16	0,242	0,753	14,215	742	0
32	0,115	0,686	23,424	219	0	32	0,16	0,687	24,155	358	0	32	0,274	0,727	10,211	742	0

COI Codon position 2						**CYTB Codon position 2**						**16S**					
NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P
4	0,084	0,777	35,08	220	0	4	0,064	0,785	55,713	358	0	4	0,227	0,803	25,229	628	0
8	0,093	0,733	28,658	220	0	8	0,057	0,738	53,39	358	0	8	0,239	0,763	17,766	628	0
16	0,08	0,651	27,983	220	0	16	0,061	0,693	48,835	358	0	16	0,256	0,741	13,228	628	0
32	0,81	0,686	30,124	220	0	32	0,066	0,687	48,062	358	0	32	0,275	0,715	9,996	628	0

COI Codon position 3						**CYTB Codon position 3**						**Tyrosine Kinase**					
NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P	NumOTU	Iss	Iss.c	T	DF	P
4	0,586	0,777	5,636	219	0	4	0,664	0,785	4,666	357	0	4	0,324	0,8	17,207	574	0
8	0,569	0,733	5,326	219	0	8	0,667	0,738	2,867	357	0,0044	8	0,327	0,759	12,198	574	0
16	0,578	0,651	2,522	219	0,0124	16	0,674	0,693	0,806	357	0,4207	16	0,351	0,733	8,596	574	0
32	0,577	0,686	4,003	219	0,0001	32	0,675	0,687	0,538	357	0,5912	32	0,392	0,71	5,957	574	0

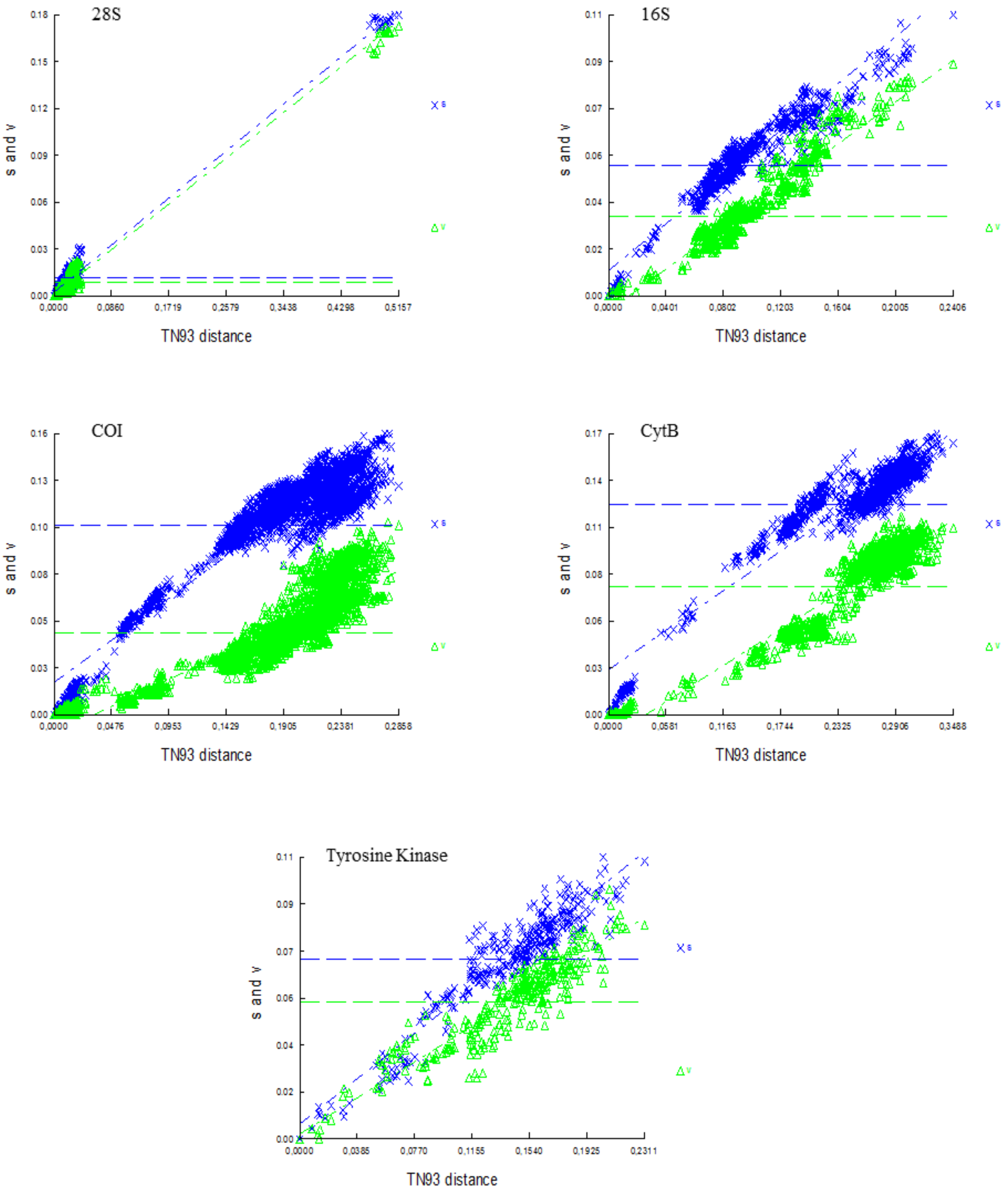


Fig. 2.1. Saturation plots for the five genetic markers chosen. The TN93 model of best-fit was used and the mean and lines of best-fit were shown.

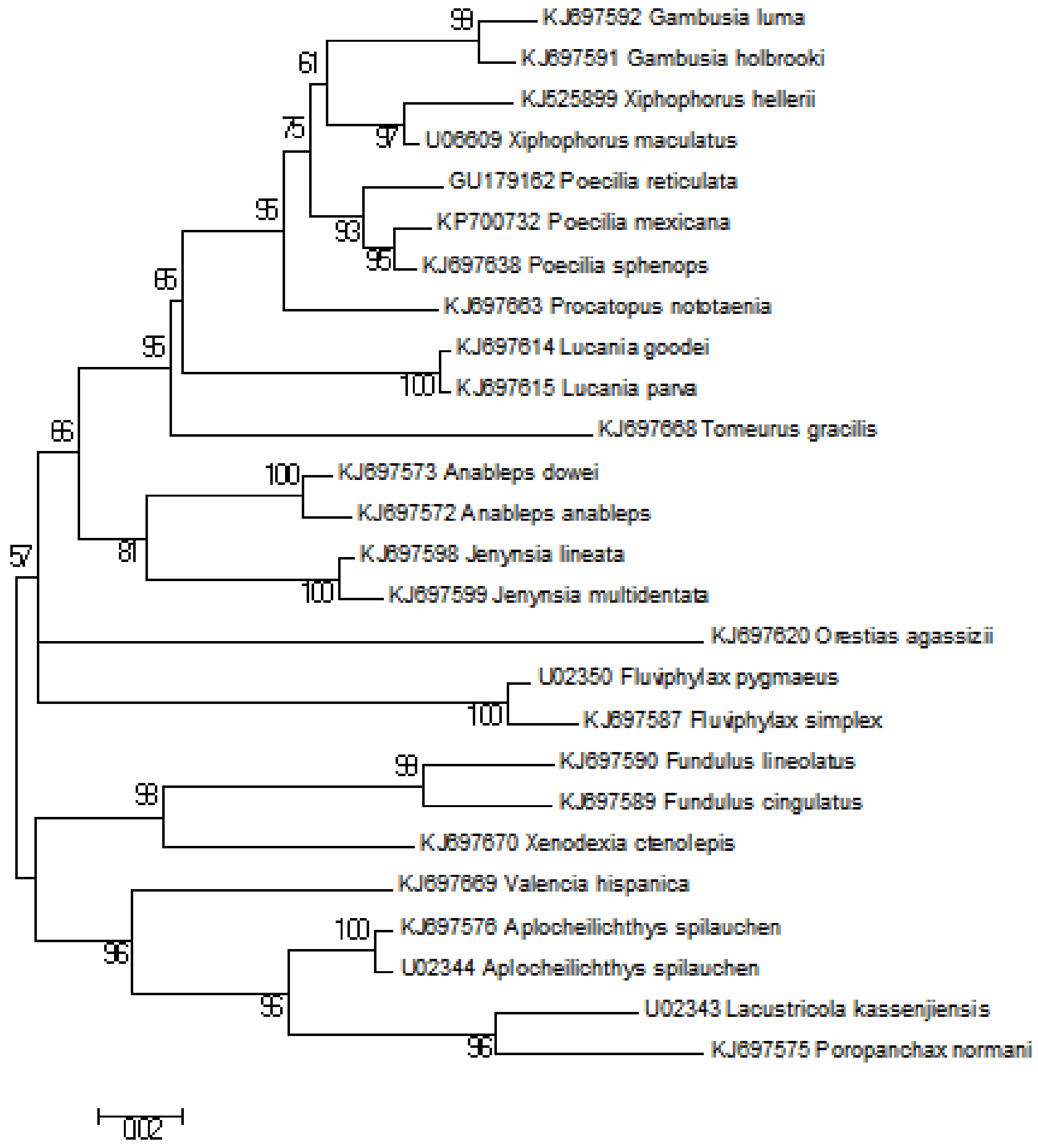


Fig. 2.2. Maximum likelihood tree (bootstrap repetitions = 1000) of possible out-groups used in the study with the Tyrosine Kinase genetic marker.

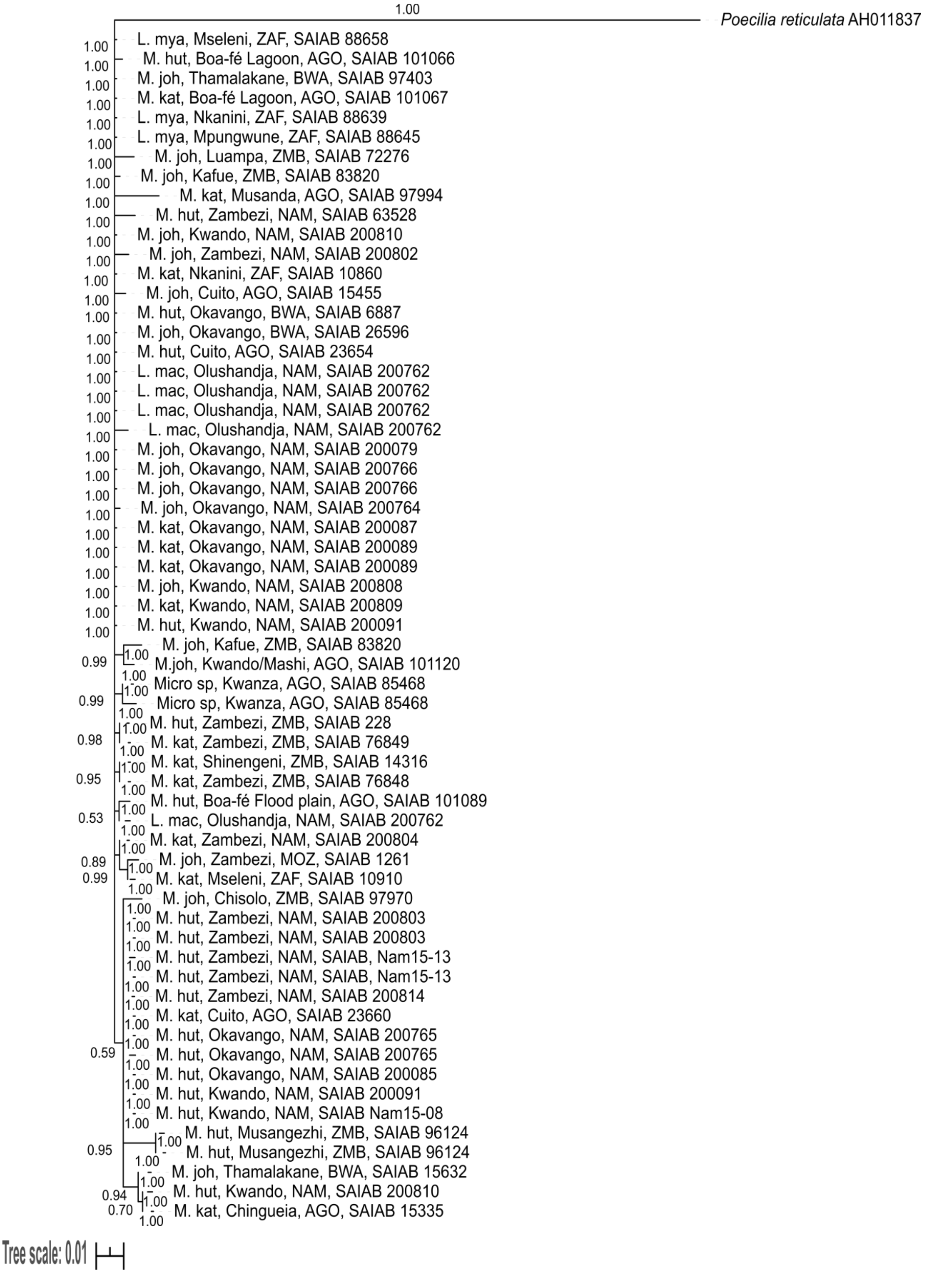


Fig. 2.3. Bayesian Inference phylogram for 28S built using a HKY+I+G model (out-group = *Poecilia*). Bootstrap values were calculated using a heuristic tree search using 1000 replicates. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

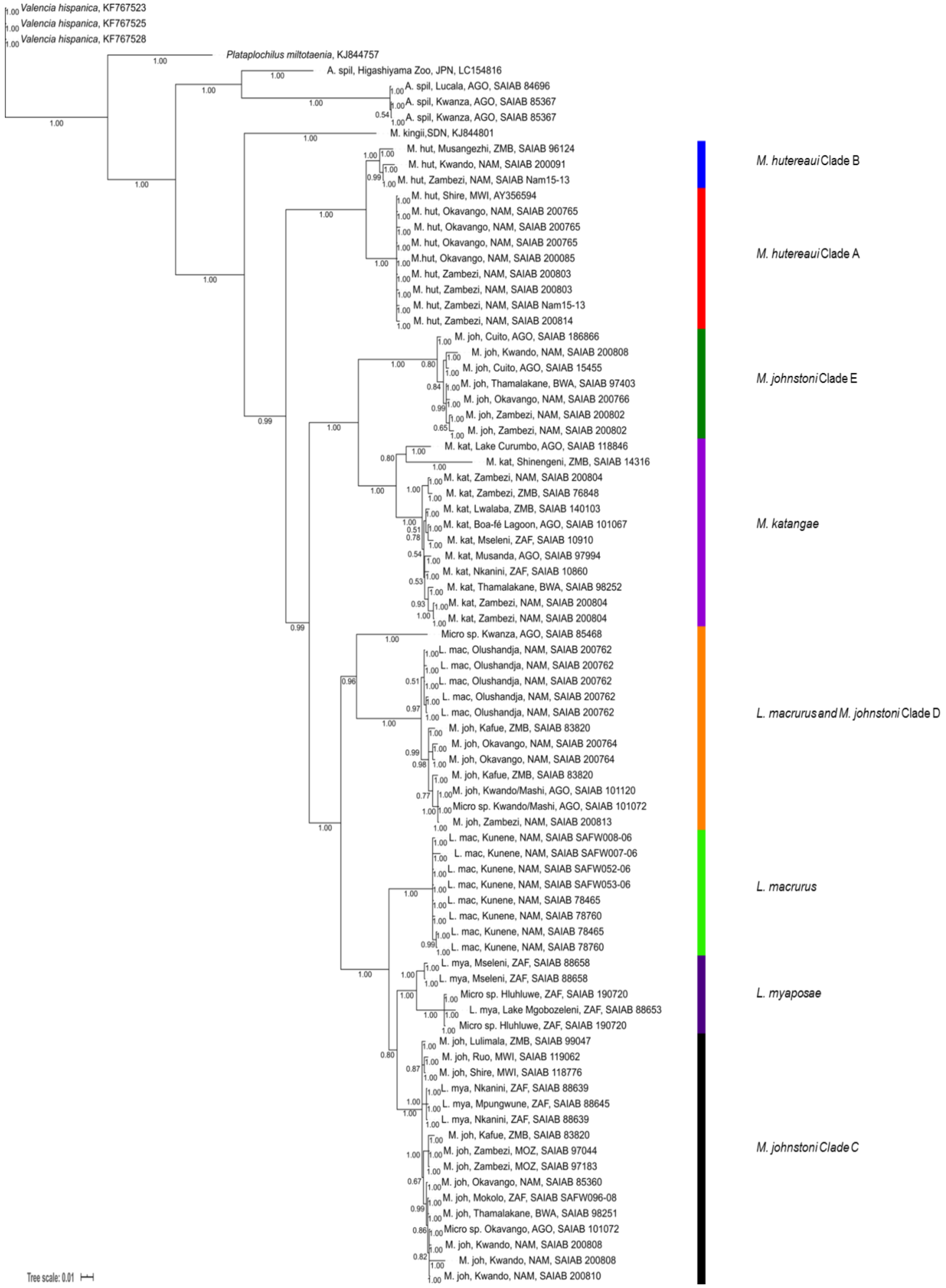


Fig. 2.4. Bayesian Inference phylogram for Total Molecular Data (28S, COI, Cyt*b* 16S and TyrK) partitioned by gene with COI and partitioned by base position; (out-group = *Valencia hispanica*). Posterior branch probabilities were calculated using a heuristic tree search using 1000 replicates. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

Fig. 2.5. Maximum likelihood phylogram for the Total Molecular Data (28S, COI, Cyt*b*, 16S and TyrK) partitioned by gene with COI and Cyt*b* partitioned by base position. (Out-group = *Valencia hispanica*). Posterior branch probabilities were calculated using a 1000 bootstrap replicates. Species names of the in-group species are shortened and all names are followed by GenBank accession and SAIAB catalogue numbers.

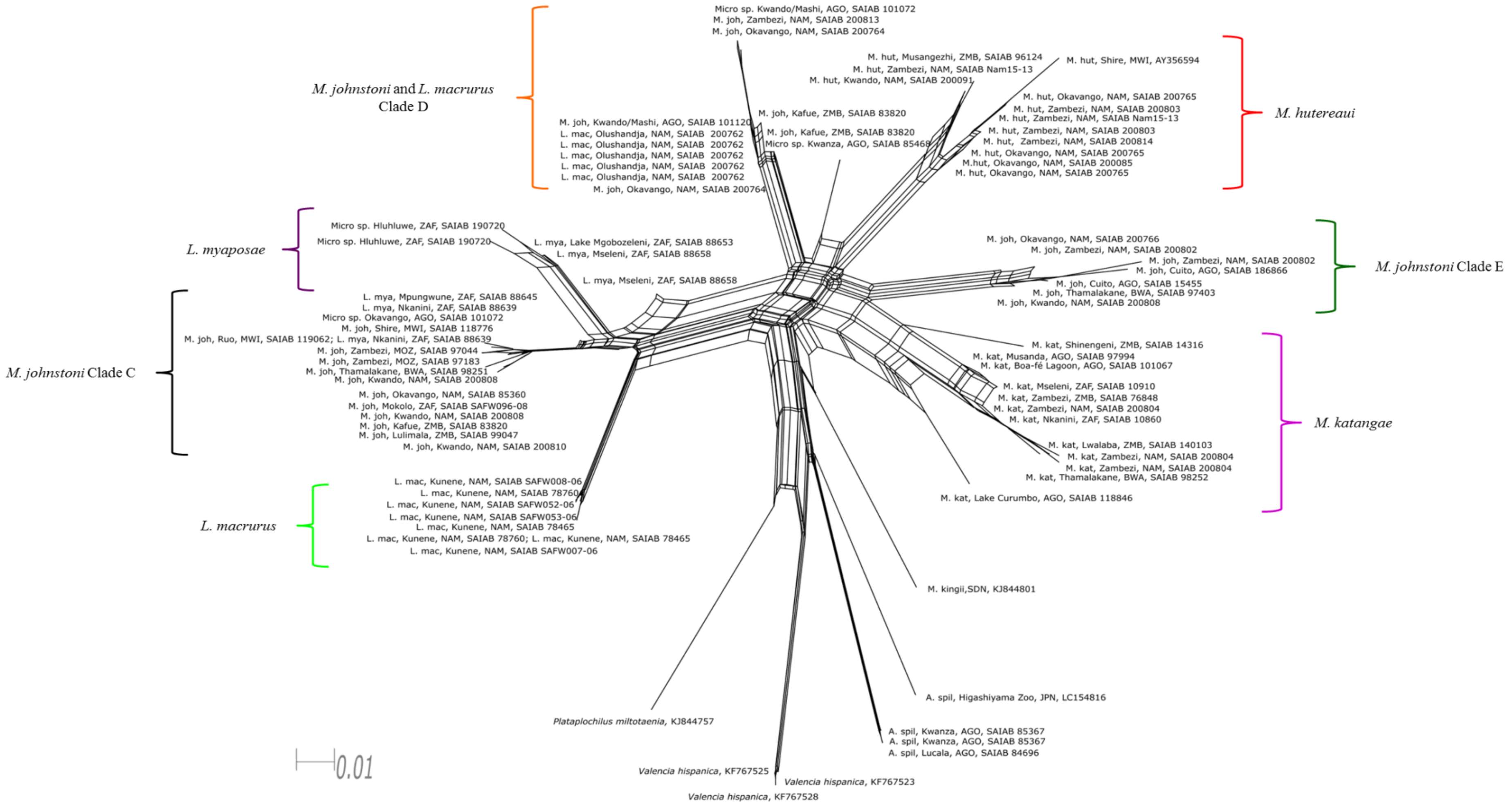


Fig. 2.6. Neighbor-net diagram representing the total molecular data (COI, 16S, 28S and Tyrosine Kinase) with uncorrected P-distances. Species names are followed with GenBank accession and SAIAB catalogue numbers.

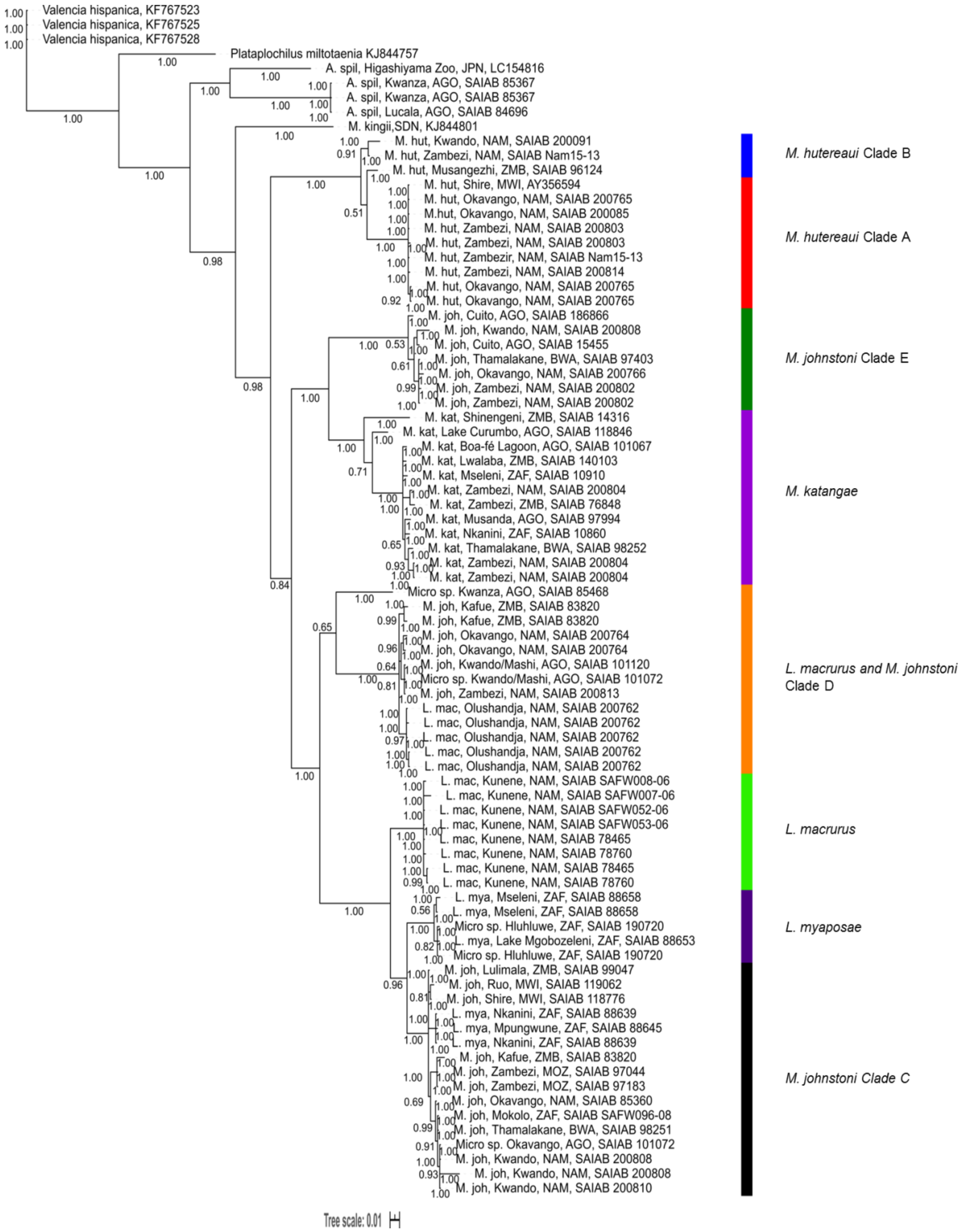


Fig. 2.7. Bayesian Inference phylogram for COI built using a HKY+I+G model and partitioned by base position (out-group = *Valencia hispanica*). Posterior branch probabilities were calculated using a heuristic tree search using 1000 replicates. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

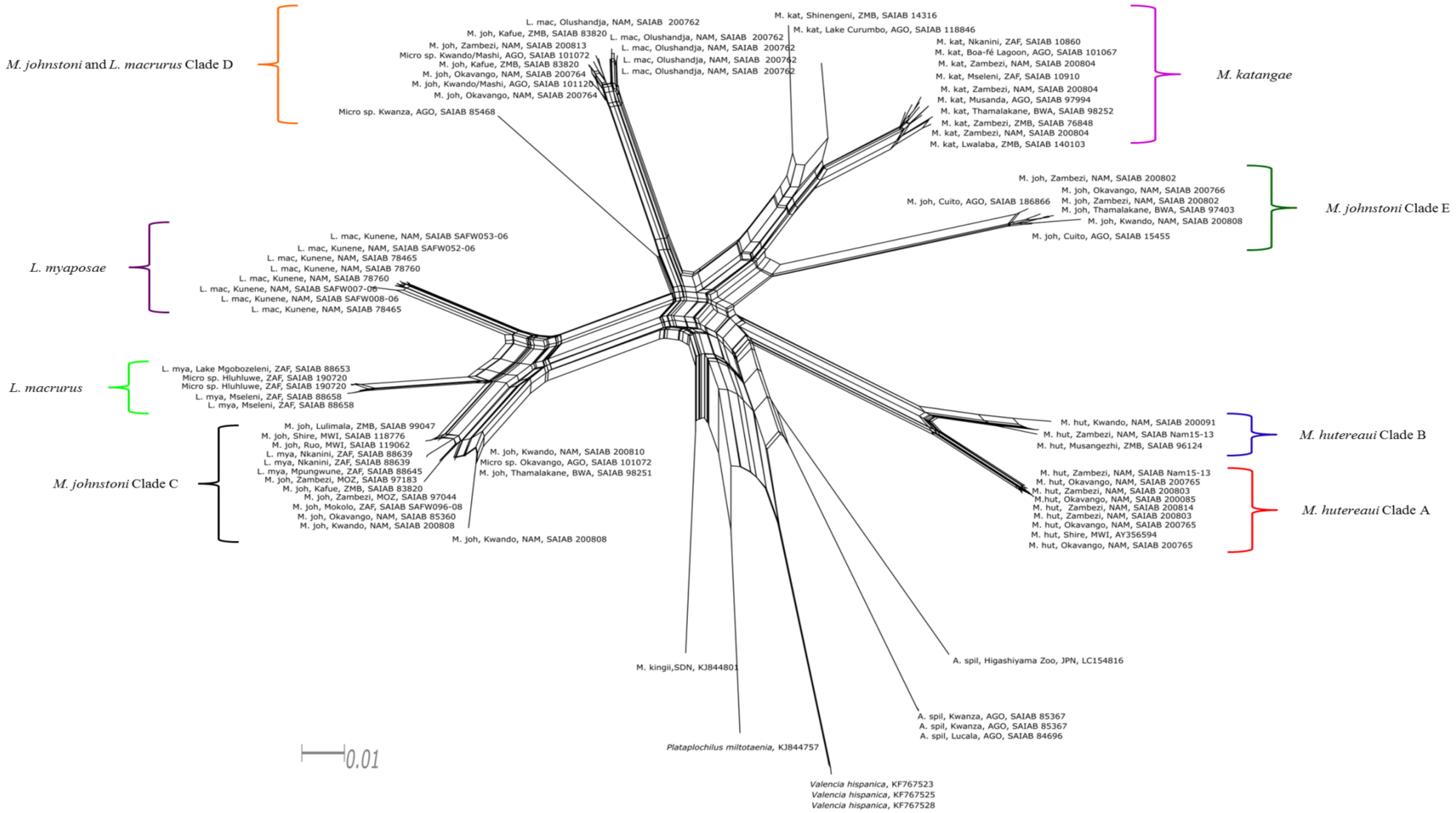


Fig. 2.8. Neighbor-net diagram representing the analysis of COI with uncorrected P-distances. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

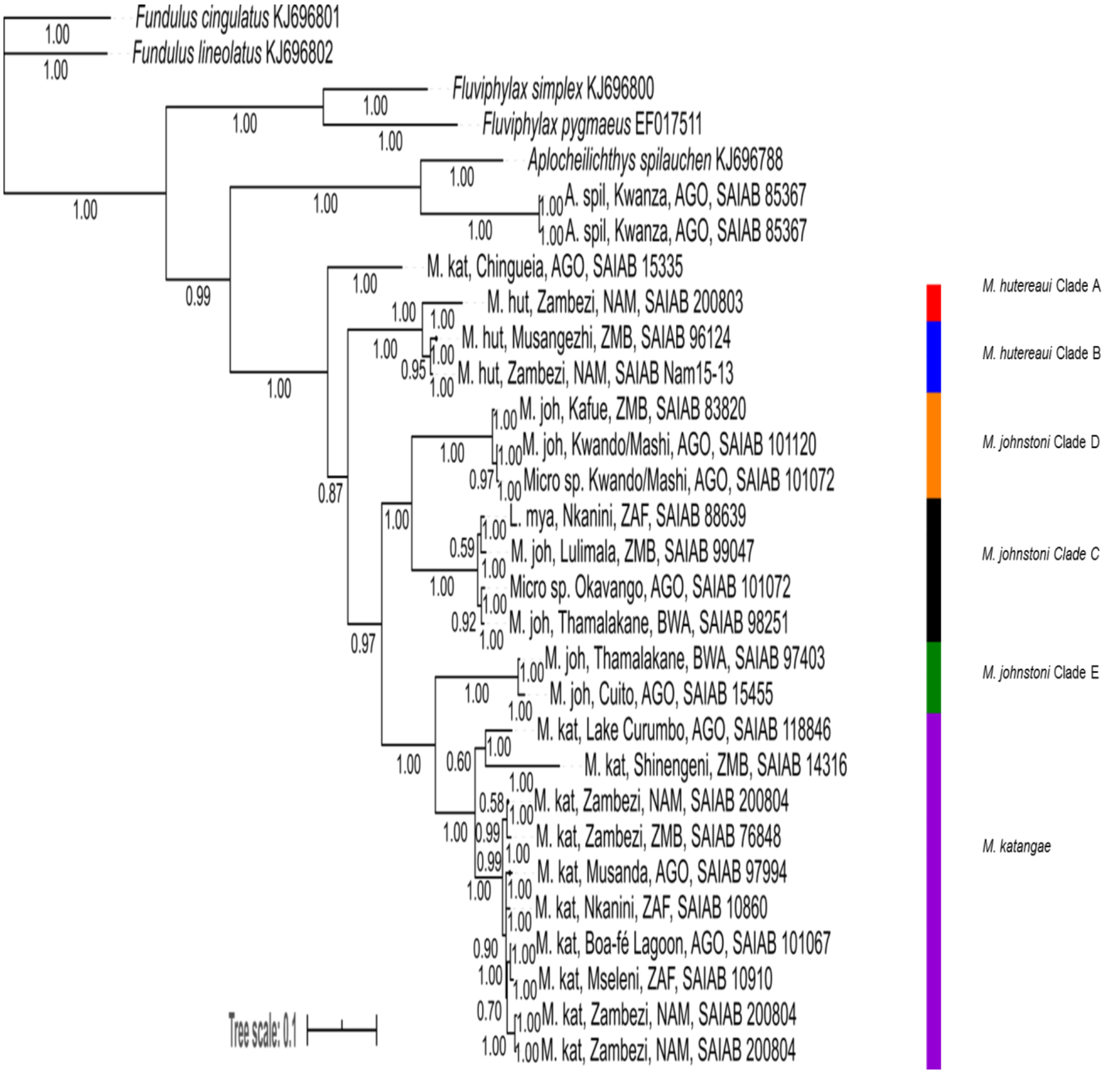


Fig. 2.9. Bayesian Inference phylogram for cyt*b* built using a GTR+I+G model and partitioned by base position (out-group = *Fundulus cingulatus*). Posterior branch probabilities were calculated using a heuristic tree search using 1000 replicates. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

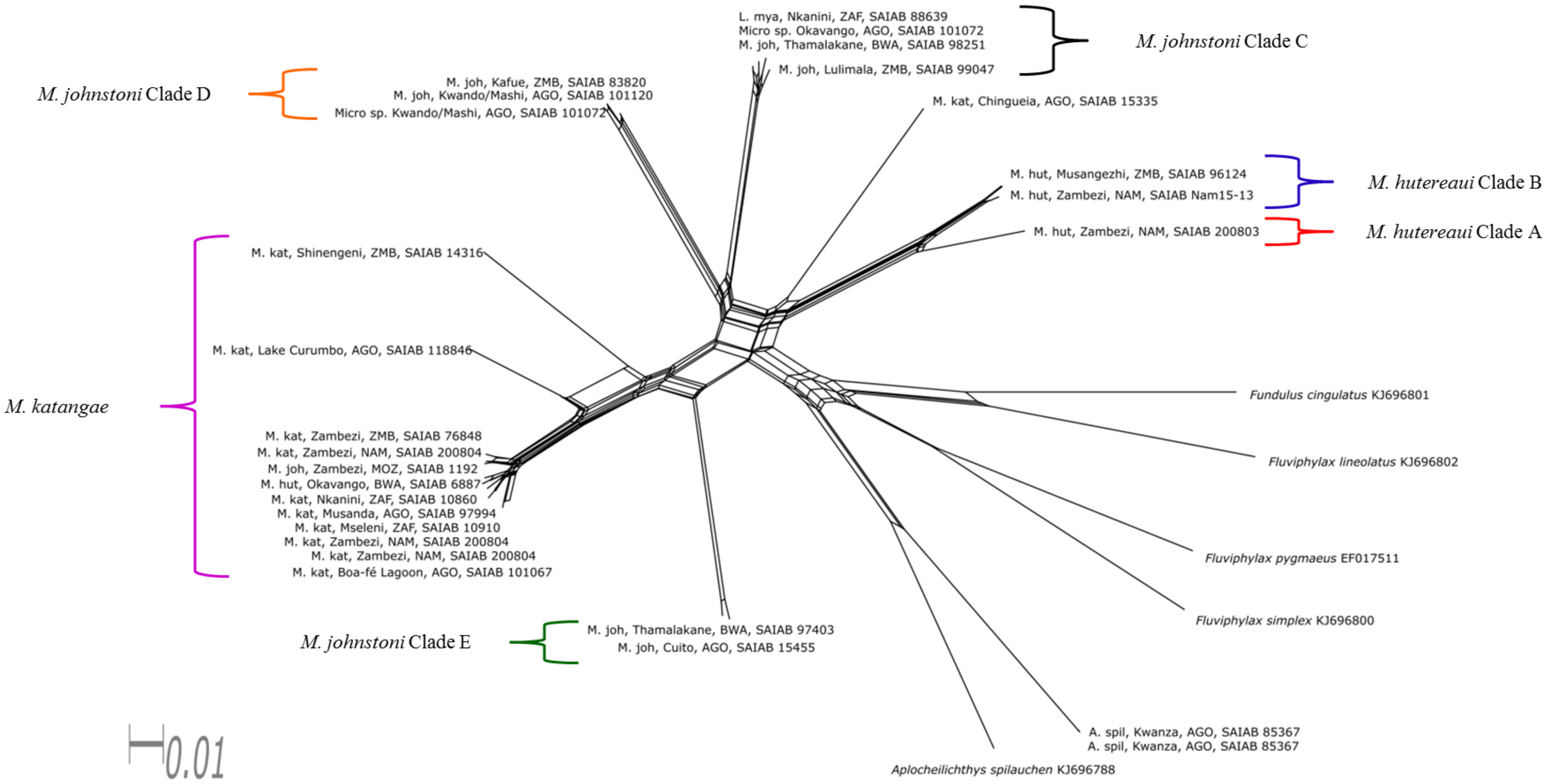


Fig. 2.10. Neighbor-net diagram representing the analysis of cyt*b* with uncorrected P-distances. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

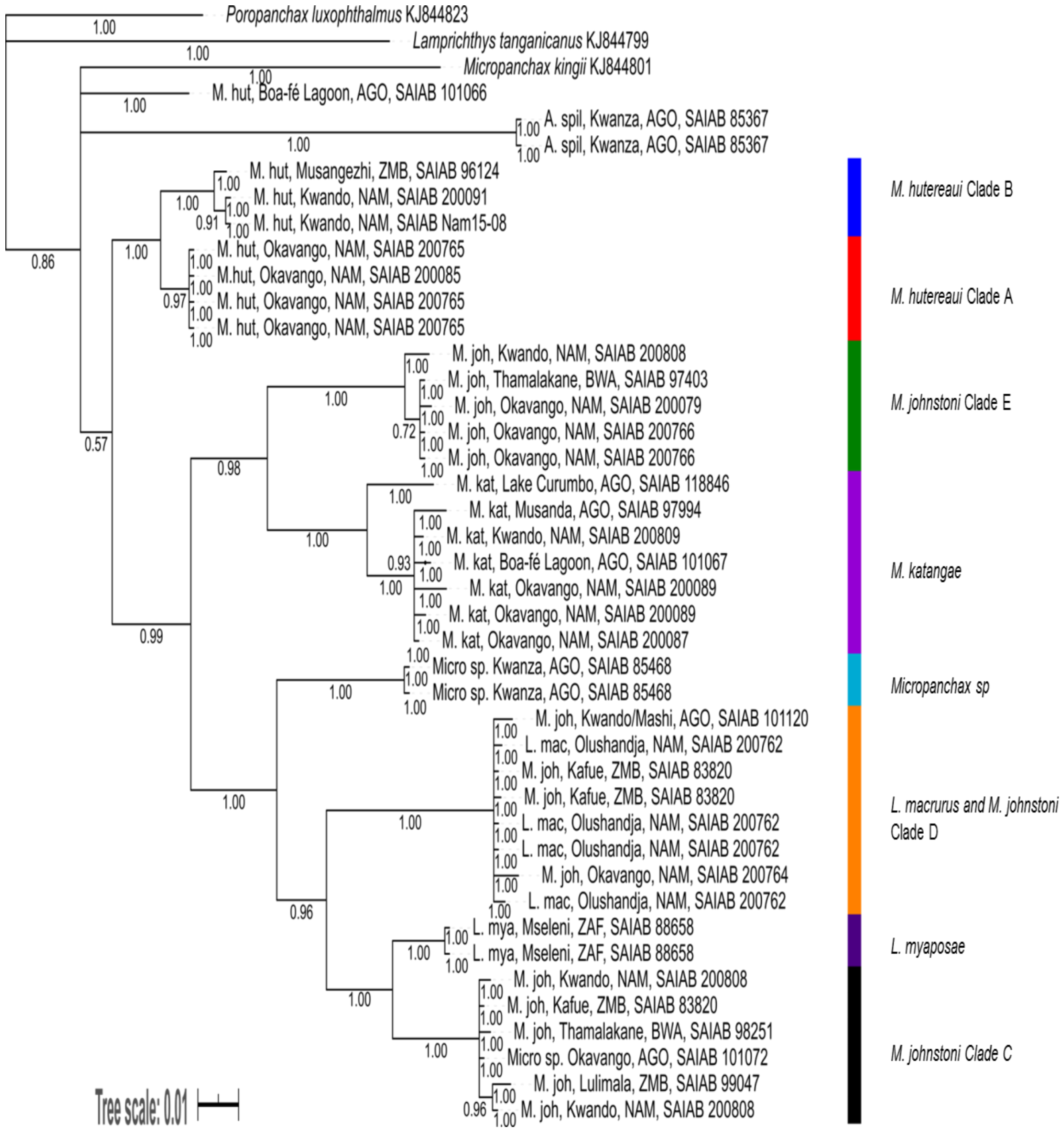


Fig. 2.11. Bayesian Inference phylogram for 16S built using SYM+I+G model (out-group = *Poropanchax luxophthalmus*). Posterior branch probabilities were calculated using a heuristic tree search using 1000 replicates. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

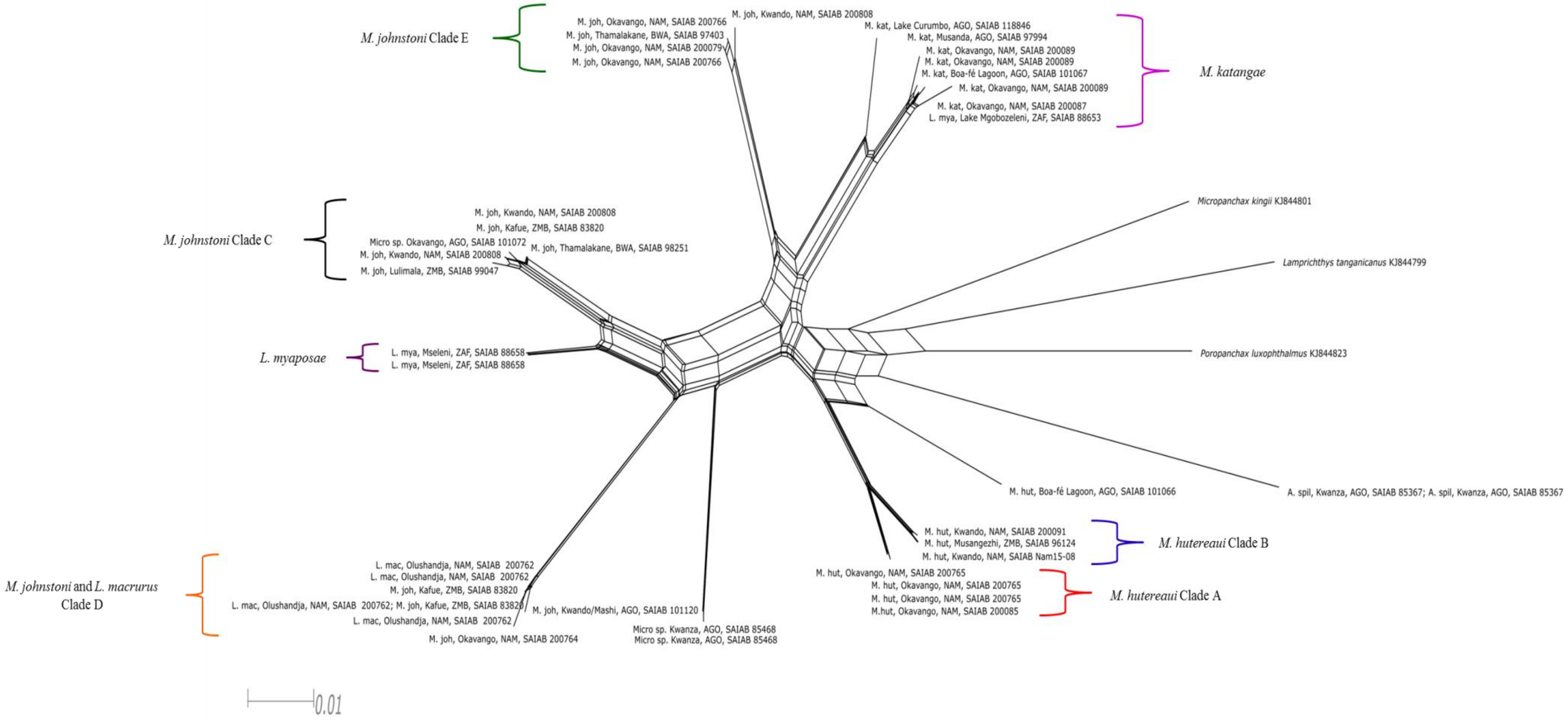


Fig. 2.12. Neighbor-net diagram analysis of 16S with uncorrected p-distances. Species' names are followed by GenBank accession and SAIAB catalogue numbers.

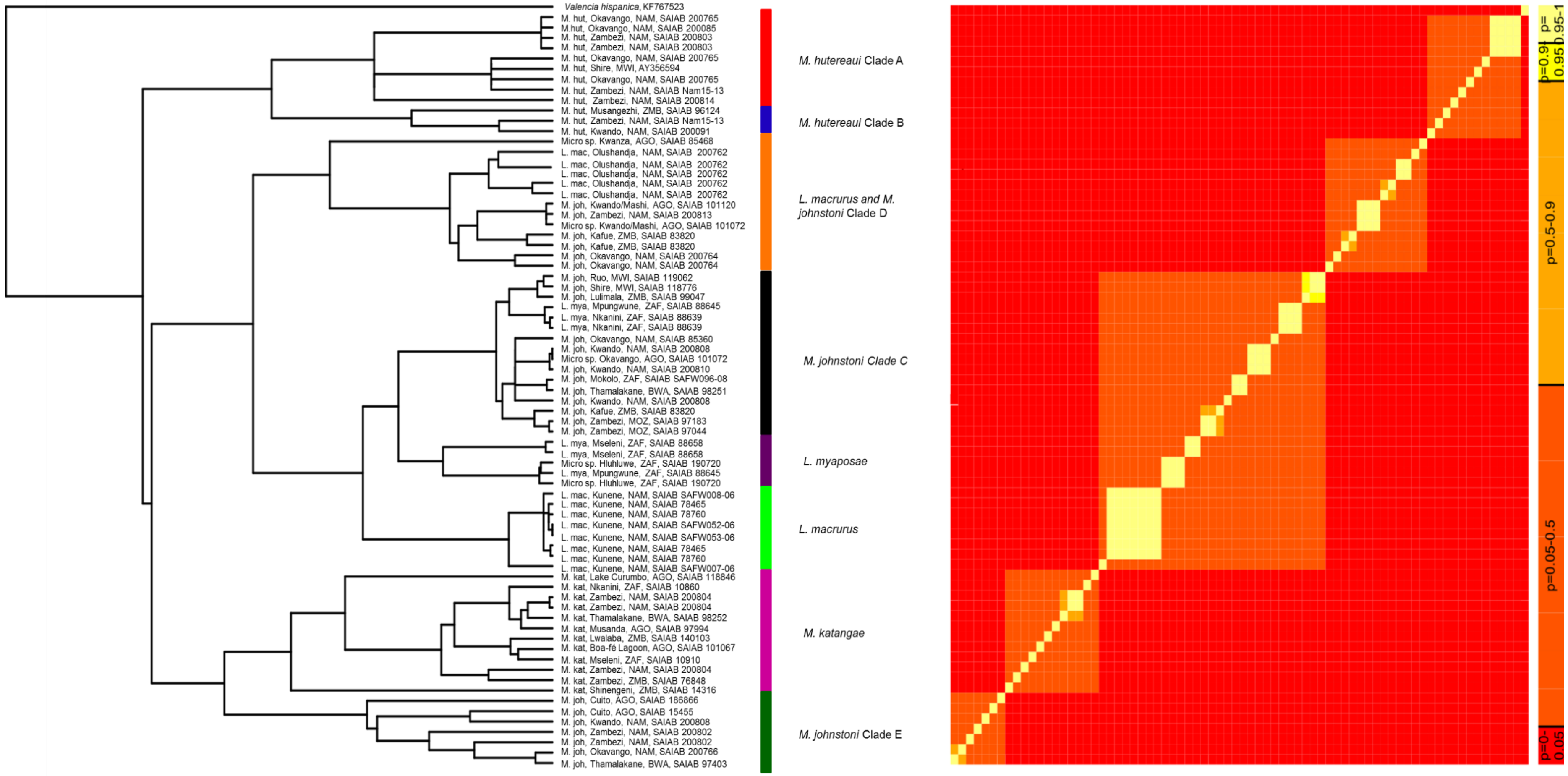


Fig. 2.13. Bayesian Generalized Mixed Yule-Coalescent ultrametric phylogram showing the results of the species delimitation with different species clades separated denoted by the red areas (P value 0 – 0.5) (Out-group = *Valencia hispanica*). Out-groups were excluded from the analysis as not to skew the results of analysis.

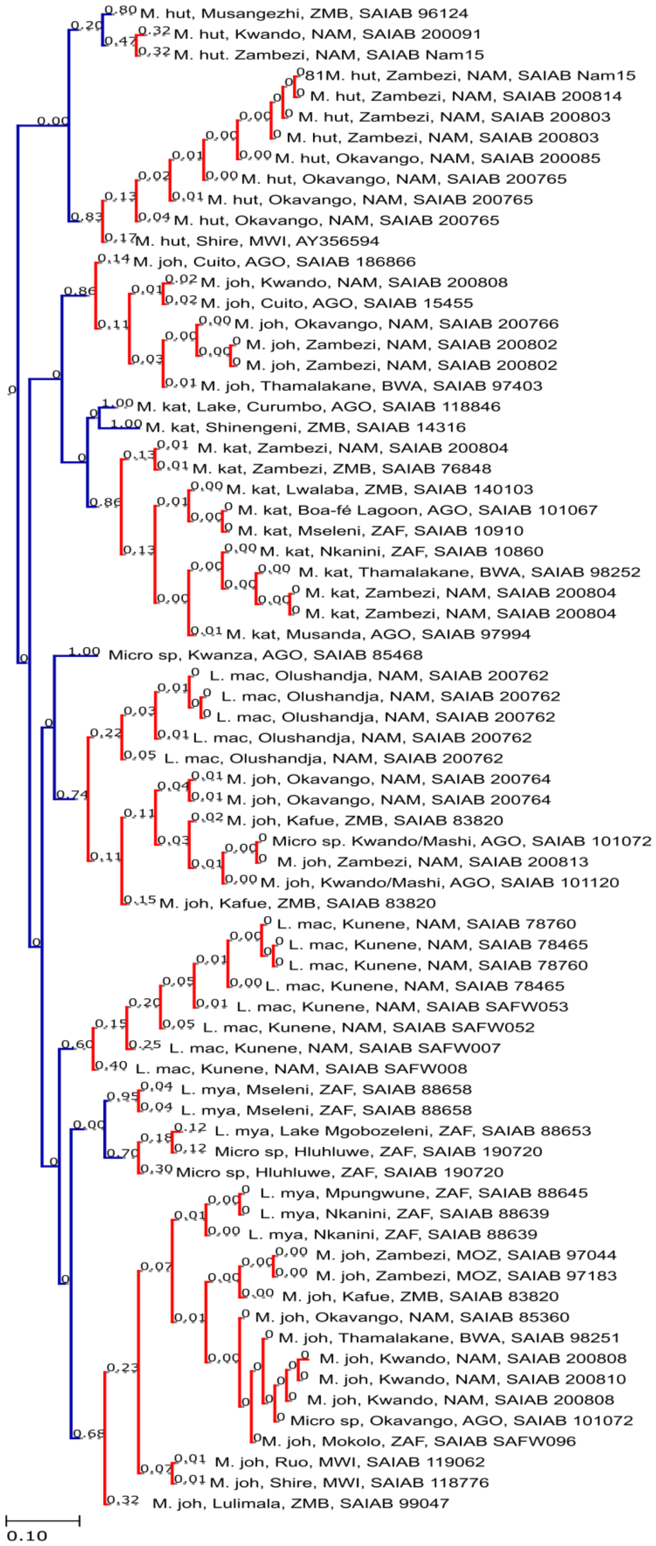


Fig. 2.14. Bayesian Poisson Tree Processes phylogram analysis to delimit species within the in-group. Tree shows Bayesian support values to delimit species (Out-group = *Valencia hispanica*). Out-group species were excluded from the analysis as not to skew the results of the analysis.

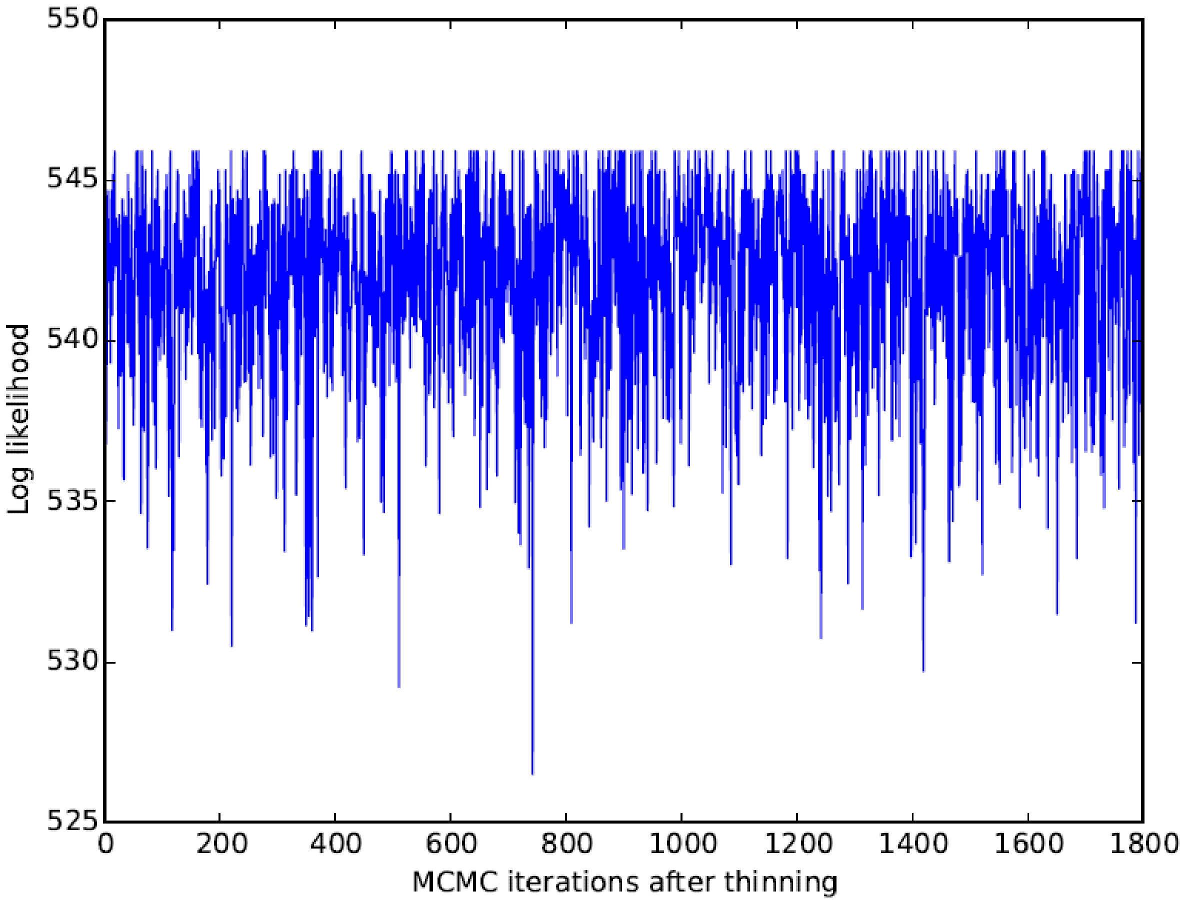


Fig. 2.15. MCMC iterations of the bPTP analysis showing that the tree reached convergence.

3

Morphometric analysis of the southern African topminnows (Procatopodidae: Procatopodinae), focused on three species complexes.

3.1. Introduction

3.1.1. Southern African Procatopodidae

The topminnows of the Procatopodidae Fowler, 1916 include several genera belonging to two subfamilies, Aplocheilichthyinae and Procatopodinae (Skelton 2001, 2002; Lucinda 2003; Marshall 2011; Braganca *et al.* 2018) and can be found in several river systems in southern Africa (Skelton 2001; Marshall 2011). The southern African species all lay eggs on vegetation in the marginal areas of rivers; the eggs are fertilised externally and do not survive desiccation (Skelton 2001). Topminnows acquired their name from their habitat and feeding behaviour, living in the shallow marginal areas of rivers and lakes, feeding from the water's surface, primarily eating small invertebrates. This makes topminnows important in controlling the spread of mosquito larvae and an important dietary source for predatory fish (Skelton 2001; Steenkamp *et al.* 2001).

The Aplocheilichthyinae consists of one genus found in southern Africa, *Aplocheilichthys* Bleeker 1863. The Procatopodinae consist of substantially more genera, with the important genera in southern Africa being *Lacustricola* Myers 1924 and *Micropanchax* Myers 1924. This study focused on two species of *Lacustricola*, *L. myaposae* (Boulenger 1908) and *L. macrurus* (Boulenger 1904) and three species of *Micropanchax*, *M. johnstoni* (Günther, 1894), *M. katangae* (Boulenger 1912) and *M. hutereaui* (Boulenger, 1913) that show subtle variations in colouration and morphology which could indicate the presence of cryptic species or, in some cases of complete geographic isolation, completely new species (Bickford *et al.* 2006). The previous chapter summarised molecular evidence that such cryptic taxa actually exist amongst these populations, and posed the question of how there lineages can be distinguished phenotypically.

The traditional taxonomy of the African Procatopodidae was based on their external morphological features, dentition, frontal sensory systems, and colouration with a large focus on osteological features to distinguish new species in previous studies (Parenti 1981; Huber 1982; Ghedotti 2000). The species-level taxonomy of the African topminnows has not been significantly updated in the last 20 years and was previously poorly described (Ghedotti 2000; Skelton 2001; Steenkamp *et al.* 2001; Lucinda & Reis 2005). There is a chance that some of the cryptic taxa revealed in the previous chapter have been described base on these subtle differences, and then placed in synonymy because the differences were regarded as too subtle. For this reason, the type material of *Micropanchax baudoni* (Myers 1924) was included in the study to assess its current status as a synonym of *M. hutereaui*.

3.1.2. Morphometrics

Morphometrics is an integral part in taxonomy and systematics, and comparing anatomical features of different organisms plays an important role in the biological sciences (Adams *et al.* 2004). Describing the qualitative morphology of organisms was the historical base of taxonomic classifications and understanding biological diversity (Cadrin & Friedland 1999; Adams *et al.* 2004). The twentieth century brought about a revolution of quantitative scientific methods and more advanced statistical methods were developed. This has allowed tremendous advances in the study of morphometrics which allow us to better describe body shape and size (Cadrin & Friedland 1999; Adams *et al.* 2004). These advances have allowed taxonomists to use phenotypic variation such as morphological and meristic differences to distinguish and identify different fish species. Morphological characters have successfully been used in fisheries science as evidence of stock discrimination and taxonomic relationships (Adams *et al.* 2004; Will & Rubinoff 2004).

Environmental factors such as habitat type and biotic and abiotic ecological factors may influence phenotypic characteristics of fish populations (Wessels 2009; Welsh *et al.* 2013). Phenotypic variation and morphological differences are thus not a guarantee of genetic isolation because phenotypic changes may have arisen due to environmental conditions and therefore cannot always be accepted as evidence of genetic differentiation (Turan 1999; Wessels 2009). Phenotypic plasticity allows fish and other organisms to adapt to environmental variation and behavioural changes which may arise such as changes in physiology and morphology (Turan 1999). Evidence of genetic differentiation causing changes in the phenotype would have to be collected to show that environmental conditions were not the cause of the phenotypic variation (Wessels 2009).

Multivariate morphometrics has shown that phenotypic variation induced by environmental factors has advantages in stock identification of species with large population structure and low genetic drift. The use of new approaches to morphometrics and a suite of morphological characters have provided strong support for the use of morphometrics as an independent identification tool. This, coupled with the use of genetic identification, could provide the most accurate way to differentiate relationships within and between species (Turan 1999; Will & Rubinoff 2004; Mojekwu & Anumudu 2015).

Strauss and Booksteins' (1982) box truss method of placing measurements at homologous landmarks further advanced the discriminating power and effectiveness of morphometric methods from the older traditional measurement methods. Traditional morphometrics included the use of meristic variables such as scale and vertebrae counts, with linear measurements occurring on the fish without a coordinate or landmark system (Strauss & Bond 1990). These methods identify fewer diagnostic characters than newer methods and may not distinguish between species that show morphological similarity in shape and colour (Kerschbaumer & Sturmbauer 2011). The box truss method allowed the measurements to cover the landmark topography evenly while allowing the detection of shape changes in orthogonal directions (Strauss and Bookstein 1982). This method allowed physical form to be accurately mapped or archived and reproduced on other specimens, making it a powerful tool for the classification and identification of species, characterisation of growth and allometry within and between populations, and the recognition of and compensation for random measurement error (Strauss and Bookstein 1982). The newer geometric morphometrics uses a system of landmarks to identify structures on the body of the fish without the use of measurements in two-dimensional or three-dimensional planes (Kerschbaumer & Sturmbauer 2011).

For this study, it was felt that the box truss method would efficiently distinguish between species and populations. Numerous studies have shown the effectiveness of multivariate morphometrics and the use of the box truss system for species delineation, identification and differences in fish stocks. Douglas and Matthews (1992) used multivariate morphometrics in an attempt to find a relationship between morphology and ecology of fish species from the families Cyprinidae, Percidae, Catostomidae, Ictaluridae and Cotidae. Although this study did not focus on taxonomic relationships of the fish, they found the morphometric results still placed the species into taxonomic groups with this occurring at family level too (Douglas & Matthews 1992). Ingenito and Buckup (2005) used a mixture of traditional measurements, truss-based measurements and meristic counts to describe a new species, *Parodon moreirai* Ingenito & Buckup, 2005 (Family Parodontidae), from the headwater streams in the northeastern slopes of the Serra da Mantiqueira in Brazil. These morphometric methods were used to assess the morphology of the new species and to place it in the correct genus. The final generic placement was based on the presence of teeth (Ingenito and Buckup 2005). The morphometric analysis was also used to uncover diagnostic features for identification.

Hossain *et al.* (2010) used truss-based morphometrics to quantify differences in the shape of three geographically isolated stocks of the carp *Labeo calbasu* (Cyprinidae) in Bangladesh, specifically two wild populations in the Jamuna and Halda Rivers, and to distinguish these wild populations from a hatchery population. The hatchery stock showed a form intermediate to those of the other two stocks, perhaps because the hatchery originally obtained breeding material from both wild stocks (Hossain *et al.* 2010). The two wild stocks showed different morphologies that could have arisen due to geographical isolation.

Istead *et al.* (2015) reared the gibbose Centrarchids, *Lepomis macrochirus, Lepomis gibbosus* and *Ambloplites rupestris*, in different water flow conditions to test the fishes' phenotypic responses and used truss-based morphometrics to show that all three species developed more streamline bodies in flowing water compared to fish reared in static water. The study also found that phenotypic responses were species-specific and varied in degree in each of the three species (Istead *et al.* 2015).

Ridden *et al.* (2016) undertook a phylogeographic and morphometric study to assess the status of *Mesobola brevianalis* (Family Danionidae) and its taxonomic synonyms. Molecular phylogenetic analysis revealed a number of distinct clades and truss-based morphometrics identified unique characteristics of each clade. The study found that the genera *Mesobola* and *Engraulicypris* were synonyms and allowed the restoration of the species *E. gariepinus*, which was found to be the only morphologically distinct clade in the truss-based analysis. Live colouration was used as unique identifiers for other genetically distinct lineages. As a result of the genetic and morphometric analysis, two new species were described for the genus *Engraulicypris*, *E. howesi* Riddin *et al.*, 2016 and *E. ngalala* Riddin *et al.*, 2016.

3.1.3. Aims

This study aims to clarify the morphological patterns and differences of the southern African topminnow taxa found in the molecular phylogenetic analysis using Principal Component Analysis (PCA) and to clarify, using Discriminant Function Analysis (DFA), possible identifying characters for the molecular clades that would be diagnostic without genetic testing. Once identifying characters are known, new species can be described.

3.2. Materials and methods

Samples were collected from eleven countries across central and southern Africa, from 24 river systems (Table 3.1; Fig 3.1). Fish were caught using seine netting, electrofishing, and hand-held D-nets to sample vegetated areas along river banks, anaesthetized using a combination of clove oil and ethanol in water and placed in formalin to fix the tissue for short-term storage. Fish were photographed in the field to record their live colouration because storage in formalin often degrades the live colours of the fish if left for extended periods. Fish were placed in separate jars with a waterproof label that included the date and method of capture and location such as GPS coordinates and river details (Table 3.1). The samples were later transferred to 70% ethanol for long-term storage, catalogued and kept at the South African Institute for Aquatic Biodiversity (SAIAB), where they could be easily accessed.

Specimens were identified in the field by collectors or identified in the collection facility at SAIAB. Specimens collected from different locations and at different dates were identified by different people, often depending on who was working at SAIAB at the time. Newer material was identified by R. M. van Zeeventer and I. R. Bills. Data from Chapter 2 were used to detect field misidentifications where possible.

Specimens of each species (Table 3.1) were placed in a white-based container with 70% ethanol to act as a photographic stage. Photographs were taken with a Nikon D3100 digital SLR camera (14.1 megapixels) with a 55 mm lens. The photographs showed the left side of each specimen and included a 50 mm scale bar and a waterproof label indicating the specimen's number, photograph number and taxon. After being photographed, each specimen was labelled with a small piece of waterproof paper to indicate its specimen number and placed into the storage jar so that further photographs could be taken if necessary. Where possible, groups were tested based on the outcome of the phylogenetic analysis in Chapter 2.

In total, 933 specimens were photographed, including type specimens from other institutions (Table 3.1).

Photographs of all of the specimens, including topotypes, holotypes and syntypes wherever possible, were imported into imaging analysis software called *analysis docu* (Soft Imaging Systems: http://www.soft-imaging.net/). The measurements were taken digitally to provide a record of the shape data to be stored and recalled as necessary. It also made re-measuring certain aspects of specimens easier and more repeatable because the process projected a three-dimensional specimen onto two dimensions in a consistent way. The advancement of image processing software allowed for measurement error to be decreased and more accurate measurements to be taken (Cadrin & Friedland 1999).

Ten landmarks were identified on each fish (Fig 3.2 A, Table 3.2). Landmarks are points on the body of the fish that indicate the same anatomical features of all sampled individuals and are normally anatomically homologous and easy to locate unambiguously (Strauss & Bookstein 1982; Schaefer et al. 2011). Examples of landmarks are points on the fish's body where different structures intersect, such as anal pores or the insertions of fins. The landmarks are then connected by line segments that form a box truss network (Strauss & Bookstein 1982; Wessels 2009). The box truss network in this study consisted of 21 measurements (Fig 3.1 A). Six additional traditional linear measurements were taken to help to provide traditional diagnostic features of the fish species and to provide extra measurements that may be standardized across all of the species (Fig 3.1 A & B Table 3.2). All measurements were input into an Excel spreadsheet with the following information: specimen number, current taxon name, catalogue number, country, river, river system and latitude and longitude. To linearise any allometric covariation, the measurements were log-transformed before analysis (Strauss & Bookstein 1982).

STATISTICA v11 (Statsoft, 2012) was used to perform statistical analyses of the data. Principal component analysis (PCA) (Bookstein *et al.* 1985; Wold *et al.* 1987; Kocovsky *et al.* 2009) was used to explore the general pattern of morphological variation between genera, between species and within species, spread across different river systems and sections within the same river system, for example the Upper and Lower Zambezi River.

Discriminant Function Analysis (DFA) (Cadrin 2000) of the log-transformed data was used to test the morphometric diagnosability of groups shown to be distinct from the PCA and to determine whether any characteristics distinguished them. Of the total 933 specimens, only 922 were used in the final analyses. The specimens were removed from the analysis due to important structures such as caudal and dorsal fins being damaged or entirely missing, or degradation during storage that did not allow for all of the measurements to be taken. The last two measurements, length of the middle caudal ray (measurement 26) and total length, were excluded from the analysis to ensure that the available type material was included in all analyses.

3.3. Results

3.3.1. Morphometric variation across southern African genera

The PCA for the morphometric variation of the four genera tested in this analysis captured 90.37% of the variation in Factors 1 to 3. Factor 1, summarising 81.30% of the variation, showed all negative eigenvector coefficients, which expressed the overall size of the specimens, and Factors 2 (6.35%) and 3 (2.72%) described the shape of the specimens (Table 3.4), which was clearly a small component of the overall variation.

Figure 3.2 depicted the weightings of the eigenvectors of Factors 2 and 3 in relation to the box truss, indicating where the most morphometric change occurred. The eigenvector coefficients of Factor 2 showed the greatest change around measurements L3 and L4, which relate to the position of the dorsal fin and the length of the tail (Fig 3.1 A). The eigenvector weightings of Factor 3 showed the greatest change in measurements L6, L7 and L20, which relate to changes in the length of the tail and depth of the body between measurement L3 and L7 (Table 3.5; Fig 3.2 B).

Despite the changes shown by the eigenvector coefficients of Factors 2 and 3, their scatterplot showed no clear discriminations of the genera, with most specimens occupying the same factor space (Fig 3.4).

Five of the 25 measurements (L1, L10, L11, L12 and L19: Table 3.2) were removed from the DFA as they were shown to be redundant and below the minimum tolerance limit of 0.01. The classification matrix showed that the bulk of the specimens in each genus were correctly identified by the DFA, with a large overlap occurring between the memberships of *Micropanchax* and *Lacustricola*. *Aplocheilichthys* was the only genus to be consistently identified correctly (Table 3.3 A). Figure 3.2 depicted the box truss with the coefficients of the DFA eigenvectors of Roots 1 and 2 to show which measurements were potentially diagnostic features. The coefficients of Root 1 showed that body length (measurements L2 and L9), head depth and body depth (measurements L15 and L20) played a factor in identifying genera, but the coefficients were relatively low (Fig 3.2 C). The coefficients of Root 2 showed the length of the tail (measurements L4 and L6), head depth (measurement L15), body depth (measurements L13, L16, L18 and L21) and head length (measurement L22) had the highest impact in identifying the genera (Table 3.6; Fig 3.2 D).

The DFA showed that the four genera could be distinguished partially (Wilks' Lambda = 0.30; F = 22.17, $p < 0.05$). Canonical analysis identified three statistically significant discriminant functions ($\chi^2 = 1093.06$, $\chi^2 = 460.74$ and 141.19, and df = 60, 36 and 17 respectively, $P < 0.05$ for all factors). The first discriminant function accounted for 63.01% of the discriminant capacity, and Root 2 for 10.55%. The scatterplot of Root 1 vs. Root 2 showed a similar picture to that of the classification matrix, with slight separation of the genera in morphospace. *Aplocheilichthys* was the only genus to occupy its own region of factor space, with a few specimens from *Micropanchax* that were estimated to belong to *Aplocheilichthys*. *Micropanchax*, *Lacustricola* and *Hypsopanchax* showed large overlap in the middle of the factor space (Fig 3.5).

3.3.2. Morphometric variation between southern African species

The PCA for the morphometric variation of the African species, *M. katangae*, *M. johnstoni*, *M. hutereaui*, *L. myaposae*, *L. macrurus*, *L. mediolateralis*, *Micropanchax* sp., *Lacustricola* sp., *M. baudoni*, three genetic clades of *M. johnstoni* and one clade of *M. hutereaui*, seen in Chapter 2, showed that 90.37% of the variation occurred in Factors 1 to 3. Factor 1 summarised 81.30% of the variation and showed all negative eigenvector coefficients, which summarised the overall size of the specimens. Factors 2 (6.35%) and 3 (2.72%) described the shape of the specimens as in the genus-level analysis (Table 3.5). The PCA therefore showed the same shape changes between the specimens of the different species as the PCA for the variation between different species and the different clades of *M. johnstoni* and *M. hutereaui*, and the weightings of the eigenvectors of the PCA in relation to the box truss network were the same as the PCA for the variation between genera (Fig. 3.3).

The scatterplot for PCA Factors 1 and 2 showed small differences between the different groups but was not presented in the thesis because Factor 1 was based on the size of the specimens used. The scatterplot of PCA Factor 2 and 3 demonstrated limited morphometric variation between the groups (Fig. 3.6). Each species cluster was found to be more dispersed over the factor space, with slight separation between *M. katangae* and *M. johnstoni*, but *M. hutereaui* and *M. johnstoni* showed extensive overlap. The specimens from Clades C, D and E of the phylogenetic analysis plotted in the same region of factor space as the main cluster of *M. johnstoni*, but Clades C and E plotted higher in the factor space than Clade D. This showed that slight differences in shape occurred between Clades C and E, and Clade D but was small and shape change was not enough to separate them from the rest of *M. johnstoni* as in the phylogenetic analysis. Clade C included the type specimens of *M. johnstoni* (Fig. 3.6). The main group of *M. hutereaui* included the type specimens and overlapped with Clade A nearer to the top of the factor space. The overlap between *M. hutereaui* and Clade A included the two type specimens of *M. baudoni* and would agree with the current treatment of *M. baudoni* as a synonym of *M. hutereaui*. Specimens from Clade B could not be included in the morphometric analysis and were not represented in the PCA.

The species of *Lacustricola* (*L. myaposae*, *L. macrurus*, *L. mediolateralis* and the unidentified *Lacustricola* specimens) all grouped within the same factor space as *M. johnstoni*, indicating that these species share a similar body shape. The unidentified *Micropanchax* not in Clade C formed two groups, one group plotted in the same factor space as *M. johnstoni* and the other group plotted nearer to *M. katangae* in the factor space (Fig 3.6). These specimens plotted nearer to *M. katangae* showed a similar relationship to Clade E seen in the phylogram of the previous chapter. *Lacustricola pumilus* and *L. matthesi* plotted in the same factor space at the top of the *M. katangae* factor space and showed a close relationship to *L. myaposae* but less close to *L. macrurus*. *Hypsopanchax jubbi* plotted in the center of the factor space while the unidentified specimens of *Hypsopanchax* showed a slight shape difference from *H. jubbi* and plotted higher in the factor space (Fig. 3.6). Finally *A. spilauchen* showed no real shape differences from the rest of the species and plotted centrally in the factor space.

Four of the 25 measurements (L1, L10, L11 and L12: Table 3.2) were removed from the DFA as they were shown to be redundant and below the minimum tolerance limit of 0.01. The classification matrix of the DFA was only able to assign specimens of *A. spilauchen*, *H. jubbi*, and *L. pumilus* consistently correctly. The rest of the groups' classification rates ranged between 85% and 5% correct (Table 3.4). The phylogenetic Clades C, D and E showed very low classification scores, with most of the specimens in Clades C and E assigned to the main

M. johnstoni group and Clade D with only 60% of its specimens assigned to the Clade D group. Clade A was split with six specimens assigned to the main *M. hutereaui* group, one specimen in the *L. macrurus* group and seven in the Clade A group (Table 3.4).

The DFA partially distinguished the species (Wilks' Lambda = 0.21; F = 10.60, $p < 0.05$). Canonical analysis identified statistically significant discriminant functions for Roots 1 and 2 (χ^2 = 3518.92 and 2296.76, and df = 378 and 340, respectively; $P < 0.05$ for both factors). The first discriminant function accounted for 46.69% of the discriminant capacity, and the second for 18.32%. The eigenvector of Root 1 showed that the measurements for the length of the tail, position of the anal fin, depth of the head, depth of the body before the tail and eye diameter were the most important for identification (measurements L6, L7, L15, L17, L20 and L23; Fig 3.3). The eigenvector of Root 2 showed that the length of the specimens' back and dorsal fin position (measurement L3), depth of the caudal peduncle (measurement L14), the size and depth of the head (measurements L9 and L15), and depth of the body before the tail (measurement L20) were important for identification and could possibly be used as diagnostic features (Table 3.7) (Fig 3.3).

The scatterplot of Roots 1 and 2 showed how the different groups plotted in the factor space but only showed a large overlap of the main *M. hutereaui* group and *M. katangae* with the two groups differing slightly from *M. johnstoni* (Fig. 3.7). The species of *Lacustricola*, *L. myaposae*, *L. macrurus* and *L. mediolateralis*, and the unidentified *Lacustricola* specimens, all grouped in the same factor space as *M. johnstoni*, which showed that the southern African species of *Lacustricola* have similar body shapes to that of the *Micropanchax* species, which would make them difficult to identify morphologically. *Lacustricola myaposae* showed resolution into two groups which border each other in the factor space. The phylogenetic Clades C, D and E showed no real distinction in the morphological factor space and plotted in the same region as the *M. johnstoni* group. Clade A plotted with the specimens of *M. hutereaui* in the factor space and showed no diagnostic shape differences (Fig 3.7). The unidentified specimens of *Micropanchax* formed two groups, one tightly clustered and showing similarity to *M. katangae*. The second group was more spread across the factor space, with some specimens plotted near *M. hutereaui* and *A. spilauchen*. *Aplocheilichthys spilauchen* showed no overlap with any of the other specimens except for one lone specimen of *M. hutereaui* (Fig. 3.7).

The specimens of *L. matthesi*, *L. pumilus* and *M. baudoni* overlapped largely with the species groups of *M. katangae* and *M. hutereaui* in the factor space. The specimens of *H. jubbi* and the unidentified *Hypsopanchax* plotted in the center of the factor space with a slight shape change between the two groups (Fig. 3.7).

3.3.3. Variation within species between river systems

3.3.3.1. Micropanchax johnstoni

The PCA for the morphometric variation of the river system populations tested in this analysis showed that 90.47% of the variation occurred in Factors 1 to 3. Factor 1 summarised 81.48% of the variation, and showed all negative eigenvector values, which summarised the overall size of the specimens. Factors 2 (6.50%) and 3 (2.49%) described the shape of the specimens. The eigenvector coefficients for Factor 2 showed similar measurements to the previous analysis were important in the shape change with most important changes occurring with the position of the dorsal fin, anal fin, length of the tail, and depth of the body between the dorsal and anal fins (measurements L3, L4, L7 and L20). Factor 3 showed large eigenvector coefficients for the head length, length of the tail, body depth between the dorsal and anal fin, depth of the caudal region, eye diameter and length of the anal fin base (measurements L1, L4, L6, L7, L18, L20, L21, L23). The scatterplot for Factors 1 and 2 showed most of the specimens from the river systems grouped in the same region of the factor space but a group of seven specimens from the Shire River System grouped on their own. Factors 2 and 3, showed a large dispersal of specimens from all the river systems throughout the factor space but indicating that the morphology of *M. johnstoni* is difficult to distinguish between river systems (Fig. 3.8). In the scatterplot for Factor 2 and 3, the seven specimens of the Shire River System grouped more closely with the rest of the specimens. These seven specimens were the syntype specimens for *M. johnstoni* (Fig. 3.8).

Two of the 25 measurements (L15 and L19: Table 3.2) were removed from the DFA comparing the river systems of *M. johnstoni* as these measurements did not meet the minimum tolerance level for the analysis. The classification matrix showed a large overlap between all of the river systems and suggested that there was no morphological discrimination across the distribution of *M. johnstoni* except for the Mhlatuze River population (Table 3.7 A).

The DFA was able to show some separation between some, but not all, groups that showed no separation (Wilks' Lambda = 0.02; F = 5.43, $p < 0.05$). Canonical analysis identified statistically significant discriminant functions for Roots 1 and 2 ($\chi^2 = 1055.52$ and 757.20, and df = 230 and 198, respectively; $P < 0.05$ for both factors). The canonical variable showed that 35.16% of the variation occurred in Root 1 and 22.06% occurred in Root 2. The eigenvectors of Root 1 showed that the most diagnostic features for the specimens were found in the length of the top of the head region (measurement L1), depth of the main body section (measurement L16), length of the tail (measurement L4), length of the anterior body region (measurement L8), length of the bottom of the head (measurement L9) and the depth of the caudal region (measurement L18). The eigenvectors from Root 2 confirmed the diagnostic role of the length of the bottom of the head, depth of the main body section and caudal region (measurements L9, L16 and L18) but added that the dorsal body length, anterior head length, head depth, position of the anal fin and the length of tail played a role (measurements L2, L6, L7, L10 and L11), and the depth of the body between the dorsal and anal fin, and the depth of tail, played a significant role (measurements L20 and L21).

The scatterplot for Roots 1 and 2 showed that there was discrimination between the Upper and Lower Zambezi systems and the Upper Zambezi system from the other river systems with a slight overlap occurring with the Shire and Limpopo Systems. The Okavango, Kwando, Lower Zambezi, Lake Bangweulu, Luapula and Congo systems grouped together in the same factor space making it difficult to distinguish between them by morphology (Fig. 3.9). The specimens from the Mkuze Swamps, Shire and Limpopo systems shared a similar shape and groups together but closely grouped to the other specimens showing that the changes in morphology would be difficult to distinguish in the field. The specimens from the Mhlatuze System and one specimen from the Limpopo System in South Africa displayed similar morphologies and grouped together, with no overlap from the other river systems (Fig 3.9).

3.3.3.2. Micropanchax hutereaui

The PCA for *M. hutereaui,* including eight *Micropanchax* sp. and two *M. baudoni* specimens showed that 90.61% of the variation occurred in Factors 1 to 3. Factor 1 summarised 77.25% of the variation, and showed all negative eigenvector values, which summarised the overall size of the specimens. Factors 2 (10.34%) and 3 (3.02%) described the shape of the specimens. The eigenvectors for both Factors 2 and 3 showed similar patterns to the previous analysis, with the most variation occurring near the dorsal fin, anal fin and in the tail, with caudal length and body depth playing a minor role (measurements L3, L4, L6, L7, L20, and L25).

A scatterplot of Factors 2 and 3 for the three groups, *M. hutereaui*, *Micropanchax* sp. and *M. baudoni,* showed that *M. hutereaui* had a widespread across the factor space and showed two areas where specimens grouped the most with the type material for *M. hutereaui* plotted in the middle between the two groups (Fig. 3.10). *Micropanchax baudoni* grouped with the *M. hutereaui* group on the positive side of the factor space and *Micropanchax* sp. showed discrimination from the rest of the specimens towards the top of the factor space (Fig. 3.10).

When the scatter plot for the PCA was coloured in relation to the river systems rather than the taxa, the Congo River System was split into two groups, one of which overlapped with the Upper Zambezi System and some specimens from the Okavango System (Fig. 3.11). There were differences between the Upper Zambezi and the Lower Zambezi, with twelve Lower Zambezi specimens grouping with Upper Zambezi specimens. The Kwando System specimens grouped on the positive side of the factor space with the Lower Zambezi, Okavango and *M. baudoni* (Chad River Basin) specimens (Fig. 3.11).

Five of the 25 measurements (L1, L2, L6, L21 and L22: Table 3.2) were removed from the DFA for the variation between *M. hutereaui*, *Micropanchax* sp. and M. *baudoni* as they were shown to be redundant and below the minimum tolerance limit of 0.01. The classification matrix of the DFA showed an overall classification success of 99.55% (Table 3.7 B). The DFA was able to discriminate between the three groups (Wilks' Lambda = 0.46 and $F = 4.78$, $p < 0.05$). Canonical analysis identified statistically significant discriminant functions for Root 1 and 2 ($\chi^2 = 163.64$ and 30.98, and df = 40 and 19, respectively; $P < 0.05$ for both factors).

The canonical variable analysis showed that 88.93% of the variation occurred in Root 1 and 11.07% occurred in Root 2. The eigenvectors for the variation between the three groups showed that depth of the caudal peduncle (measurements L5 and L18), depth of the head region (measurements L11 and L12) and body depth (measurements L17 and L20) had the most diagnostic value in Root 1. In Root 2, the eigenvectors showed that the position of the dorsal fin (measurement L3), length between the head region and anal fin (measurement L8), depth of the head region (measurement L11) and body depth (measurement L19) had the most diagnostic value. The scatterplot of Roots 1 and 2 showed clear discrimination between *M. hutereaui*, *Micropanchax* sp. and *M. baudoni* (Fig. 3.12).

Four of the 25 measurements (L1, L2, L21 and L22: Table 3.2) were removed from the DFA of *M. hutereaui* river systems as these measurements did not reach the minimum tolerance level for the analysis. The classification matrix (Table 3.7 C) showed that most specimens could be correctly assigned.

The DFA was able to show the discrimination between the river systems (Wilks' Lambda = 0.10; F = 5.48, $p < 0.05$). Canonical analysis identified statistically significant discriminant functions for Roots 1 and 2 (χ^2 = 474.82 and 229.86, and df = 105 and 80, respectively; $P < 0.05$ for both factors). The canonical variable showed that 61.79% of the variation occurred in Root 1 and 23.82% occurred in Root 2. The eigenvectors for Root 1 showed that the position of the dorsal and anal fins (measurement L3, L7 and L8) and body depth (measurements L12 and L16) had the most diagnostic value. The eigenvector for Root 2 showed that body depth (measurements L11, L13, L15 and L20) and eye diameter had the most diagnostic value. The scatterplot for both Roots 1 and 2 showed the discrimination of the different river systems and the overlap of the systems in the factor space (Fig 3.13). The scatterplot showed that the Congo system specimens grouped the most on their own but still showed similar body shape to the other river systems, which shared more of the same factor space (Fig. 3.13). The DFA for both between the species and the river systems split the groups on the scatterplot and was able to assign the specimens into their respective groups with confidence but showed that the characters used to diagnose the specimens are difficult to identify in the field.

3.3.3.3. Micropanchax katangae

The PCA for the morphometric variation of the four genera tested in this analysis showed that 87.04% of the variation occurred in Factors 1 to 3. Factor 1 of the PCA summarised 75.89% of the variation, showed all negative eigenvector values, which summarised the overall size of the specimens. Factors 2 (7.66%) and 3 (3.49%) described the shape of the specimens. The eigenvectors scores for Factor 2 displayed that the measurements most responsible for shape change were the position of the dorsal fin and length of the tail (measurements L3 and L4), position of the anal fin (measurement L7) and the depth of the body between the dorsal and anal fin (measurement L20). The eigenvector scores of Factor 3 suggested most of the shape change occurred in the dorsal and anterior caudal length (measurements L4 and L6), the length of the anal fin base (measurement L7) and the depth of the body and caudal regions (measurement L18 and L20).

The scatter plot of Factors 1 and 2 showed that there was a gradient in change in the river systems with specimens from the Lower Zambezi, Upper Zambezi, Kwando, St Lucia, Lake Xingute systems and the two type material specimens from the Luapula System grouped in the same region of the factor space. Specimens from the Okavango System grouped in this area but showed a spread across most of the factor space. The Luapula System showed the furthest split from the main group and the Lake Sibaya System plotted between the Luapula System and the overlapped river systems at the bottom of the factor space (Fig 3.14). The scatterplot of Factors 2 and 3 showed the same groupings as the scatterplot of Factors 1 and 2 but with tighter grouping. The scatterplot also showed that the Lake Sibaya specimens have a shape closely related to the Okavango specimens (Fig 3.14).

None of the 25 measurements of the DFA for the variation between the different river systems were removed from the DFA. The classification matrix (Table 3.7 D) showed that most specimens could be correctly assigned.

The DFA was able to show some discrimination between the groups and some groups that showed no discrimination from each other (Wilks' Lambda = 0.04; $F = 3.55$, $p < 0.05$). Canonical analysis identified statistically significant discriminant functions for Roots 1 and 2 ($\chi^2 = 610.94$ and 364.34, and df = 200 and 168, respectively; $P < 0.05$ for both factors). The canonical variable showed that 40.73% of the variation occurred in Root 1 and 19.75% in Root 2. The eigenvector coefficients of Root 1 displayed that body depth (measurements L16, L19 and L20) and length of the dorsal body and caudal regions (measurements L1 and L4) to have the most diagnostic ability. The eigenvectors of Root 2 showed that measurements in the head region (measurements L10 and L11), length of the dorsal body region (measurement L2), anterior distance between the pectoral fin and head (measurement L9), and body depth (measurements L13 and L19) had important diagnostic value.

The scatterplot showed that the Okavango, Upper Zambezi, Lower Zambezi, Congo, Lake Xingute, St Lucia systems and the type material of *M. katangae* plotted in the same region of the factor space. The Kwando System showed some specimens plotted with this main group but moved more positively on Root 2. Lake Sibaya specimens plotted lower on Root 2 but showed no clear grouping together and formed pairs in the factor space (Fig.

3.15). The Luapula system plotted positively on Root 1 and Root 2 and demonstrated discrimination from the other specimens from the other river systems (Fig. 3.15).

3.3.3.4. Lacustricola myaposae

The PCA for the morphometric variation of the four genera tested in this analysis showed that 88.32% of the variation occurred in Factors 1 to 3. Factor 1 of the PCA summarising 76.79% of the variation, showed all negative eigenvector values, which summarised the overall size of the specimens. Factors 2 (7.38%) and 3 (4.15%) described the shape of the specimens. The eigenvector coefficients for Factor 2 showed the position of the anal fin (measurement L7 and L8) and body depth (measurement L13, L16 and L20) had the greatest variation in the specimens. The eigenvector coefficients for Factor 3 showed the greatest variation in the position of the dorsal fin and length of the tail (measurement L3, L4 and L6).

The scatterplot of Factors 2 and 3 displayed overlap between the specimens but three groups formed with the coastal specimens forming one group, Lake Piti system forming one group and Mhlathuze System forming the last group. Four of the Lake Sibaya specimens plotted within the Lake Piti group, five specimens grouping with the Mhlathuze system and one specimen grouping close to the coastal specimens (Fig. 3.16).

Five of the 25 measurements (L10, L11, L22, L23 and L24: Table 3.2) of the DFA for the variation between the different river systems of *L. myaposae* were removed from the DFA as these measurements did not reach the minimum tolerance level for the analysis. The classification matrix was able to assign all of the specimens into their respective river systems correctly (Table 3.7 E). The DFA was able to split the four groups (Wilks' Lambda = 0.01; $F = 7.80$, $p < 0.05$) and canonical analysis identified statistically significant discriminant functions for Roots 1 and 2 ($\chi^2 = 229.45$ and 99.59, and df = 60 and 38, respectively; $P < 0.05$ for both factors). The canonical variables showed that 80.18% of the variation occurred in Root 1 and 10.66% in Root 2.

The eigenvector coefficients of Root 1 suggested that body depth between the dorsal and anal fins showed the most diagnostic value (measurements L17 and L20), length of the top of the head and head depth (measurements L1 and L15), and position of the anal fin (measurement L7) played a more minor role for discriminating the different groups. The eigenvector scores of Root 2 showed body depth (measurements L13, L16, L19 and L21) had the greatest role in discriminating the groups, with the forward change in dorsal fin position playing a more minor role (measurements L2 and L3). The scatterplot of Roots 1 and 2 showed clear discrimination of the Coastal specimens and Lake Piti System which each formed its own group. The Lake Sibaya and Mhlathuze System grouped close to one another with one specimen from the Mhlathuze System grouped in the middle of the Lake Sibaya specimens (Fig 3.17). The DFA and canonical analysis showed that the specimens could be discriminated morphologically by system but the difference is very subtle and nearly impossible to use in the field.

3.3.3.5. Lacustricola macrurus

The PCA for the morphometric variation of the four genera tested in this analysis showed that 78.92% of the variation occurred in Factors 1 to 4. Factor 1 of the PCA summarised the overall 81.30% of the variation, showed all negative eigenvector values, which summarised the overall size of the specimens.. Factors 2 (12.61%) and 3 (10.35%) described the shape of the specimens. The eigenvector coefficients of Factors 2 and 3 showed that the length of the tail changed the position of the dorsal and anal fins (measurements L3, L4, L6 and L7), the length of the bottom of the body between the anal fin and the head region (measurement L8), size of the head (measurement L10) body and tail depth (measurements L16, L20 and L21), and the eye diameter (measurement L23).

The scatterplot for Factors 2 and 3 showed the specimens spread across most of the factor space and formed two main groups close to one another. Four specimens from the Cunene System split away from the rest of the group. Three of these specimens were from Epupa on the Cunene River and one specimen from Ruacana Falls on the Cunene River (Fig. 3.18).

Six of the 25 measurements (L1, L5, L7, L9, 17 and L19: Table 3.2) of the DFA for the variation between the different river systems of *L. macrurus* were removed from the DFA as these measurements did not reach the minimum tolerance level for the analysis. The classification matrix was able to assign all of the specimens into their respective river systems correctly (Table 3.7 F). The DFA was able to discriminate between the two groups (Wilks' Lambda = 0.12; F = 7.50, $p < 0.05$) and canonical analysis identified one statistically significant discriminant function ($\chi^2 = 59.72$ and df = 19, $P < 0.05$). The canonical variables showed that all of the variation occurred in Root 1. The eigenvector scores of Root 1 showed that head depth (measurement L12), body depth (measurement L13) and caudal length (measurement L22) played important role in diagnosing the two groups.

3.3.3.6 Micropanchax sp.

The PCA for the morphometric variation of the four genera tested in this analysis showed that 96.91% of the variation occurred in Factor 1 to Factor 3. Factor 1 of the PCA summarising 94.07% of the variation, showed all negative eigenvector values, which summarised the overall size of the specimens. Factor 2 (1.82%) and Factor 3 (1.02%) described the shape of the specimens. The eigenvector coefficients for Factor 2 and Factor 3 showed similar results to the previous analyses. The eigenvector coefficients of the two factors showed that the position of the dorsal and anal fin (Measurements L3, L4, L6 and L8), depth of the body and tail (Measurements L5, L13, L14, L15 and L17) and length of the head (Measurements L10 and L24) were most subject to change.

The scatterplot of Factor 1 and Factor 2 showed that the different groups, *M. johnstoni* type material and *M. katangae* type material, plotted separately in the factor space with the exception of the *M. baudoni* type material, which grouped with the unidentified *Micropanchax* sp. The scatterplot of Factor 2 and Factor 3 showed the four groups plotted in the same region of the factor space with only four specimens of the unidentified specimens plotting in different regions of the factor space (Fig. 3.19).

Eleven of the 25 measurements (L1, L7, L9, L11, L15, L16, L18, L19, L20, L24 and L25: Table 3.2) the DFA for the variation between the different river systems of *L. macrurus* were removed from the DFA as they were shown to be redundant and less than the minimum tolerance limit of 0.01. The classification matrix was able to assign all the specimens into their respective river systems correctly leaving the overall classification success of the classification matrix at 100%. The DFA was able to discriminate between the four groups (Wilks' Lambda = 0.01; F = 5.27, $p < 0.05$) and canonical analysis identified two statistically significant discriminant functions ($\chi^2 = 126.57$ and 46.99, and df = 42 and 26 respectively, $P < 0.05$ for both factors). The canonical variables showed that 85.02% of the variation occurred in Root 1 and 9.78% of the variation occurred in Root 2.

The eigenvector coefficients of Root 1 showed that body length, tail length, position of the anal fin and body depth (measurements L6, L8, L13 and L21) had important diagnostic ability but the most important measurement in Root 1 was caudal length (measurement L22). The eigenvector coefficients of Root 2 showed tail length (measurement L6), position of the dorsal fin (measurement L4), and body depth (measurements L12, L13 and L14) were important in discriminating between the four groups. The scatterplot of Roots 1 and 2 showed that *M. johnstoni*, *Micropanchax* sp. and *M. katangae* plotted in the middle of Root 2 but were separated along the Root 1 axis. *Micropanchax baudoni* plotted positively in the factor space for both Roots 1 and 2 (Fig. 3.20).

3.4. Discussion

This morphometric study found variation between the species and some variation within the species between the river systems. Some of the populations defined in the haplotype trees did not present informative morphometric values to distinguish them, with some species showing variation relating to geography. The measurement of the holotypes and syntypes of *M. johnstoni* (Günther 1894), *M. katangae* (Boulenger 1912), *M. hutereaui* (Boulenger 1913), *M. baudoni* (Myers 1924), *L. mediolateralis* (Poll 1967) and *H. jubbi* (Poll & Lambert 1965) allowed the defining specimens for those species to be included in the analyses.

3.4.1. Species

3.4.1.1. Micropanchax johnstoni

The morphometric study of the Clades C, D and E and *M. johnstoni sensu stricto* confirmed that it is very difficult to tell the specimens apart using morphology. The specimens all showed elongated, slender bodies compared to *M. katangae* and *M. hutereaui* and attained a greater caudal length than *M. hutereaui* (Skelton 2001). *Micropanchax johnstoni* was shown to be distinguishable from *M. hutereaui* in a new study (Braganca *et al.* in review) by the absence of barred dorsal, anal and caudal fins. *Micropanchax hutereaui* also showed the presence of a reticulate pattern on scale margins Braganca *et al.* in review). The study by Braganca *et al.* (in prep) used cleared, stained specimens to determine the position the unpaired fins. *Micropanchax johnstoni* showed a shorter dorsal fin base, and narrower body and caudal peduncle than *L. myaposae*, with the dorsal fin starting closer to the caudal region in *M. johnstoni* (Braganca *et al.* in review). The first proximal radial of the dorsal fin in *M. johnstoni* occurs between vertebrae 16 and 17 while the first proximal radial of the dorsal fin in *L. myaposae* occurs between vertebrae 13 and 14 (Braganca *et al.* in review). The position of the anal fin showed a slight overlap between *M. johnstoni* and *L. myaposae*, with the first proximal radial of the anal fin occurring between the pleural rib of vertebrae 11 and 13 in *M. johnstoni* and 11 and 12 in *L. myaposae* (Braganca *et al.* in review).

The live colouration of the different populations of *M. johnstoni* was very similar, with only a few differences. Most of the specimens showed blue or turquoise colours along the body and yellow in the dorsal, anal and caudal fins (Skelton 2001; Marshall 2011). Specimens from the Kwando and Zambezi Rivers were very similar with dark eyes and darker bodies than the Okavango River specimens. The Zambezi River specimens lacked yellow in the anal fins. Specimens from the Okavango were overall brighter blue along the body with the upper iris displaying a turquoise stripe (Skelton 2001; Marshall 2011).

3.4.1.2. Micropanchax katangae

Micropanchax katangae showed a deeper body than *M. johnstoni* and attained a longer caudal length than *M. hutereaui* with the most distinctive changes occurring in the live colouration of the fish. The morphometric analysis also showed that the dorsal fin and anal fin have shorter distances between the insertion points of each fin. Most of the specimens of M. katangae showed that their morphologies were very closely overlapped but the Lake Sibaya and Luapula River Systems showing changes in morphology with most of the changes occurring in body depth. The most defining and recognizable characteristic of the colouration of *M. katangae* was the black stripe often along the lateral line of the body of the fish, which displays a zigzag shape (Skelton 2001).

3.4.1.3. Micropanchax hutereaui

Clade B could not be included in the analysis due to inavailability of specimens. Clade A showed no morphometric difference to the main group of *M. hutereaui*. The tightly-grouped specimens of *Micropanchax* sp. thought to possibly be the Pigmy Topminnow showed a close relationship to *M. hutereaui* but had a deeper body. Specimens of *M. hutereaui* from different river systems that had been stored for long periods of time could not be told apart using live colouration as most specimens had lost their colouration during storage. Specimens of *M. hutereaui* photographed and collected in the field showed small deviations in colour. *Micropanchax hutereaui* is generally lightly coloured with a translucent yellowish body commonly with blue reflections on the scales (Skelton 2001). The undescribed Pygmy Topminnow was darker in colour with a more light brown body and only the postero-ventral scales have blue in them (Skelton 2001). Both of these species displayed a turquoise iris.

The specimens of *M. hutereaui* from the Kwando River showed a colouration similar to that of the Pygmy Topminnow. The two differed in the iris, which was dark coloured with a white stripe at the bottom, and more blue around the operculum and lateral line scales (Skelton 2001; Marshall 2011). Specimens collected from the Okavango River differed in that they displayed a completely blue iris with a darker brown body and yellow, green, and blue around the operculum and lateral line scales. The dorsal, anal and caudal fins were yellow without brown spots and the ends of the fins appeared clear (Skelton 2001; Marshall 2011).

3.4.2. Genera

Not all of the genera were morphologically distinct in this study, with only *Aplocheilichthys* assigned consistently correctly. The morphometric results for the different genera indicated that there are taxonomic problems within the family Procatopodinae in southern Africa. *Aplocheilichthys* correctly assigned in the morphometric analysis supports Huber (1982) and Ghedotti (2000) in that *A. spilauchen* is monotypic but the phylogenetic analysis (Chapter 2) did not concur.

The large morphometric misplacement of *Lacustricola* specimens into the *Micropanchax* group was seen with the molecular analysis in Chapter 2, with the close relationship between *M. johnstoni*, *L. macrurus* and *L. myaposae*, which often showed specimens in the same clade together and can often be distinguished in the field only due to live colouration (Skelton 2001). *Lacustricola myaposae* showed a colouration similar to *M. johnstoni*. Fins of the mature males of *L. myaposae* males distinctly showed a black colouration with yellow border and yellow spots in the back. *Lacustricola macrurus* was characterized by a light translucent brown body, while the upper iris is bright white and the lower iris is a duller white. The colour of the body showed light blue sheen along the lateral line scales with a bright blue behind the operculum and all the fins were yellow except the pectorals (Skelton 2001; Marshall 2011). The specimens from Olushandja Dam live colouration had similarities to *L. macrurus* in the iris of the eye and yellow of the fins but showed a darker body with a bright blue along the body below the lateral line.

The *Lacustricola* species also showed large morphological similarities with the species of *Micropanchax* in the species-level morphometric analysis, which agreed with the results of the genus-level analysis. This result suggests that the southern African species *L. myaposae*, *L. macrurus*, *M. johnstoni*, *M. hutereaui* and *M. katangae* need to possibly be placed in the same genus as *L. myaposae* and *L. macrurus* are found further south than the type species *L. pumilus* Boulenger, 1906, which was described from Lake Tanganyika (Boulenger 1906).

3.4.2.1. Micropanchax

The specimens labelled as *Micropanchax* sp. were not assigned to their own group but were rather split between *M. johnstoni*, *M. katangae*, *M. hutereaui* and *L. macrurus*, with the bulk of specimens assigned to the large *M. johnstoni* group seen in the analysis. The unidentified specimens used in the molecular analysis of Chapter 2 showed a similar result, with specimens placed in different clades. The rest of the species and genera showed specimens that were assigned to different groups and illustrated that morphology of each species was not distinctive and was, in some instances, shared. This could be due to convergent adaptation to sharing the same habitat within the same river system (Langerhans *et al.* 2003).

The ordinations suggested that most of the morphometric variation between the species, although very small, occurred around the caudal area and the area of the body between the dorsal and anal fins, with the DFA also indicating a difference in position of the dorsal fin and changes in the eye diameter of the different species. Tail size and eye diameter have often been linked to environmental changes or habitat preference, for example differences between fish in fast-flowing waters such as rivers and fish in slow-moving water such as lakes (Welsh *et al.* 2013; Pelicice *et al.* 2015). Fish living in rivers or faster-flowing waters are often more streamlined than those primarily inhabiting slower-moving water, but this may not always be the case. Lake-dwelling fish may also develop larger eyes to feed on zooplankton and fish, especially in turbid waters, and to improve vision in low-light environments (Welsh *et al.* 2013).

Micropanchax katangae and *L. myaposae* both have lake populations in southern Africa and showed small morphological differences between river and lake specimens in some instances. Lake specimens of *M. katangae* and *L. myaposae* had less streamlined bodies than specimens from fast-flowing water in some systems (Langerhans *et al.* 2003). This diversification may be also be a result of resource composition and could be adaptive to suit a specific niche in the habitat. Phenotypic diversification can be the result of gene flow or phenotypic plasticity. Population mixing may also play a role in reducing the amount of phenotypic diversification shown in samples in different habitats or within connected habitats (Langerhans *et al.* 2003). The results from the ordinations suggested that most of the morphometric change between the different species occurred around the caudal area and the area of the body between the dorsal and anal fins, with the DFA also detecting a difference in the position of the dorsal fin and the eye diameter of the different species. *Lacustricola myaposae* occurred in the coastal region and could be exposed to differing salinity levels which could cause changes in phenotypic expression and variation (Hurwood & Hughs 1998; Warren & Pardew 1998; Engelbrecht & Mulder 1999; Pelicice *et al.* 2015).

The specimens of the study are typically found in similar habitats, generally the marginal areas of rivers and lakes within vegetation and typically not in fast-flowing waters (Skelton 2001; Marshall 2011). Most of the species showed fairly wide distributions across southern Africa, occurring in different river systems but some species had much smaller distributions and in some cases were endemic to only one or two river systems (Skelton 2001; Marshall 2011), with populations that may be connected or not. The species-level ordinations for *M. johnstoni*, *M. katangae*, *M. hutereaui* and *L. myaposae* and the PCA of *L. macrurus* showed morphometric variations between river systems within the different species, to varying but slight degrees. *Micropanchax johnstoni* showed very little geographical morphometric variation. The specimens that could be grouped according to the Clades C, D and E in the phylogenetic analysis showed that Clades C and E shared similar morphologies with the large *M. johnstoni* group which could not be matched to a specific clade. Clade C included the type material of the specimens suggesting that Clade C was the true *M. johnstoni*. These clades showed a slight morphological difference from Clade D. The different clades in the phylogenetic analysis did not show geographical isolation despite their genetic independence, and the fact that the morphometric variation is small made it very difficult to identify the specimens without genetic analysis.

Micropanchax hutereaui had eight unidentified specimens grouped with it in the species analysis, and *M. baudoni*, currently considered a synonym of *M. hutereaui*, had one of two specimens group with *M. hutereaui*. The analysis of just the three groups together showed that there was slight morphometric variation between them with each group separable from the others. There were not enough specimens of *M. baudoni* and no relevant molecular evidence to evaluate reinstating the species as a separate species and it should remain a synonym of *M. hutereaui* until further evidence can be collected.

The eight specimens of *Micropanchax* sp. were believed to belong to an undescribed species referred to as the Pigmy Topminnow (Skelton 2001). The Pigmy Topminnow showed a shorter and deeper body, and displayed different colours to *M. hutereaui* and was found more often in floodplains of the Upper Zambezi and the Okavango Delta (Skelton 2001). This potentially new species will need to be analysed in more detail with the phylogenetic analysis as specimens used in the study were not included in the phylogenetic study.

A surprising result involving *M. hutereaui* and *M. johnstoni* was the change in morphology between the Upper and Lower Zambezi River. This could be due to a number of environmental and anthropogenic factors causing the fragmentation of the gene pool, or just phenotypic plasticity to varying conditions along the length of the Zambezi River (Hurwood & Hughs 1998; Warren & Pardew 1998; Engelbrecht & Mulder 1999). Environmental barriers may also have played a role in restricting gene flow in one direction. The hydrological structure of the Zambezi River was changed significantly in 1959 and 1974 with the completion of the Kariba and Cahora Bassa Dams, respectively, and the construction of dykes along the lower Zambezi River has changed flooding patterns and flow regimes that could have an impact on the morphology of fish species (Beilfuss & Brown 2010).

Anthropogenic disturbances and changes in watercourses may increase the isolation of populations or may even change the direction of gene flow or restrict gene flow in one direction, which could change selective pressures on different populations (Díez-del-Molino *et al.* 2018). If the populations of the Zambezi or other river systems experienced low levels of gene flow between populations within a species, local selection pressures could play a larger role on that population and which phenotype is expressed and may account for the small differences seen (Østbye *et al.* 2018).

3.4.2.2. Lacustricola

Lacustricola pumilus (Boulenger, 1906) the type species of *Lacustricola*, described from Lake Tanganyika, has a distribution primarily restricted to central Africa in countries such as Tanzania and Democratic Republic of Congo (Huber 1999; Fricke *et al.* 2018; Froese & Pauly 2018), while *L. myaposae* was described from the Mposa River in Zululand, KwaZulu-Natal and is restricted to South Africa (Skelton 2001; Fricke *et al.* 2018); *L. macrurus* was described from Lake Sarmento, Marimbo, southern Angola close to the Namibian border and has a distribution restricted to the Kunene River System and the upper reaches of the Okavango System in Namibia and Angola (Skelton 2001; Fricke *et al.* 2018). *Lacustricola* was once placed as a subgenus of *Micropanchax* and as a synonym of *Micropanchax* (Ghedotti 2000) and showed a body shape closer to that of *M. katangae*, with a deeper body than *L. myaposae* and *L. macrurus*, the body shape of which tends to be more elongated, similar to *M. johnstoni*. This difference in body shape and close relationship between *L. myaposae* and *L. macrurus* with *Micropanchax* species in the phylogenetic analysis suggest that the two species may be placed in the wrong genus and should to be transferred to *Micropanchax*.

The African Procatopodidae are egg-laying with external fertilisation, which could allow for mixed fertilisation and gene exchange between populations. *Micropanchax johnstoni* co-occurred with *L. myaposae* and *L. macrurus* at some localities and this may lead to hybridization due to their egg-laying and fertilisation process (Scribner *et al.* 2000; Chiesa *et al.* 2013; Meraner *et al.* 2013). Hybrid offspring may not be infertile in the case of closely-related species of fish and this could lead to the variation in morphology seen in some of the species and the multiple clades of what was thought to be *M. johnstoni* seen in the phylogenetic analysis (Chapter 2) (Chiesa *et al.* 2013; Meraner *et al.* 2013). This morphometric similarity could arise from occupying similar environments and experiencing comparable selective pressures from the environment (Engelbrecht & Mulder 1999; Langerhans *et al.* 2003).

Parallels may have occurred in the species' morphologies due to increased gene flow or the lack of gene flow between the different river systems (Engelbrecht & Mulder 1999). Increased gene flow may lead to more homogeneous morphologies within a species and similar environmental conditions and exposure to the same predators as the other system may lead to convergent morphologies (Engelbrecht & Mulder 1999; Jackson *et al.* 2001), such as

M. katangae, where most of the samples grouped exception for the Lake Sibaya and Luapula River systems. *Lacustricola myaposae* and *L. macrurus* showed morphometric variation and could be a result of different local habitat conditions or the lack of gene flow between populations due to limited or no connectivity between their river systems.

3.4.2.3. Hypsopanchax

Hypsopanchax jubbi (Poll & Lambert, 1965) was described from near the origin of the Upper Zambezi River in the Mwinilunga District, Zambia (Poll & Lambert 1965). The specimens of *H. jubbi* used in the study were primarily type specimens housed at the Royal Museum for Africa, Belgium. They showed a slight shape difference from the *Hypsopanchax* sp. specimens from the Democratic Republic of Congo, which could not be identified in the field and could possibly be a new species, but this could not be explored in the phylogenetic analysis (Chapter 2). Although the specimens differed in shape, primarily in the depth of the body, it was minimal and may not be enough to distinguish between the two with the naked eye. Live colouration could not be examined due to the loss of colour during storage, which is especially true for the type specimens, which have lost all colouration.

3.5. Conclusion

The morphometric analysis of the genera confirmed that *Aplocheilichthys* was a monotypic genus but this was not seen in the phylogenetic analysis. The species under *Micropanchax* had no clear relationship to *A. spilauchen*. The morphometric and phylogenetic data suggests that *Lacustricola*, which was once a subgenus of *Micropanchax*, may need to have its southern African species reassigned to *Micropanchax*. This is especially true for *L. myaposae* and *L. macrurus*, which showed very close morphometric relationships to *Micropanchax* Clades C and D. The three molecular Clades C, D and E showed that although there is very little morphometric variation to the naked eye, the three clades formed two morphometric groups with Clades C and E relatively more similar. The small differences between the river systems may be in part to low gene flow between populations that may only cross during events of extreme flooding and or through anthropogenic dispersal.

Morphometric analysis of the family Procatopodinae was relevant for this study but future studies should include more measurements. The results of the study showed that body depth, position of the dorsal and anal fin, and the length of the tail where the most important features in distinguishing between species and populations. These suggested that environmental factors played a role in how the phenotype was expressed. Some measurements that could be important to diagnosing the species were excluded due to the physical condition of some of the important specimens, which were missing some structures such as parts of their fins. The specimens stored for any length of time lose their colouration, which appears to be an important identifying tool for the different species and should be included in future studies.

3.6. References

ADAMS, D.C., RHOLF, F.J. & SLICE, D.E. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71(1)**: 5-16.

BEILFUSS, R. & BROWN, C. 2010. Assessing environmental flow requirements and trade-offs for the Lower Zambezi River and Delta, Mozambique. *International Journal of River Basin Management* **8(2)**: 127-138.

BICKFORD, D., LOHMAN, D. J., SODHI, N. S., NG, P. K. L., MEIER, R., WINKER, K., INGRAM, K. K. & DAS, I. 2006. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22**: 148-155.

BLEEKER, P. 1863. *Mémoire sur les poisons de le côte de Guinée.* (Vol. 1). Loosjes.

BOOKSTEIN, F.L., CHERNOFF, B., ELDER, R.L., HUMPHRIES, J.M., SMITH, G.R. & STRAUSS, R.E. 1985. Morphometrics in evolutionary biology: the geometry of size and shape change, with examples from fishes. Special publication 15. *The Academy of Natural Sciences of Philadelphia*.

BOULENGER, G.A. 1904. Descriptions of new West-African freshwater fishes. *Annals and Magazine of Natural History* **14**: 16-20.

BOULENGER. G.A. 1906. Fourth contribution to the ichthyology of Lake Tanganyika. Report on the collection of fishes made by Dr. W. A. Cunnington during the third Tanganyika Expedition, 1904-1905. *Transactions of the Zoological Society of London* **17()**: 537-601.

BOULENGER, G.A. 1908. On a collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum* **1**: 219-235.

BOULENGER, G.A. 1912. Description d'un poisson nouveau du genre *Haplochilus* provenant du Katanga. *Revue de Zoologie Africaine* **2**: 47-48.

BOULENGER, G.A. 1913. Sur une petite collection de poisons recueillis dans l'Uelé par la mission dirigée par M. hutereau. *Revue de Zoologie Africaine* **2**: 155-161.

BRAGANÇA, P.H.N., AMORIM, P.F. & Costa, W.J.E.M. 2018. Pantanodontidae (Teleostei, Cyprinodontiformes), the sister group to all other cyprinodontoid killifishes as inferred by molecular data. *Zoosystematics and Evolution* **94(1)**: 137-145.

BRAGANCA, P.H.N., VAN ZEEVENTER, R.M., BILLS, I.R., TWEDDLE, D. & CHAKONA, A. Diversity of the southern Africa *Lacustricola* Myers, 1924 and redescription of *Lacustricola johnstoni* (Boulenger, 1906) and *Lacustricola myaposae* (Boulenger, 1908) (Cyprinodontiformes: Procatopodidae). *Zookeys*, submitted.

CADRIN, S.X. 2000. Advances in morphometrics identification of fishery stocks. *Reviews in Fish Biology and Fisheries* **10**: 91-112.

CADRIN, S.X. & FRIEDLAND, K.D. 1999 . The utility of image processing techniques for morphometric analysis and stock identification. *Fisheries Research* **43**: 129-139.

CHIESA, S., FILONZI, L., VAGHI, M., PAPA, R. & MARZANO, F.N. 2013. Molecular barcoding of an atypical cyprinid population assessed by Cytochrome b gene sequencing. *Zoological Science* **30**: 408-413.

DÍEZ-DEL-MOLINO, D., GARCÍA-BERTHOU, E., ARAGUAS, R.A., ALCARAZ, C., VIDAL, O., SANZ, N. & GARCÍA-MARÍN, J.L. 2018. Effects of water pollution and river fragmentation on population genetic structure of invasive mosquitofish. *Science of the Total Environment* **637**: 1372-1382.

DOUGLAS, M.E. & MATTHEWS, W.J. 1992. Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage. *Nordic Society Oikos* **65**: 213-224.

ENDLER, J. A. 1982. Convergent and divergent effects of natural selection on colour patterns in two fish faunas. *Evolution* **36**: 178-188.

ENGELBRECHT, G.D. & MULDER, P.F.S. 1999. Extremely high genetic differentiation between two populations of the river goby, *Glossogobius callidus* (Smith, 1937). *Water SA* **25(1)**: 85-90.

FRICKE, R., ESCHMEYER, W.N. & VAN DER LAAN, R. (eds). 2018. Catalog of fishes: genera, species, references (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 12 March 2018.

FROESE, R. & PAULY, D. EDS. 2018. World Wide Web electronic publication (http://www.fishbase.org), version (02/2018). Electronic version accessed 06 July 2018.

GHEDOTTI, M.J. 2000. Phylogenetic analysis and taxonomy of the poecilioid fishes (Teleostei: Cyprinodontiformes). *Zoological Journal of the Linnean Society* **130**: 1-53.

GÜNTHER, A. 1984. Second report on the reptiles, batrachians, and fishes transmitted by Mr. H. H. Johnston, C. B., from British Central Africa. *Proceedings of the Zoological Society of London* **1893** (4): 616-628, Pls. 53-57.

HOSSAIN, M.A.R., NAHIDUZZAMAN, M., SAHA, D., KHANAM, M.U.H. & ALAM, M.S. 2010. Landmark-based morphometric and meristic variations of the endangered carp, *Kalibaus Labeo calbasu*, from stocks of two isolated rivers, the Jamuna and Halda, and a hatchery. *Zoological Studies* **49**(4): 556-563.

HUBER, J. H. 1982. *A review of the cypronodont fauna of the coastal plain in Rio Muni, Gabon, Congo, Cabinda, and Zaire, with taxonomic shifts in* Aphyosemion, Epiplatys and *West African Procatopodins.* British Killifish Association, London.

HURWOOD, D.A, & HUGHS, J.M. 1998. Phylogeography of the freshwater fish, *Mogurnda adspersa*, in streams of northeastern Queensland, Australia: evidence for altered drainage patterns. *Molecular Ecology* **7**: 1507-1517.

INGENITO, L.F. & BUCKUP, P.A. 2005. A new species of *Parodon* from the Serra da Mantiqueira, Brazil (Teleostei: Characiformes: Parodontidae). *Copeia* **2005**(4): 765-771.

ISTEAD, A.E., YAVNO, S. & FOX, M.G. 2015. Morphological change and phenotypic plasticity in response to water velocity in three species of Centrarchidae. *Canadian Journal of Zoology* **93**(11): 879-888.

JACKSON, D.A., PERES-NETO, P.R. & OLDEN, J.D. 2001. What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. *The Canadian Journal of Fisheries and Aquatic Sciences* **58**: 157-170.

KERSCHBAUMER, M. & STURMBAUER, C. 2011. The utility of geometric morphometrics to elucidate pathways of cichlid fish evolution. *International Journal of Evolutionary Biology*: 8 pages.

KOCOVSKY, P.M., ADAMS, J.V. & BRONTE, C.R. 2009. The effect of sample size on the stability of Principal Components Analysis of truss-based fish morphometrics. *Transactions of the American Fisheries Society* **138**: 487-496.

LANGERHANS, R.B., LAYMAN, C.A., LANGERHANS, A.K. & DEWITT, T.J. 2003. Habitat-associated morphological divergence in two Neotropical fish species. *Biological Journal of the Linnean Society* **80**: 689-698.

LUCINDA, P.H. 2003. Family Poeciliidae. *Check list of the freshwater fishes of south and Central America*, pp. 555-581.

LUCINDA, P.H.F. & REIS, R.E. 2005. Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. *Neotropical Ichthyology* **3**: 1-60.

MARSHALL, B. 2011. *The fishes of Zimbabwe and their biology*. Smithiana Monographs. **3**: i-vi + 1-290.

MERANER, A., VENTURI, A., FICETOLA, G.F., ROSSI, S., CANDIOTTO, A. & GANDOLFI, A. 2013. Massive invasion of exotic *Barbus barbus* and introgressive hybridization with endemic *Barbus plebejus* in Northern Italy: where, how and why? *Molecular Ecology* **22**: 5295-5312.

MOJEKWU, T.O. & ANUMUDU, C.I. 2015. Advanced techniques for morphometric analysis in fish. *Journal of Aquaculture Research and Development* **6(8)**: 354-359.

MYERS, G.S. 1924. A new poeciliid fish of the genus *Micropanchax*, from Ubangi. *American Museum Novitates* 122: 1-3.

ØSTBYE, K., TAUGBØL, A., RAVINET, M., HARROD, C., PETTERSEN, R.A., BERNATCHEZ, L. & VØLLESTAD, L.A. 2018. Ongoing niche differentiation under high gene flow in a polymorphic brackish water threespine stickleback (*Gasterosteus aculeatus*) population. *BMC Evolutionary Biology* **18(1)**: 1-14.

PARENTI, L. R. 1981. A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History* **168**: 335-557.

PELICICE, F.M., POMPEU, P.S. & AGOSTINHO, A.A. 2015. Large reservoirs as ecological barriers to downstream movements of Neotropical migratory fish. *Fish and Fisheries* **16**: 697-715.

POLL, M. & LAMBERT, J.G. 1965. Contribution à l'étude systématique et zoogéographique des Procatopodinae de l'Afrique centrale (Pisces, Cyprinodontidae). *Bulletin des séances. Académie royale des sciences dÓutre-Mer. Nouv. sér.* 615-631.

POLL, M. 1967. Contribution à la faune ichthyologique de l'Angola. Publicações Culturais, Companhia de Diamantes de Angola (DIAMANG). *Lisboa* **75**: 1-381, Pls. 1-20.

RIDDEN, M.A., BILLS, I.R. & VILLET, M.H. 2016. Phylogeographic, morphometric and taxonomic re-evaluation of the river sardine, *Mesobola brevianalis* (Boulenger, 1908) (Teleostei, Cyprinidae, Chedrini). *ZooKeys* **641**: 121-150.

SCRIBNER, K.T., PAGE, K.S.& BARTRON, M.L. 2000. Hybridization in freshwater fishes, a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* **10(3)**: 293-323.

SKELTON, P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town.

SKELTON, P.H. 2002. Changes to the scientific and common names of southern African freshwater fishes. *African Journal of Aquatic Science* **27**: 171-174.

STEENKAMP, M.K.J., ENGELBRECHT, G.D. & MULDER, P.F.S. 2001. Allozyme variation in a Johnston's topminnow, *Aplocheilichthys johnstoni*, population from the Zambezi River system. *Water SA* **27**: 53-55.

STRAUSS, R.E. & BOOKSTEIN, F.L. 1982. The truss: body form reconstructions in morphometrics. *Systematic Zoology* **31**: 113-135.

TOBLER, M. & HASTINGS, L. 2011. Convergent patterns of body shape differentiation in four different clades of poeciliid fishes inhabiting sulphide springs. *Evolutionary Biology* **38**: 412-421.

TURAN, C. 1999. A note on the examination of morphometric differentiation among fish populations: the truss system. *Turkish Journal of Zoology* **23**: 259-263.

WARREN, JR, M.L. & PARDEW, M.G. 1998. Road crossings as barriers to small-stream fish movement. *Transactions of the American Fisheries Society* **127**: 637-644.

WELSH, D.P., ZHOU, M., MUSSMANN, S.M., FIELDS, L.G., THOMAS, C.L., PEARISH, S.P., KILBURN, S.L., PARKER, J.L., STEIN, L.R., BARTLETT, J.A., BERTRAM, C.R., BLAND, T.J., LASKOWSKI, K.L., MOMMER, B.C., ZHUANG, X. & FULLER, R.C. 2013. The effects of age, sex, and habitat on body size and shape of the blackstripe topminnow, *Fundulus notatus* (Cyprinodontiformes: Fundulidae) (Rafinesque 1820). *Biological Journal of the Linnean Society* **108**: 784-789.

WESSELS, G. 2009. *Meristic and morphometric variation among sardine Sardinops sagax around the coast of southern Africa.* M.Sc thesis, University of Cape Town, Cape Town, South Africa.

WILL, K.W. & RUBINOFF, D. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**: 47-55.

WOLD, S., ESBENSEN, K. & GELADI, P. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* **2**(1-3): 37-52.

Table 3.1. Sample catalogue numbers with relevant locality information and type specimens represented in bold.

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
Aplocheilichthys spilauchen						
84605	Angola	Kwanza River	Kwanza System	-9.69	14.42	10
84665	Angola	Kwanza River	Kwanza System	-9.69	14.42	6
84696	Angola	Lucala River	Kwanza System	-9.52	14.39	10
75242	Benin	Aho River	Lake Ahémé	6.34	2.09	10
Micropanchax johnstoni						
99047	Zambia	Lulimala River	Congo System	-12.29	30.62	10
200812	Namibia	Kwando River	Kwando System	-17.79	23.34	10
200799	Namibia	Kwando River	Kwando System	-	-	7
200808	Namibia	Kwando River	Kwando System	-17.88	23.32	10
76838	Zambia	Lake Bangweulu	Lake Bangweulu	-11.36	29.56	6
76854	Zambia	Lake Bangweulu	Lake Bangweulu	-11.37	29.56	7
78549	South Africa	Mokolo River	Limpopo System	-24.11	27.80	5
99212	South Africa	Nzhelele River	Limpopo System	-22.42	30.28	2
76618	Zambia	Luapula River	Luapula System	-12.06	29.67	10
76684	Zambia	Luapula River	Luapula System	-10.54	28.66	10
65263	South Africa	-	Mhlatuze System	-28.77	32.08	4
27805	South Africa	Mkuze Swamps	Mkuze System	-27.59	32.41	10
190590	Angola	Cuito River	Okavango System	-16.62	19.05	10
186834	Angola	Cuito River	Okavango System	-15.17	19.19	10
186878	Angola	Cuito River	Okavango System	-15.14	19.19	10
101088	Angola	Kwando River	Okavango System	-17.47	23.08	10
187043	Botswana	Lake Ngami	Okavango System	-20.36	22.95	10
66443	Botswana	Okavango River	Okavango System	-18.41	21.89	1
200801	Namibia	Okavango River	Okavango System	-	-	1
200784	Namibia	Okavango River	Okavango System	-	-	10

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
200771	Namibia	Okavango River	Okavango System	-	-	10
186644	Angola	Quebe River	Okavango System	-14.94	17.72	10
66752	Botswana	Thamalakane River	Okavango System	-18.27	21.80	9
66750	Botswana	Thamalakane River	Okavango System	-	-	10
66655	Botswana	Thaoge River	Okavango System	-18.88	22.39	10
34384	Malawi	Shire River	Shire System	-15.07	35.22	1
41623	Malawi	Shire River	Shire System	-15.07	35.13	5
BMNH 1893.11.15.92-99	**Malawi**	**Shire River**	**Shire System**	**-**	**-**	**7**
35820	Malawi	Upper Shire River	Shire System	-14.45	33.27	10
71283	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	10
190409	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	4
193611	Zambia	Lui River	Upper Zambezi System	-16.22	23.34	10
72078	Zambia	Zambezi River	Upper Zambezi System	-16.15	23.29	5
193587	Zambia	Munde Stream	Zambezi System	-14.69	22.62	9
200802	Namibia	Zambezi River	Zambezi System	-17.56	24.46	10
200813	Namibia	Zambezi River	Zambezi System	-17.56	24.46	3
81254	Mozambique	Zambezi River	Zambezi System	-15.75	31.61	10
81036	Mozambique	Zambezi River	Zambezi System	-	-	10
Micropanchax katangae						
88640	South Africa	Nkanini River	Kosi Bay	-26.95	32.78	1
97285	Democratic Republic of Congo	Karivia River	Congo System	-11.65	27.29	5
200798	Namibia	Kwando River	Kwando System	-17.87	23.31	5
200809	Namibia	Kwando River	Kwando System	-17.88	23.32	5
88657	South Africa	Mseleni River	Lake Sibaya	-27.36	32.53	10
64724	Mozambique	Lake Xingute	Lake Xingute	-26.51	32.81	10
81526	Democratic Republic of Congo	Kisayu River	Luapula System	-10.71	25.93	8
81549	Democratic Republic of Congo	Lake Nzilo	Luapula System	-10.77	25.74	12

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
BMNH 1912.7.10.3-4	**Democratic Republic of Congo**	**Lumbumbashi River**	**Luapula System**	**-**	**-**	**2**
66793	Botswana	Boro River	Okavango System	-19.54	23.06	10
97257	Democratic Republic of Congo	Cuito River	Okavango System	-11.65	27.36	3
101097	Angola	Cuito River	Okavango System	-17.51	20.07	10
101073	Angola	Kwando River	Okavango System	-17.53	23.19	10
18878	Botswana	Okavango River	Okavango System	-20.00	23.33	1
66932	Bostwana	Okavango River	Okavango System	-18.74	22.20	6
200787	Namibia	Okavango River	Okavango System	-	-	2
186731	Angola	Tributary Chingueia	Okavango System	-13.69	17.06	10
86625	South Africa	Hluhluwe River	St. Lucia	-28.14	32.31	1
83158	South Africa	St. Lucia	St. Lucia	-27.63	32.56	1
190390	Namibia	Bukalo Channel	Upper Zambezi System	-17.72	24.52	6
73493	Zambia	Chitunta River	Upper Zambezi System	-	-	10
190404	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	5
193693	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	5
72904	Zambia	Litoya River	Upper Zambezi System	-11.50	24.38	10
97941	Zambia	Chisolo River	Zambezi System	-12.04	25.44	10
140191	Zambia	Chisolo River	Zambezi System	-12.09	25.33	10
189975	Namibia	Kwando River	Zambezi System	-	-	10
189976	Namibia	Kwando River	Zambezi System	-	-	10
140155	Zambia	Shinengeni Stream	Zambezi System	-12.36	25.33	10
60854	Mozambique	Zambezi River	Zambezi System	-18.50	35.65	3
200804	Namibia	Zambezi River	Zambezi System	-17.56	24.46	7
Micropanchax hutereaui						
99078	Zambia	Congo River	Congo System	-12.01	30.42	10
99038	Zambia	Lulimala River	Congo System	-12.36	30.61	10
99046	Zambia	Lulimala River	Congo System	-12.29	30.62	10

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
78292	Central African Republic	Ouaka River	Congo System	6.16	20.80	1
78313	Central African Republic	Ouaka River	Congo System	6.21	20.71	7
78314	Central African Republic	Ouaka River	Congo System	6.16	20.71	15
78317	Central African Republic	Ouaka River	Congo System	6.21	20.69	10
78322	Central African Republic	Ouaka River	Congo System	6.06	20.81	7
BMNH 1912.12.6.14	**Democratic Republic of Congo**	**Uelé River**	**Congo System**	**-**	**-**	**1**
200797	Namibia	Kwando River	Kwando System	-17.87	23.31	10
200810	Namibia	Kwando River	Kwando System	-17.88	23.32	3
101178	Angola	Cuito River	Okavango System	-16.62	19.05	5
101071	Angola	Kwando River	Okavango System	-17.53	23.19	6
101089	Angola	Kwando River	Okavango System	-17.47	23.08	5
190631	Angola	Kwando River	Okavango System	-17.57	23.26	9
66784	Botswana	Maunachira River	Okavango System	-19.20	23.46	10
66934	Botswana	Okavango River	Okavango System	-18.74	22.20	2
200774	Namibia	Okavango River	Okavango System	-	-	6
200781	Namibia	Okavango River	Okavango System	-	-	1
200786	Namibia	Okavango River	Okavango System	-	-	8
200814	Namibia	Okavango River	Okavango System	-17.56	24.46	4
66572	Botswana	Thaoge River	Okavango System	-18.96	22.39	1
66589	Botswana	Thaoge River	Okavango System	-18.96	22.38	1
66607	Botswana	Thaoge River	Okavango System	-18.96	22.40	7
71688	Zambia	Kaande Canal	Upper Zambezi System	-15.24	23.24	5
190406	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	1
71262	Zambia	Litoya River	Upper Zambezi System	-15.68	23.30	2
193609	Zambia	Lui River	Upper Zambezi System	-16.22	23.34	1
71199	Zambia	Zambezi River	Upper Zambezi System	-16.15	23.29	5
72864	Zambia	Zambezi River	Upper Zambezi System	-16.19	23.29	5
73455	Zambia	Zambezi River	Upper Zambezi System	-12.24	25.57	5

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
193548	Zambia	Kabompo River	Zambezi System	-12.45	24.94	6
193622	Zambia	Luanginga River	Zambezi System	-14.54	22.25	2
193673	Zambia	Lungwebungu River	Zambezi System	-14.00	22.65	2
193588	Zambia	Munde Stream	Zambezi System	-14.69	22.62	5
60855	Mozambique	Zambezi River	Zambezi System	-18.50	35.65	3
60860	Mozambique	Zambezi River	Zambezi System	-18.55	35.64	10
200803	Namibia	Zambezi River	Zambezi System	-17.56	24.46	10
Micropanchax baudoni						
AMNH 20936	**Central African Republic**	**Gribingi River**	**Shari System**	**-**	**-**	**1**
Micropanchax sp.						
97528	Democratic Republic of Congo	Niari River	Congo System	-4.08	12.14	3
74020	Mozambique	Lugenda River	Rovuma System	-12.45	37.60	2
190398	Namibia	Caprivi River	Caprivi System	-17.72	24.52	1
135903	Democratic Republic of Congo	-	Congo System	-12.75	28.55	2
193697	Zambia	Kataba River	Upper Zambezi System	-15.57	23.28	8
189970	Namibia	Kwando River	Zambezi System	-	-	10
88671	South Africa	-	-	-	-	10
Lacustricola macrurus						
28096	Namibia	Cunene River	Cunene System	-17.40	14.20	7
63280	Namibia	Cunene River	Cunene System	-17.25	12.44	10
28126	Namibia	Cunene River	Cunene System	-17.00	13.25	3
84794	Angola	Kwanza River	Kwanza System	-9.81	15.41	10
Lacustricola myaposae						
88658	South Africa	Mseleni River	Lake Sibaya	-27.36	32.53	10

Catalogue Number	Country	River	River system	Latitude	Longitude	Sample size
64798	Mozambique	Lake Piti	Lake Piti	-26.63	32.88	10
64777	Mozambique	Lake Piti	Lake Piti	-26.62	32.89	10
96619	South Africa	Mhlathuze River	Mhlatuze System	-28.84	31.91	10
64751	Mozambique	Lake Piti	Lake Piti	-26.54	32.90	10
96591	South Africa	Nseleni River	Coastal	-28.67	31.96	9
Lacustricola matthesi						
37315	Zambia	Saisi River	Lake Tanganyika	-9.08	31.53	10
Lacustricola mediolateralis						
162484 (R. G. Mus. Afr. centr.)	**Angola**	**Cuango-Muquehe River**	**Congo System**	**-**	**-**	**2**
162489-494 (R. G. Mus. Afr. centr.)	**Angola**	**Cuango-Muquehe River**	**Congo System**	**-10.85**	**19.3**	**2**
162488 (R. G. Mus. Afr. centr.)	**Angola**	**Cuango-Muquehe River**	**Congo System**	**-10.88**	**19.019**	**1**
Lacustricola pumilus						
135325	Tanzania	Lake Tanganyika	Lake Tanganyika	-9.40	31.58	3
Lacustricola sp.						
85070	Angola	Kwanza River	Kwanza System	-9.69	15.00	10
Hypsopanchax jubbi						
148115 (R. G. Mus. Afr. centr.)	**Zambia**	**Zambezi River**	**Upper Zambezi System**	**-11.12**	**24.10**	**1**
142129-756 (R. G. Mus. Afr. centr.)	**Zambia**	**Zambezi River**	**Upper Zambezi System**	**-11.12**	**24.10**	**7**
***Hypsopanchax* sp.**						
81581	Democratic Republic of Congo	Kii River	Lualaba System	-10.62	25.94	10
81582	Democratic Republic of Congo	Kii River	Lualaba System	-10.62	25.94	2

Table 3.2. Measurements used for the box truss and additional measurements used for morphometric analysis.

Measurement	Landmarks	Description
L1	A-B	Head length
L2	B-C	Dorsal body length
L3	C-D	Dorsal fin width
L4	D-E	Dorsal caudal length
L5	E-F	Caudal depth
L6	F-G	Anterior caudal length
L7	G-H	Anal fin width
L8	H-I	Anterior body length
L9	I-J	Pectoral fin to anterior head section
L10	A-J	Anterior head length
L11	A-I	Head depth
L12	B-I	Head depth
L13	C-H	Body depth
L14	D-G	Caudal depth
L15	B-J	Head depth
L16	B-H	Body diagonal
L17	C-G	Diagonal before caudal region
L18	D-F	Caudal diagonal
L19	C-I	Body diagonal
L20	D-H	Diagonal before caudal region
L21	E-G	Caudal diagonal
L22		Caudal length
L23		Eye diameter
L24	A to pectoral fin	Tip of snout to pectoral fin
L25		Anal fin length
L26		Total length
L27		Length of middle caudal fin ray

Table 3.3. The classification matrix of the DFA for the variation between genera.

Genera	**Genera**				
	Percent	***Micropanchax***	***Lacustricola***	***Hypsopanchax***	***Aplocheilichthys***
Micropanchax	96,5517	728	16	5	5
Lacustricola	63,2479	43	74	0	0
Hypsopanchax	85,0000	3	0	17	0
Aplocheilichthys	100,0000	0	0	0	36
Total	92,2330	774	90	22	41

Table 3.4. The classification matrix of the DFA for the variation between species.

Genera	**Genera**																			
	Percent	***M. katangae***	***M. johnstoni***	**Clade C**	***M. hutereaui***	**Clade D**	***L. myaposae***	***L. mediolateralis***	***H. jubbi***	***A. Spilauchen***	***Micropanchax* sp.**	***L. pumilus***	***Lacustricola* sp.**	***L. macrurus***	**Hypsopanchax sp.**	**Clade C**	**L. matthesi**	**Clade E**	**Clade A**	**M. baudoni**
M. katangae	76,59	157	11	0	24	2	1	1	0	0	0	0	0	0	5	3	1	0	0	0
M. johnstoni	84,65	10	193	4	7	2	9	0	0	0	0	0	0	2	0	1	0	0	0	0
Clade C	30,95	0	28	13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
M. hutereaui	77,16	29	9	0	152	0	0	0	0	0	3	0	0	0	1	0	0	0	3	0
Clade D	60	0	3	0	0	6	0	1	0	0	0	0	0	0	0	0	0	0	0	0
L. myaposae	76,27	3	11	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0
L. mediolateralis	80	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
H. jubbi	100	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0
A. spilauchen	100	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	0	0
***Micropanchax* sp.**	56,25	2	0	0	0	0	0	0	0	4	9	0	0	1	0	0	0	0	0	0
L. pumilus	100	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
***Lacustricola* sp.**	40	0	2	0	0	0	0	0	0	0	0	0	4	4	0	0	0	0	0	0
L. macrurus	86,67	1	1	0	0	0	2	0	0	0	0	0	0	26	0	0	0	0	0	0
***Hypsopanchax* sp.**	83,33	2	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
***Micropanchax* sp. Clade C**	20	1	14	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0
L. matthesi	40	5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4	0	0	0
Clade E	5	1	17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Clade A	50	0	0	0	6	0	0	0	0	0	0	0	0	1	0	0	0	0	7	0
M. baudoni	50	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Total	73,68	211	290	19	190	10	58	6	8	40	13	3	4	34	16	8	5	1	10	1

Table 3.5. Eigenvalues and Eigenvectors of the first three principle components of a PCA of the morphometric measurements for genus-level. Coefficients in **bold** indicate characters that are particularly influential in the relevant Eigenvector.

Variable	Eigenvector		
	PC1	**PC2**	**PC3**
1 A-B	-0,201085	0,006532	0,036974
2 B-C	-0,208316	0,150078	-0,140049
3 C-D	-0,148216	**-0,517975**	-0,006864
4 D-E	-0,168668	**0,434424**	-0,141134
5 E-F	-0,208700	-0,148133	0,138227
6 F-G	-0,188078	**0,215169**	**-0,377474**
7 G-H	-0,179168	**-0,270551**	**-0,482717**
8 H-I	-0,199782	**0,200872**	0,111785
9 I-J	-0,199950	0,082339	**0,247259**
10 A-J	-0,207685	0,069096	0,173523
11 A-I	-0,214856	0,063817	0,188565
12 B-I	-0,211521	-0,073867	**0,243750**
13 C-H	-0,207644	-0,195625	-0,191788
14 D-G	-0,202262	-0,184485	0,172502
15 B-J	-0,207654	-0,100681	**0,203589**
16 C-I	-0,210263	0,135354	0,159045
17 D-H	-0,198503	**-0,230765**	**0,254618**
18 E-G	-0,204066	0,199472	-0,073303
19 B-H	-0,215023	0,055410	-0,145368
20 C-G	-0,196497	**-0,288452**	**-0,320314**
21 D-F	-0,203637	0,117728	-0,150658
22 Caudal length	-0,198955	0,054210	-0,109572
23 Eye diameter	-0,202886	0,113057	0,060536
24 Head length	-0,213581	0,067143	0,007322
25 Anal fin length	-0,188978	-0,058088	0,009402
Eigenvalue	20,32565	1,58781	0,67926
Cumulative variance	81,30%	87,65%	90,37%

Table 3.6. Eigenvalues and Eigenvectors of the first three roots of a DFA of the morphometric measurements for the genus-level. Measurements A-B, A-J, A-I, B-I and C-I were excluded for showing practically no effect on body shape difference in the PCA (Table 3.3).

Variable	**Eigenvector**		
	Root 1	**Root 2**	**Root 3**
L2 B-C	**0,84810**	-0,57830	0,14973
L3 C-D	0,66040	0,81160	**1,22167**
L4 D-E	-0,47075	**1,67614**	0,68138
L5 E-F	-0,73672	-0,84551	**-1,54636**
L6 F-G	-0,06450	**2,44051**	**-2,37609**
L7 G-H	-0,16898	-0,44126	-0,68237
L8 H-I	0,47125	**-1,37233**	-0,56763
L9 I-J	**-0,98555**	0,09029	0,12022
L13 C-H	-0,00057	**2,06367**	0,48371
L14 D-G	-0,50342	-0,95923	0,74116
L15 B-J	**1,24332**	**1,44590**	0,82857
L16 B-H	-0,49519	**2,43097**	0,43576
L17 C-G	-0,15188	-0,11826	**-1,07957**
L18 D-F	-0,50668	**-1,13847**	0,73679
L20 D-H	**-1,48065**	0,00141	0,29161
L21 E-G	0,64110	**-3,25139**	0,53138
L22 Caudal length	0,37325	**-1,65322**	0,57934
L23 Eye diameter	-0,35360	-0,01428	**1,40380**
L24 Head length	0,51979	-0,45003	-0,67742
L25 Anal fin length	0,58457	0,64989	**-0,97596**
Eigenvalue	0,99734	0,41853	0,16704
Cumulative Variance	63.01%	89,45%	100%

Table 3.7. Eigenvalues and Eigenvectors of the first three roots of a DFA of the morphometric measurements for the species-level. Measurements A-B, A-J, A-I and B-I were excluded for showing practically no effect on body shape difference in the PCA (Table 3.5).

Variable	**Eigenvector**		
	Root 1	**Root 2**	**Root 3**
L2 B-C	-0,44743	-0,63934	**-1,70165**
L3 C-D	-0,40853	**-0,77517**	0,72001
L4 D-E	0,06010	0,33047	0,67913
L5 E-F	0,47881	0,68710	**-1,91712**
L6 F-G	**0,70355**	-0,42432	0,45582
L7 G-H	**1,73121**	-0,05774	0,12896
L8 H-I	0,41882	-0,01023	**-1,78636**
L9 I-J	0,29745	**0,87146**	0,27664
L13 C-H	**-0,73986**	-0,44878	**0,80149**
L14 D-G	0,09017	**1,02819**	-0,32209
L15 B-J	**-1,10853**	**-1,12897**	**0,92854**
L16 B-H	-0,11526	0,13163	**2,66903**
L17 C-G	**-0,87703**	0,34965	0,32834
L18 D-F	-0,25509	0,50645	**-1,02720**
L19 C-I	0,40995	-0,42503	**1,64817**
L20 D-H	**-0,73767**	**1,69969**	-0,31560
L21 E-G	0,15896	-0,21913	-0,66497
L22 Caudal length	0,55872	-0,25482	-0,31520
L23 Eye diameter	**0,78602**	-0,02390	-0,09938
L24 Head length	-0,12678	-0,39218	-0,03293
L25 Anal fin length	0,02056	-0,36913	0,37260
Eigenvalue	2,85343	1,11966	0,80714
Cumulative variance	46,69%	65,01%	78,22%

Table 3.7. The classification matrix of the River systems for the DFA for each species group.

A) The classification matrix of the *M. johnstoni* river systems.

River Systems	**River Systems**										
	Percent	**Mhlatuze System**	**Limpopo System**	**Zambezi System**	**Okavango System**	**Mkuze swamps**	**Shire System**	**Upper Zambezi System**	**Lake Bangweulu System**	**Luapula System**	**Congo System**	**Kwando System**
Mhlatuze System	100	4	0	0	0	0	0	0	0	0	0	0
Limpopo System	57,1429	1	4	0	0	1	0	1	0	0	0	0
Zambezi System	55,814	0	1	24	13	0	0	3	0	0	0	2
Okavango System	82,8829	0	0	5	92	4	1	3	1	1	0	4
Mkuze swamps	60	0	3	0	1	6	0	0	0	0	0	0
Shire System	81,8182	0	0	0	2	1	18	0	1	0	0	0
Upper Zambezi System	89,6552	0	0	1	1	0	0	26	1	0	0	0
Lake Bangweulu System	61,5385	0	0	1	4	0	0	0	8	0	0	0
Luapula System	50	0	0	2	8	0	0	0	0	10	0	0
Congo System	60	0	0	0	4	0	0	0	0	0	6	0
Kwando System	74,0741	0	0	0	7	0	0	0	0	0	0	20
Total	73,6487	5	8	33	132	12	19	33	11	11	6	26

B) The classification matrix of the *M. hutereaui*, *M. baudoni* and *Micropanchax* sp.

Group	**Species**			
	Percent	***M. hutereaui***	***M. baudoni***	***Micropanchax* sp. (Pygmy)**
M. hutereaui	99,5261	210	0	1
M. baudoni	100,0000	0	2	0
***Micropanchax* sp. (Pygmy)**	100,0000	0	0	8
Total	99,5475	210	2	9

C) The classification matrix of the *M. hutereaui* river systems.

River System	**River System**						
	Percent	**Congo System**	**Upper Zambezi System**	**Okavango System**	**Kwando System**	**Zambezi System**	**Chad basin**
Congo System	97,1831	69	0	0	0	2	0
Upper Zambezi System	71,8750	3	23	3	0	3	0
Okavango System	64,4444	5	2	29	3	6	0
Kwando System	69,6970	2	0	6	23	2	0
Zambezi System	47,3684	1	3	13	3	18	0
Chad basin	100,0000	0	0	0	0	0	2
Total	74,2082	80	28	51	29	31	2

D) The classification matrix of the *M. katangae* river systems.

River Systems	**River Systems**									
	Percent	**Luapula System**	**St Lucia**	**Okavango System**	**Congo System**	**Lake Sibaya**	**Zambezi System**	**Upper Zambezi System**	**Kwando System**	**Lake Xingute**
Luapula System	90,90909	20	0	1	0	0	0	0	1	0
St Lucia System	50	0	1	0	0	0	1	0	0	0
Okavango System	63,46154	0	1	33	0	1	10	4	2	1
Congo System	80	0	0	0	4	0	1	0	0	0
Lake Sibaya	90	0	1	0	0	9	0	0	0	0
Zambezi System	75,4386	0	0	6	0	0	43	6	1	1
Upper Zambezi System	66,66666	0	0	2	0	0	9	24	0	1
Kwando System	80	0	0	0	0	0	2	0	8	0
Lake Xingute	50	0	0	0	0	0	4	1	0	5
Total	72,05882	20	3	42	4	10	70	35	12	8

E) The classification matrix of the *L. myaposae* river systems.

River Systems	**River Systems**				
	Percent	**Lake Sibaya**	**Lake Piti System**	**Mhlathuze System**	**Coastal**
Lake Sibaya	100,0000	10	0	0	0
Lake Piti System	100,0000	0	30	0	0
Mhlathuze System	100,0000	0	0	10	0
Coastal	100,0000	0	0	0	9
Total	100,0000	10	30	10	9

F) The classification matrix of the *L. macrurus* river systems.

River Systems	**River Systems**		
	Percent	**Cunene System**	**Kwanza System**
Cunene System	100,0000	20	0
Kwanza System	100,0000	0	20
Total	100,0000	20	20

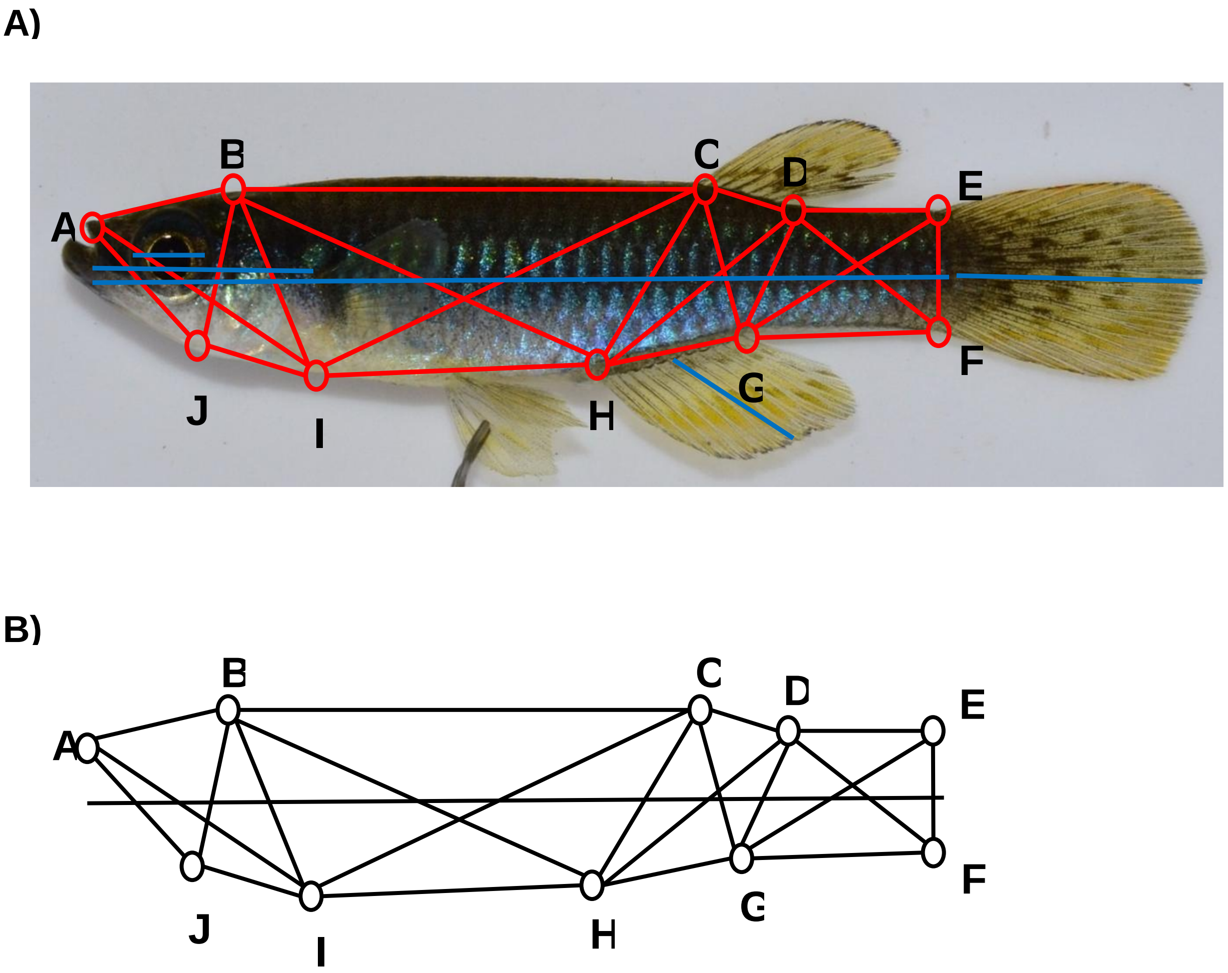


Fig. 3.1. (A) Ten landmarks defined on the body (red), with the eye and body measurements made (blue), (B) the Truss network measurements and caudal length created using the ten landmarks on the fish body.

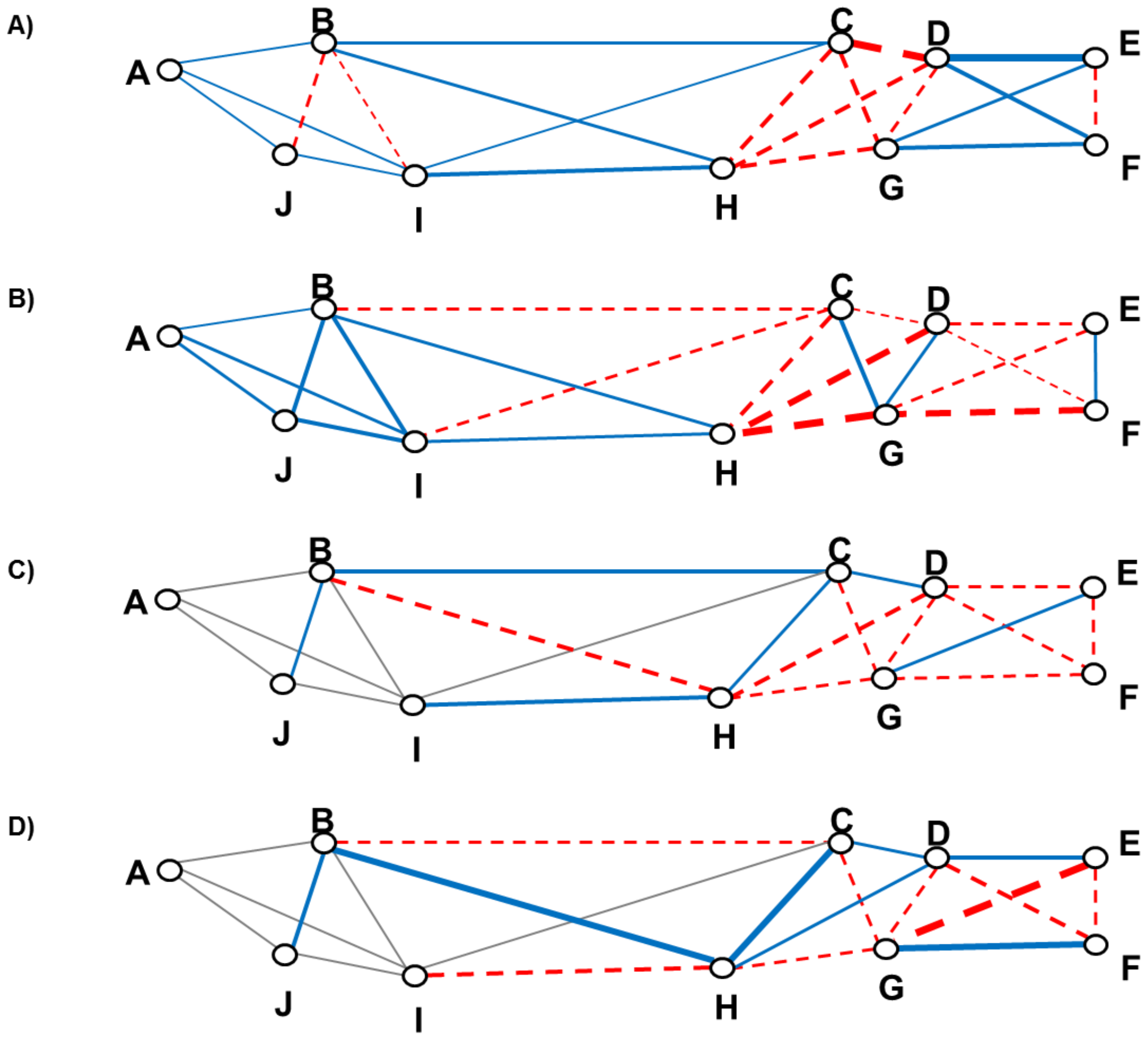


Fig. 3.2. Variation between genera. **A)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the second Eigenvector of the PCA, which summarised variation in body shape; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **B)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the third Eigenvector of the PCA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **C)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the first Eigenvector of the DFA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **D)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the second Eigenvector of the DFA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines.

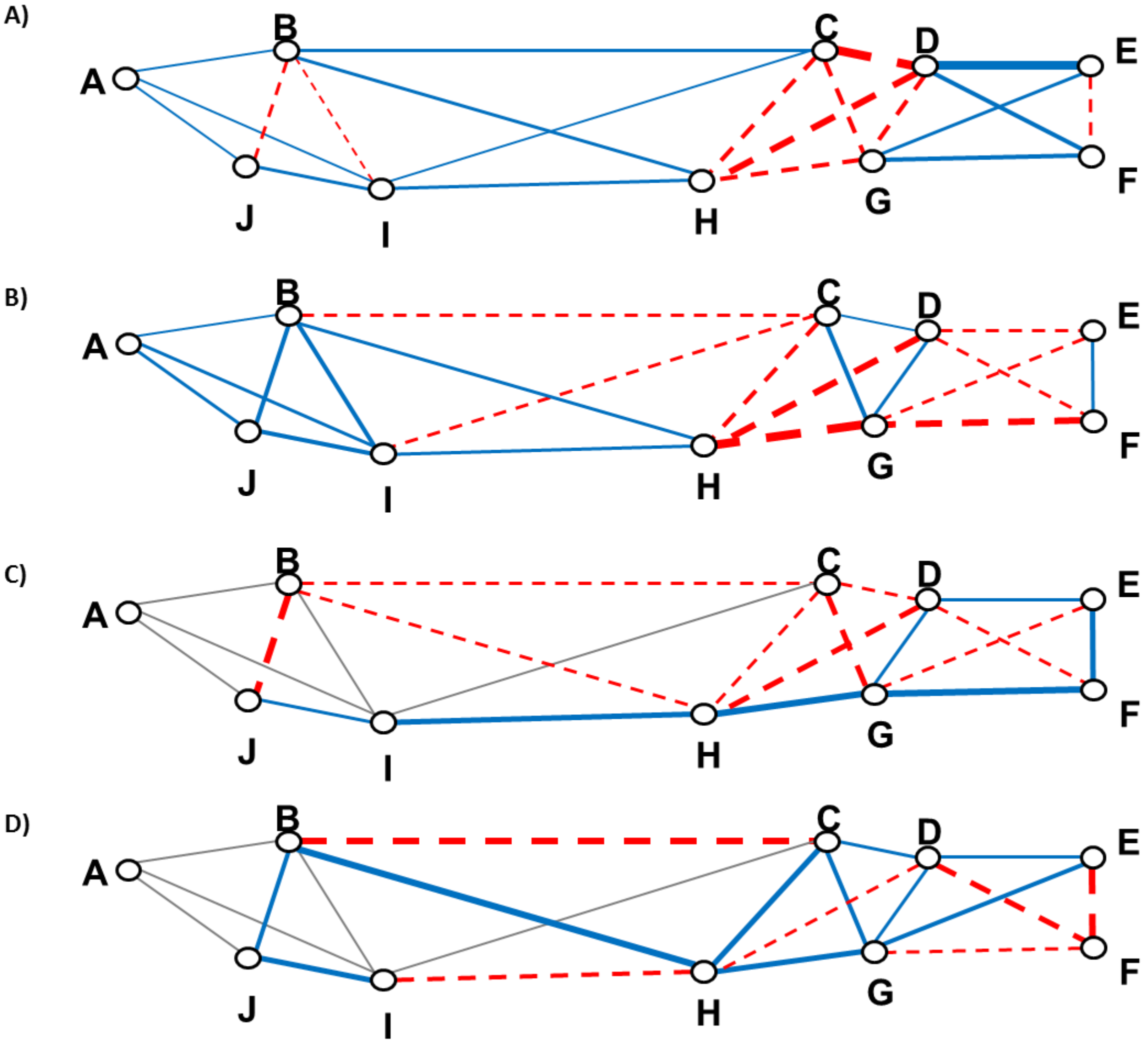


Fig. 3.3. Variation within species **A)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the second Eigenvector of the PCA, which summarised variation in body shape; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **B)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the third Eigenvector of the PCA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **C)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the first Eigenvector of the DFA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines. **D)** The box truss network plotted with the thickness of the truss edges in proportion to the magnitude of their respective factor scores in the second Eigenvector of the DFA; positive factor scores are represented by blue, solid lines and negative factor scores by red, dotted lines.

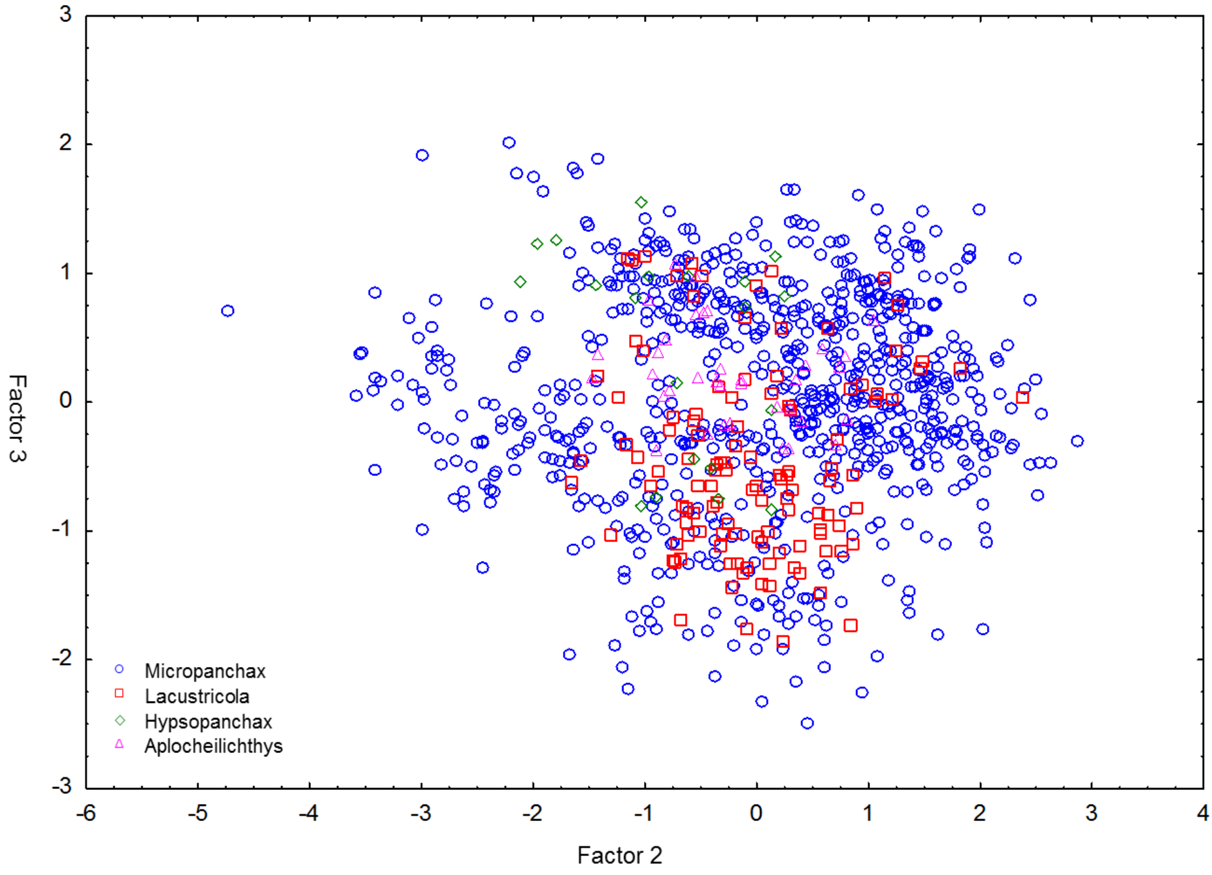


Fig. 3.4. Principal component two plotted against principal component three of the principal component analysis plotted individually for genus-level variation.

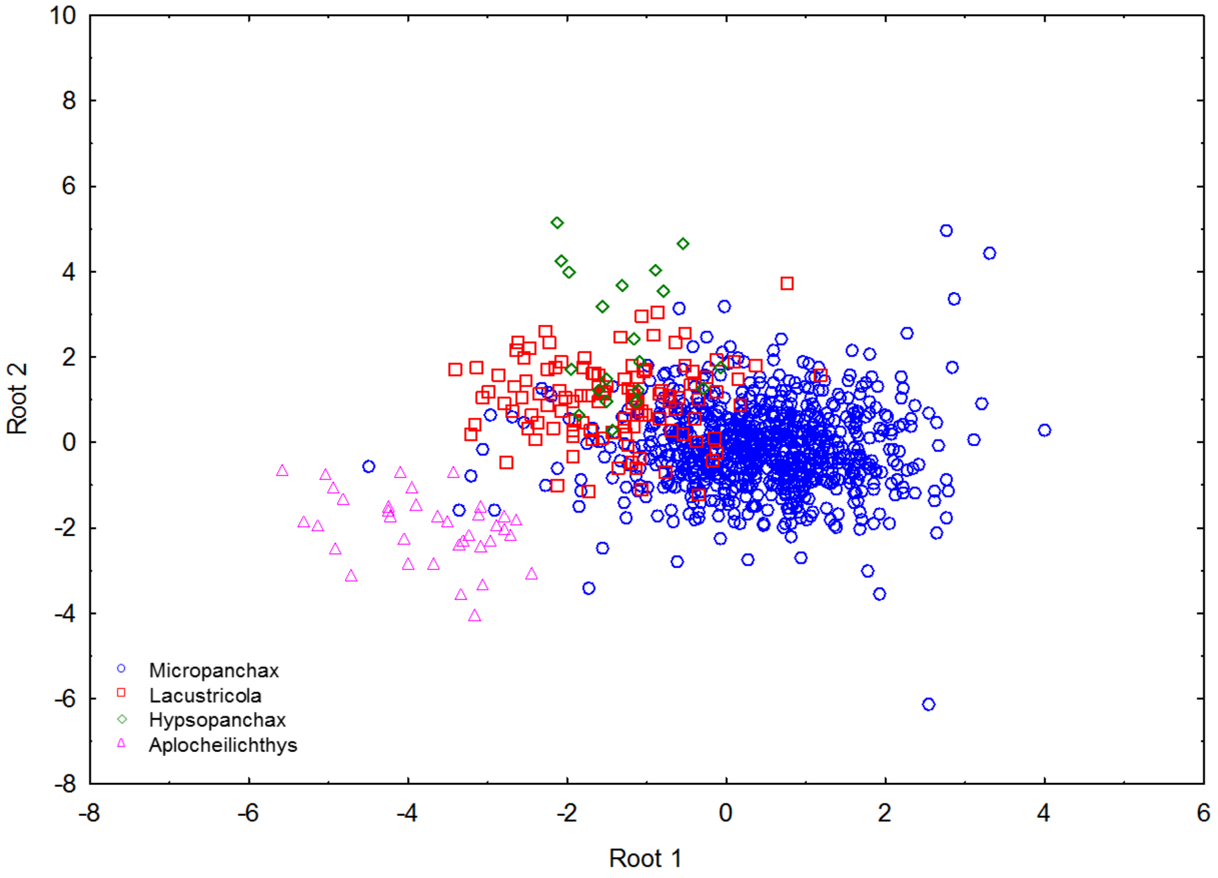


Fig. 3.5. Canonical variate root one plotted against root two of the canonical variate analysis plotted individually for genus-level variation.

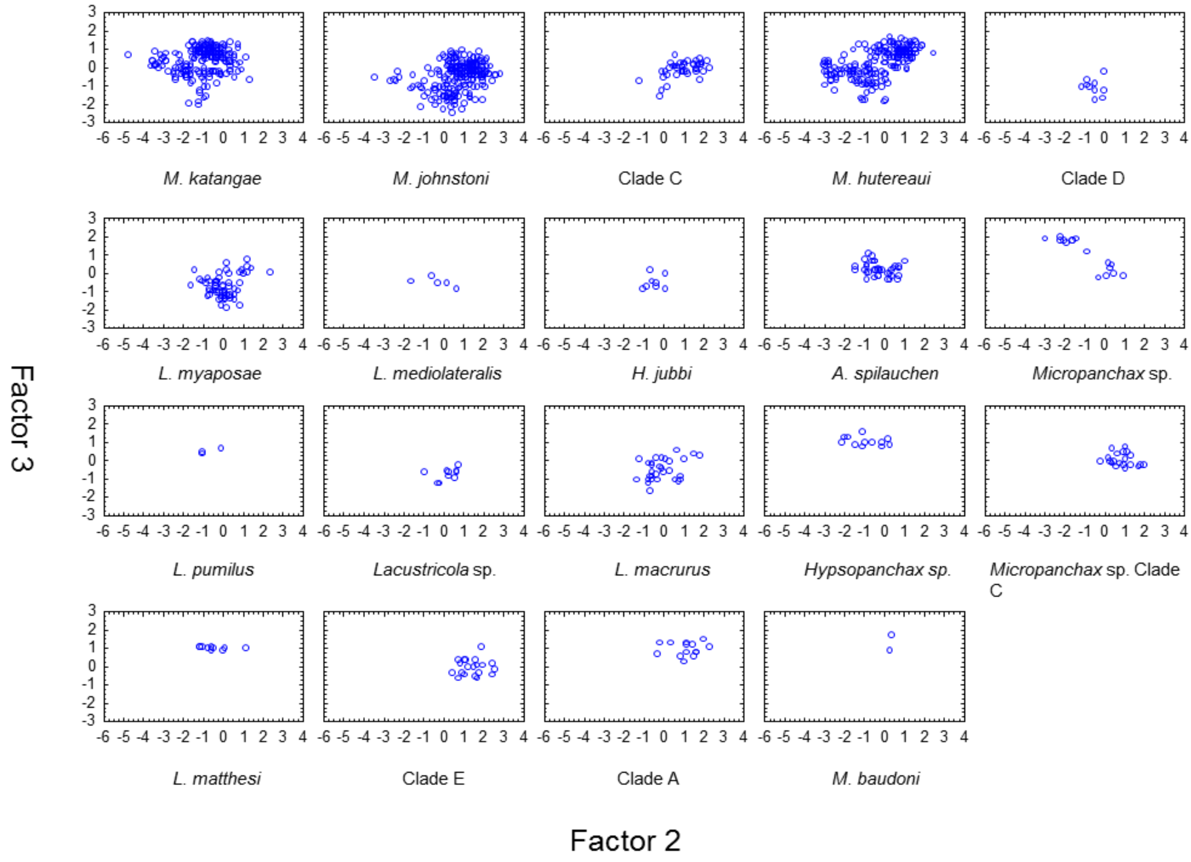


Fig. 3.6. Principal component two plotted against principal component three of the principal component analysis plotted individually for species-level variation.

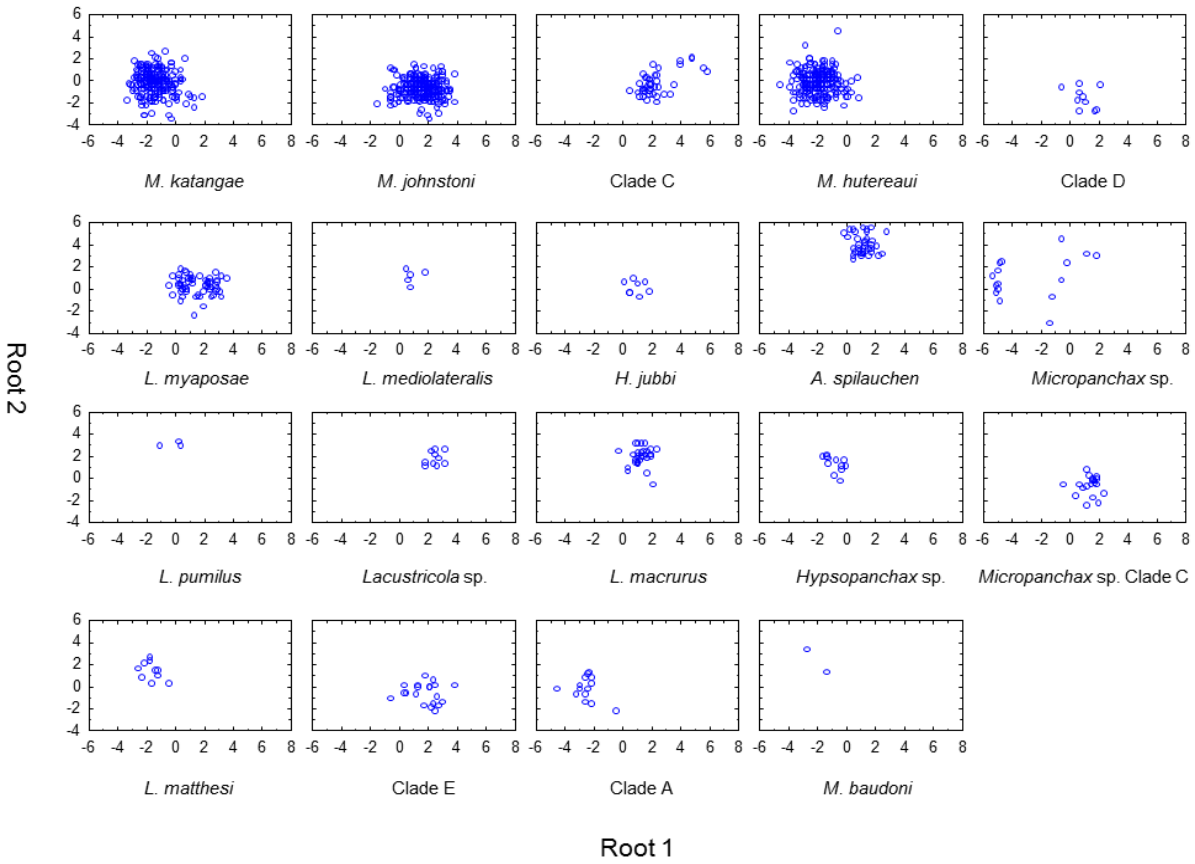
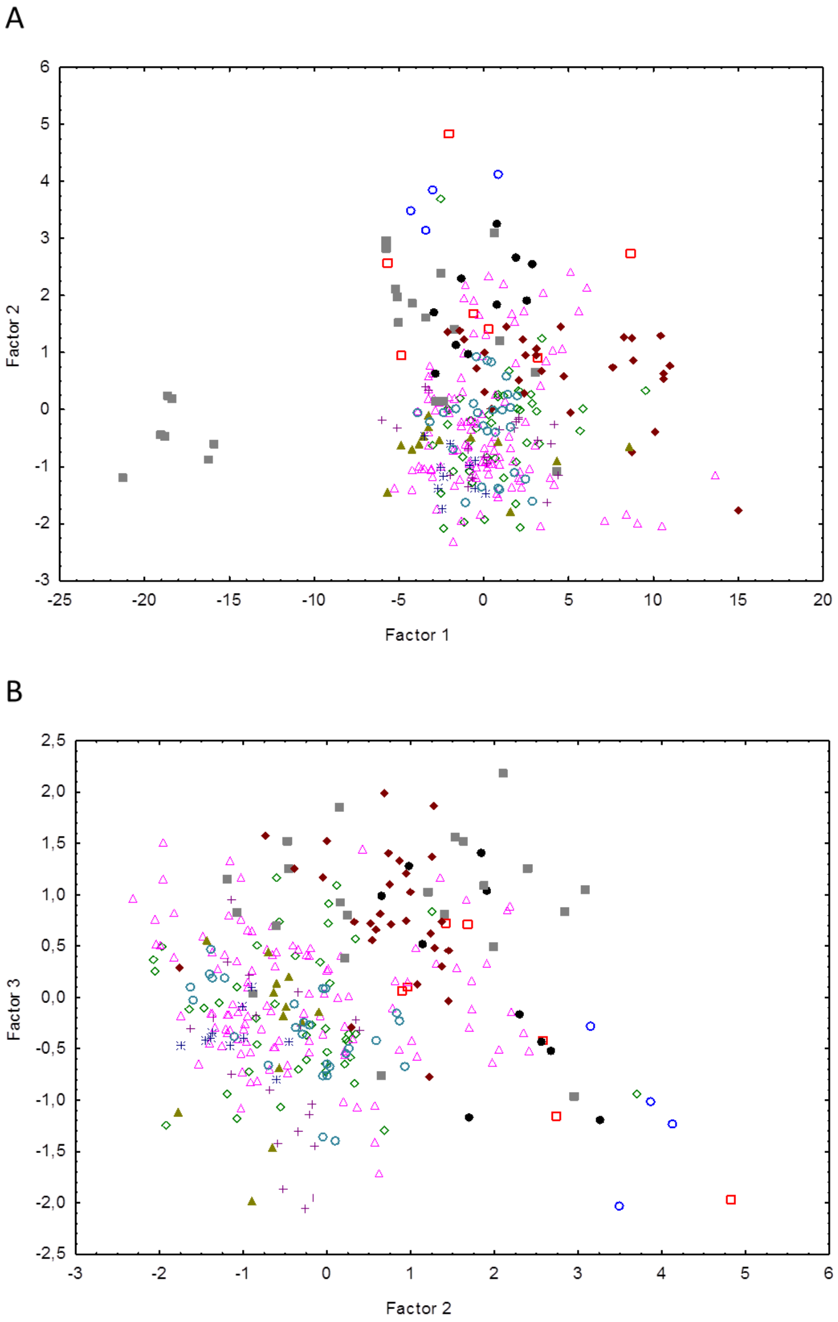
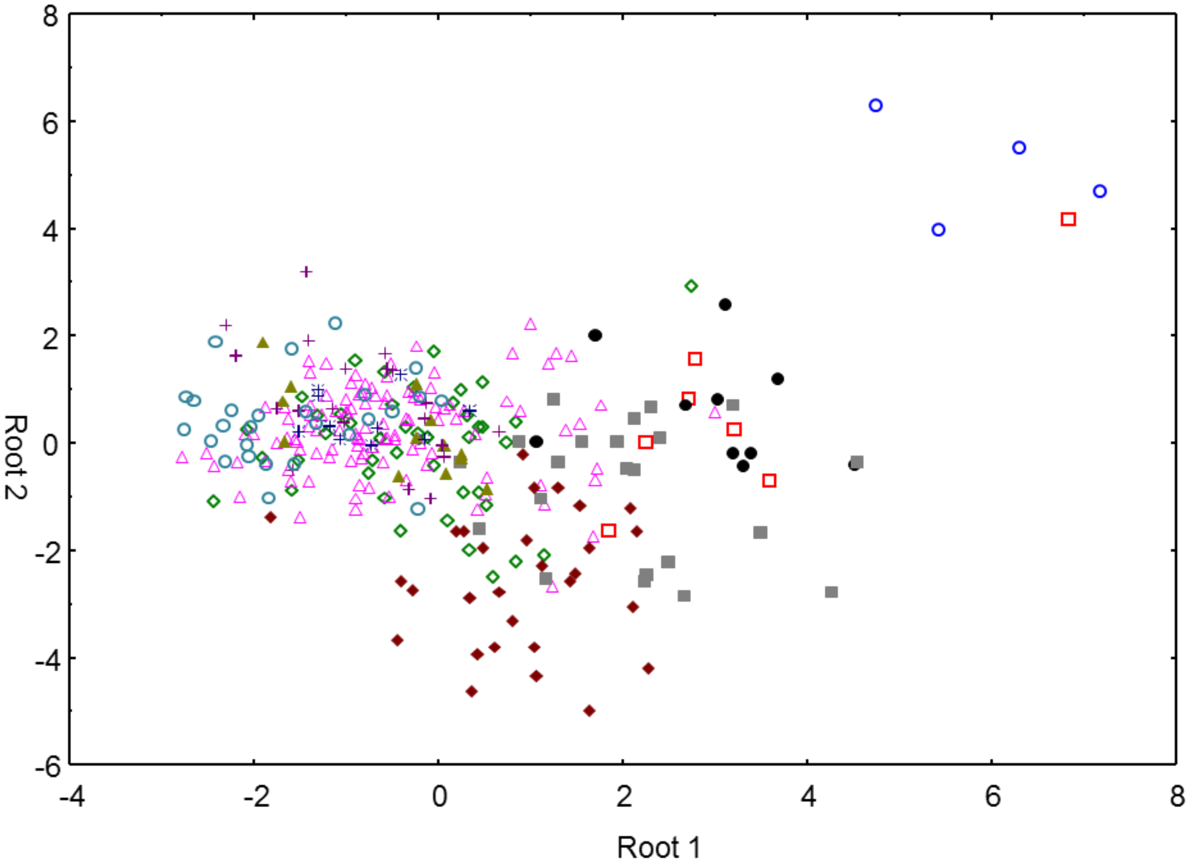


Fig. 3.7. Ordination plot of the canonical variate analysis representing the first two roots, plotted individually for the different species.



Legend: ○ - Mhlatuze river system; □ - Limpopo river system; ■ - Shire river system; △ - Okavango river system; ● - Mkuze swamps; ○ - Kwando river system; ♦ - Upper Zambezi river system; ◇ - Lower Zambezi system; ▲ - Lake Bangweulu system; +- Luapula river system; ⁑ - Congo river system

Fig. 3.8. A) Ordination plot of the principal component analysis representing the first two principal components and **B)** the ordination plot of Factor 2 and Factor 3 of the morphology of *M. johnstoni* from the different river systems.



Legend: ○ - Mhlatuze river system; □ - Limpopo river system; ■ - Shire river system; △ - Okavango river system; ● - Mkuze swamps; ○ - Kwando river system; ♦ - Upper Zambezi river system; ◊ - Lower Zambezi system; ▲ - Lake Bangweulu system; +- Luapula river system; ⋇ - Congo river system

Fig. 3.9. Ordination plot of the Canonical Variate Analysis representing the first two roots for the morphology of *M. johnstoni* from different river systems.

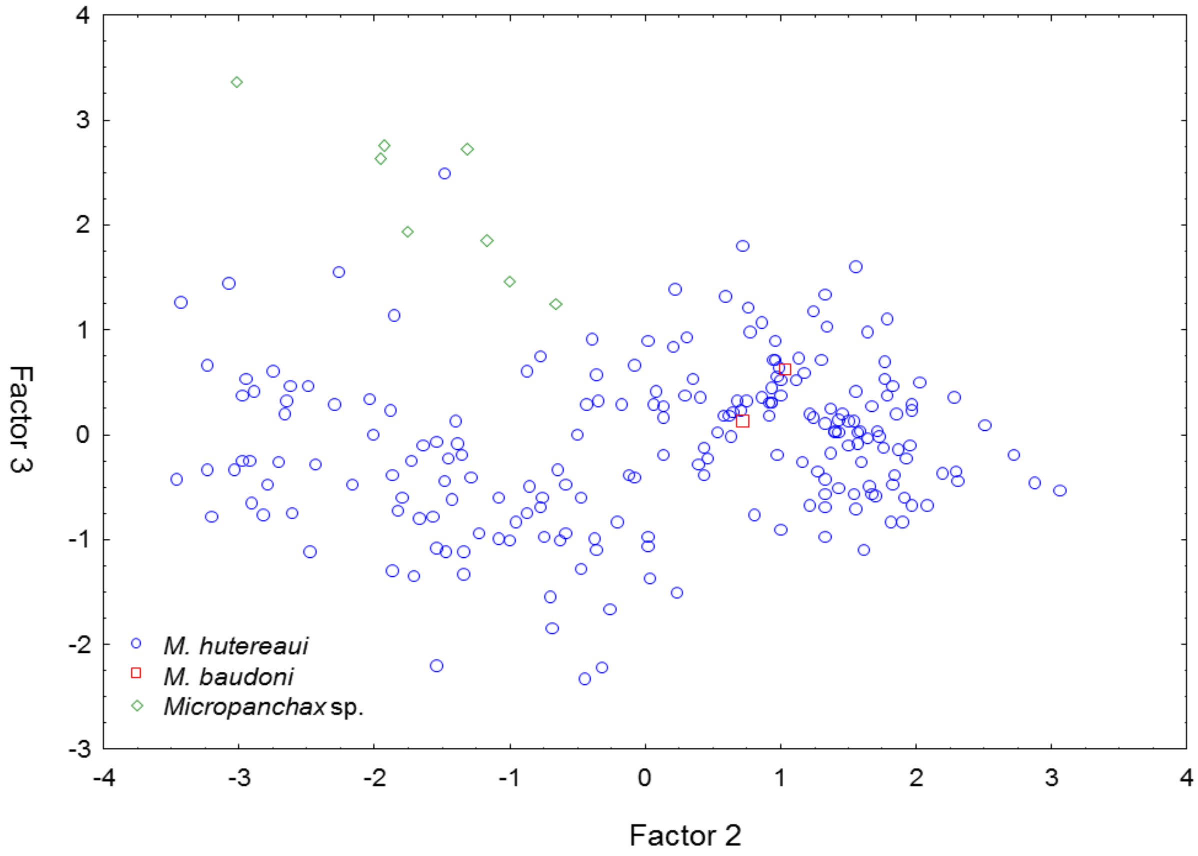
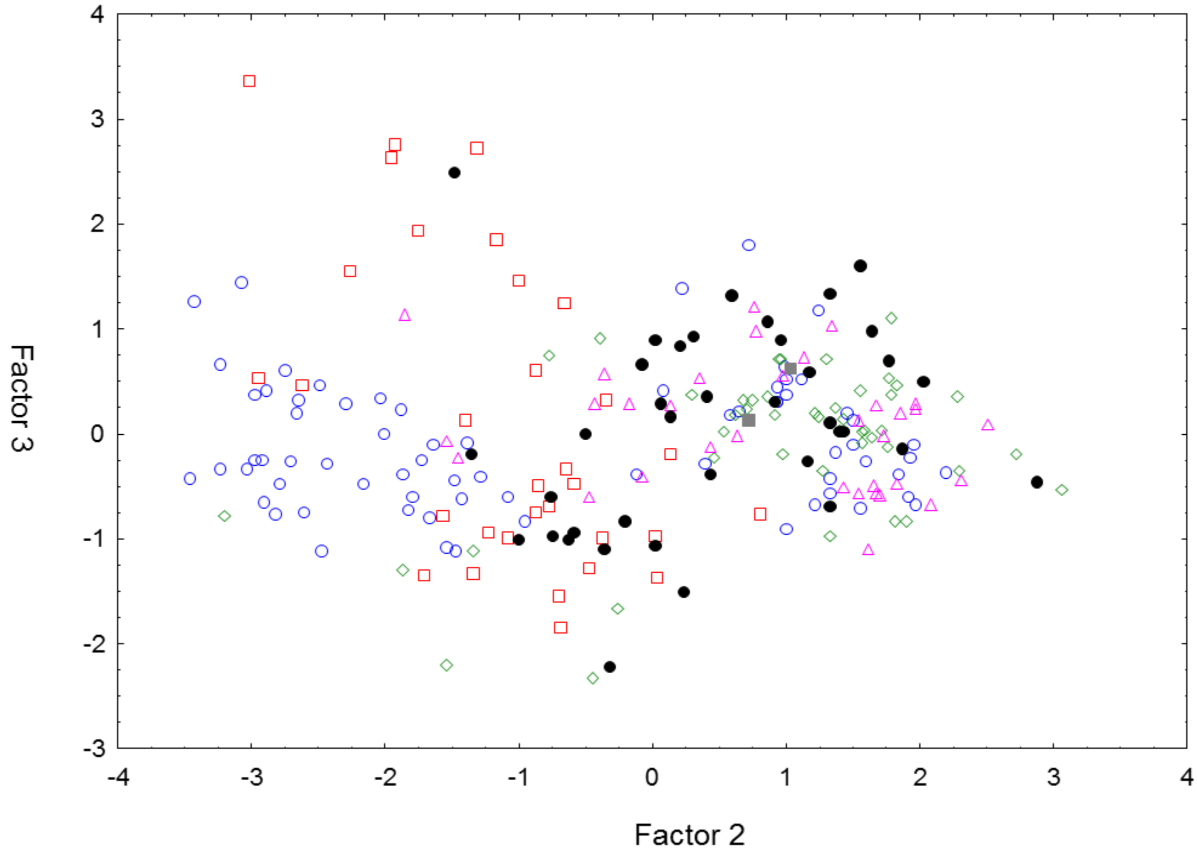


Fig. 3.10. Ordination plot of the principal component analysis representing the first two principal components of the morphology of *M. hutereaui* for the three different species groups.



Legend: ○ - Congo River system; □ - Upper Zambezi River system; ◊ - Okavango River system; △ - Kwando River system; ● - Lower Zambezi River system; ■ - Chad River Basin

Fig. 3.11. Ordination plot of the principal component analysis representing the first two principal components of the morphology of *M. hutereaui* by river system.

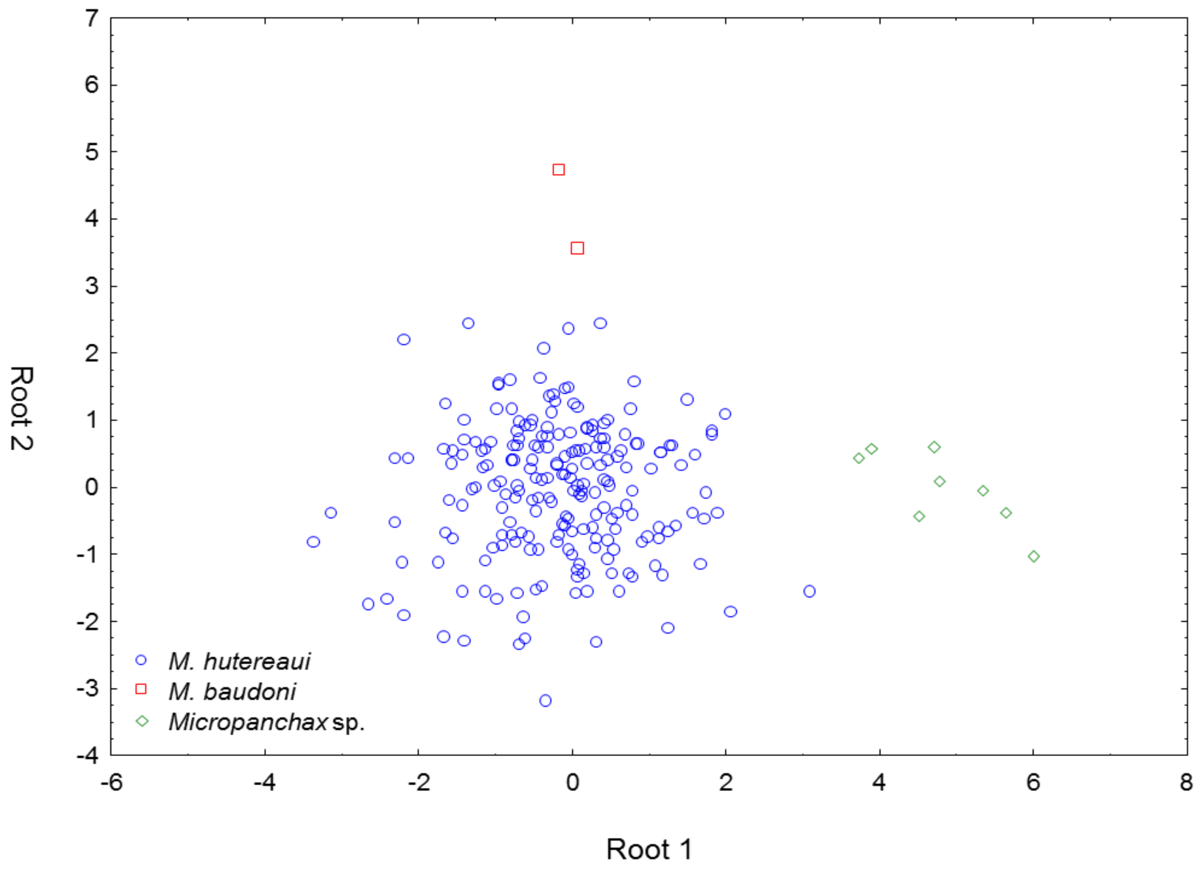
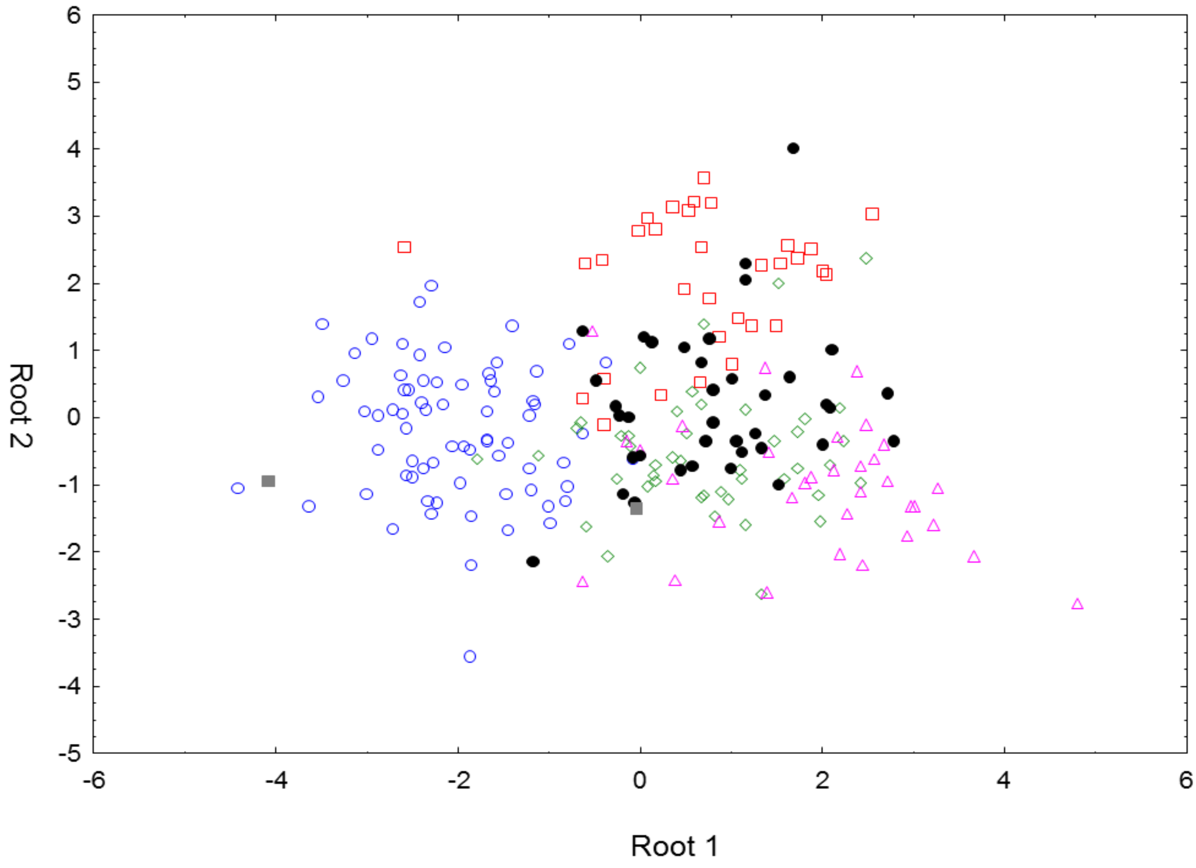
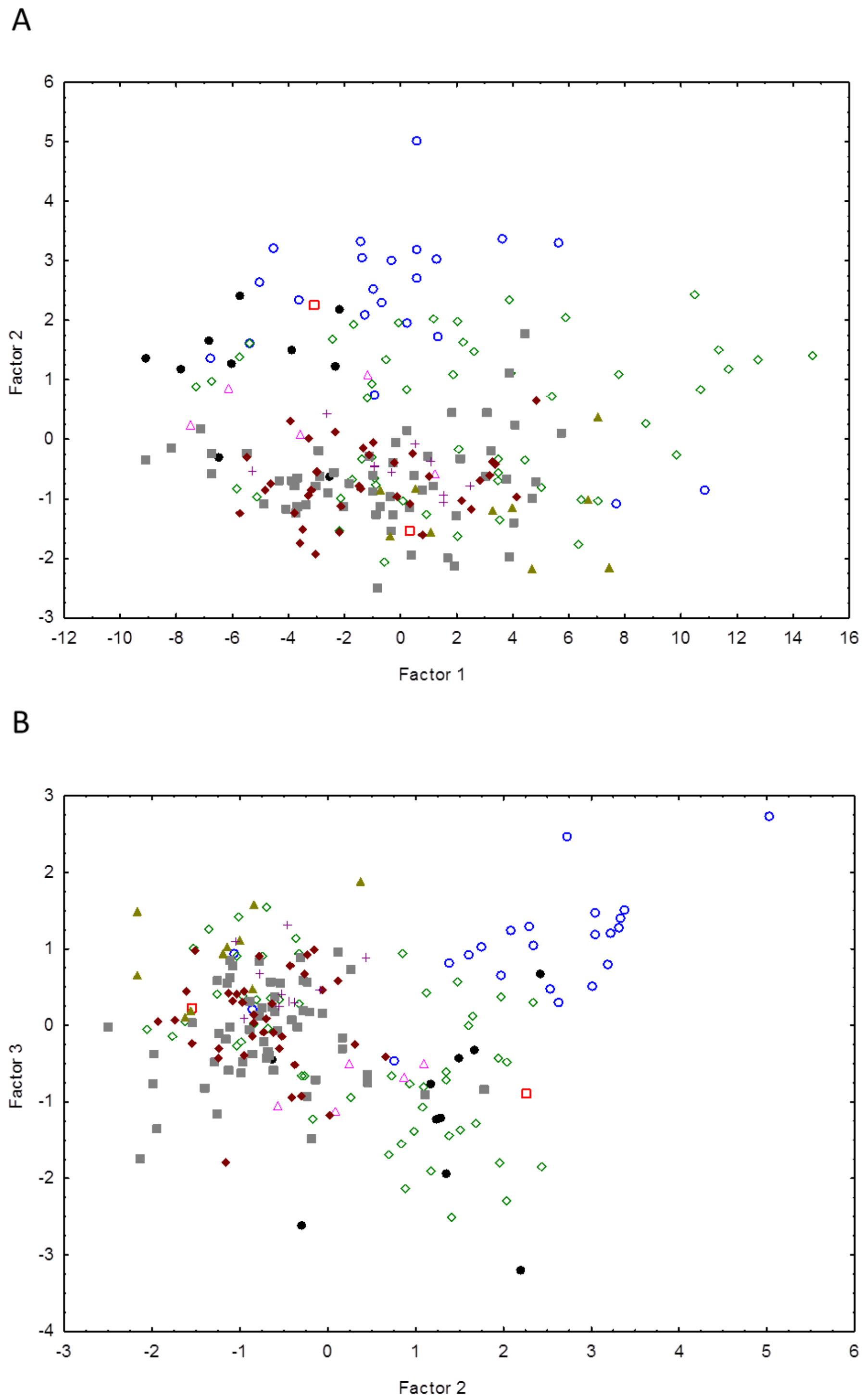


Fig. 3.12. Ordination plot of the canonical variate analysis representing the first two roots of the morphology of *M. hutereaui* for the three different species groups.



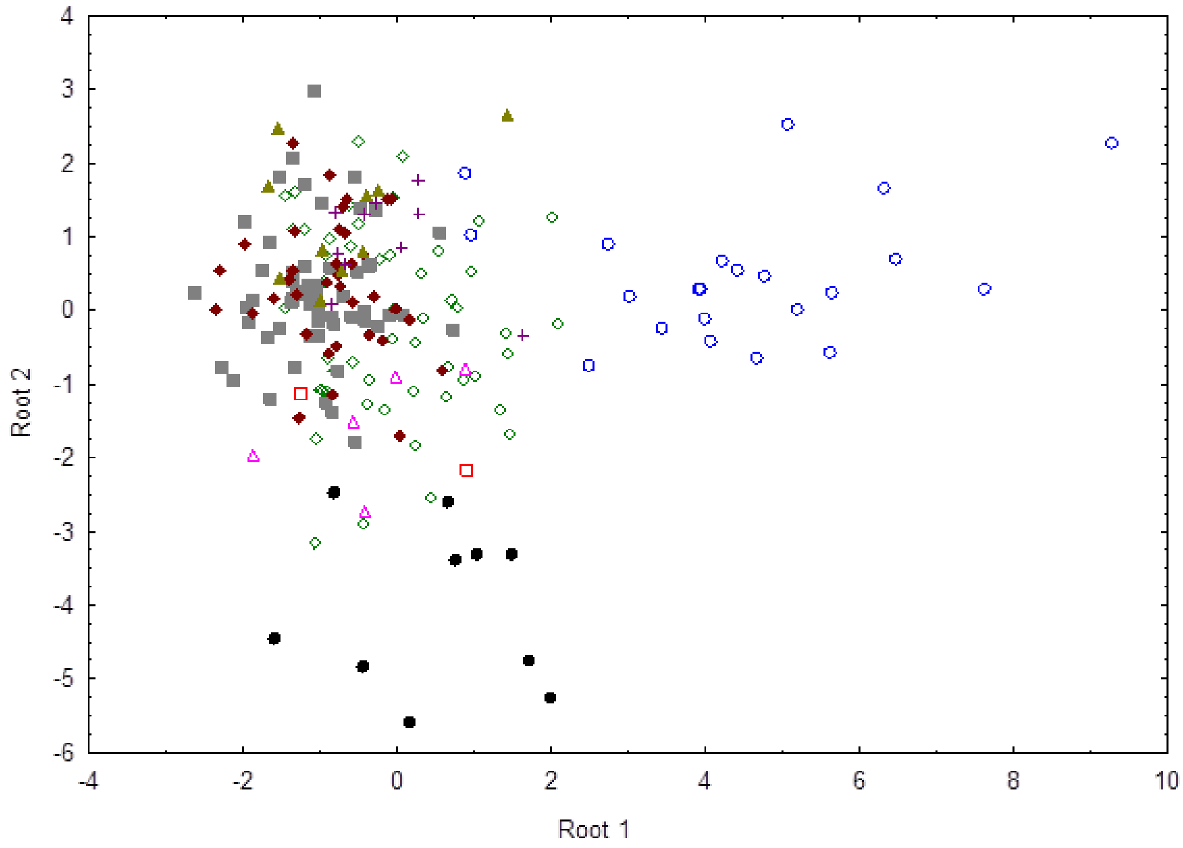
Legend: ○ - Congo River system; □ - Upper Zambezi River system; ◊ - Okavango River system; △ - Kwando River system; ● - Lower Zambezi River system; ■ - Chad River Basin

Fig. 3.13. Ordination plot of the Canonical Variate Analysis representing the first two roots for the morphology of *M. hutereaui* from different river systems.



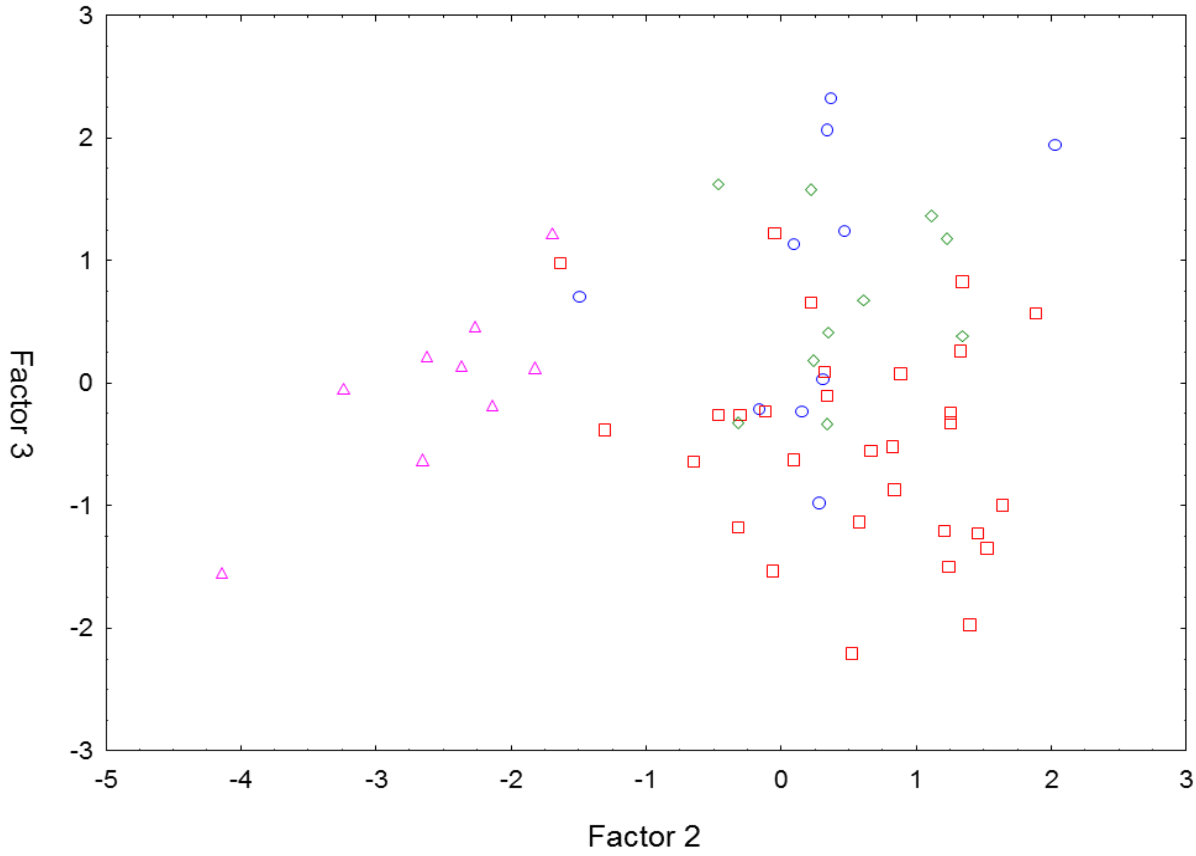
Legend: ○ - Luapula River system; □ - St Lucia system; ◊ - Okavango River system; △ - Congo River system; ▲ - Kwando River system; ● - Lake Sibaya; ♦ - Upper Zambezi River system; ■ - Lower Zambezi River system; + - Lake Xingute

Fig. 3.14. A) Ordination plot of the principal component analysis representing the first two principal components and **B)** the second and third principal components for the morphology of *M. katangae* from different river systems.



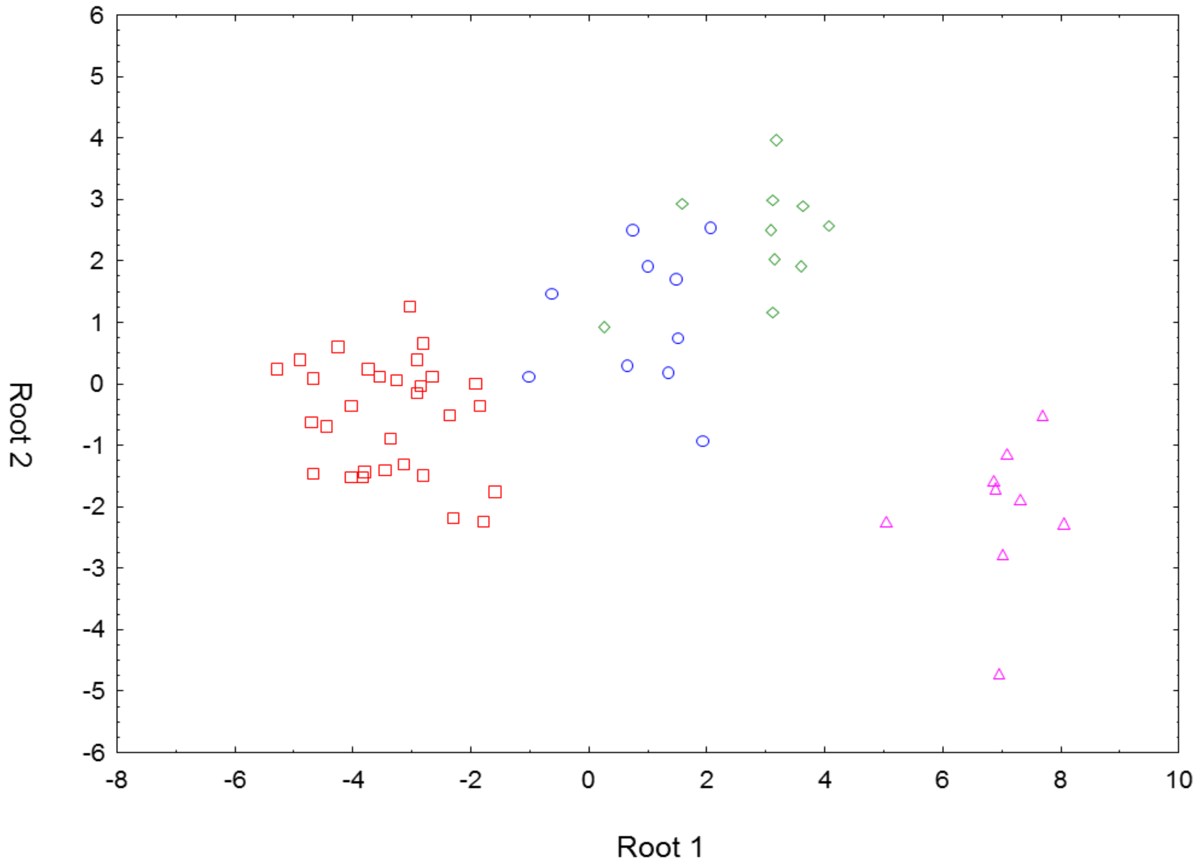
Legend: ○ - Luapula River system; □ - St Lucia system; ◊ - Okavango River system; △ - Congo River system; ▲ - Kwando River system; ● - Lake Sibaya; ♦ - Upper Zambezi River system; ■ - Lower Zambezi River system; + - Lake Xingute

Fig. 3.15. Ordination plot of the Canonical Variate Analysis representing the first two roots for the morphology of *M. katangae* from different river systems.



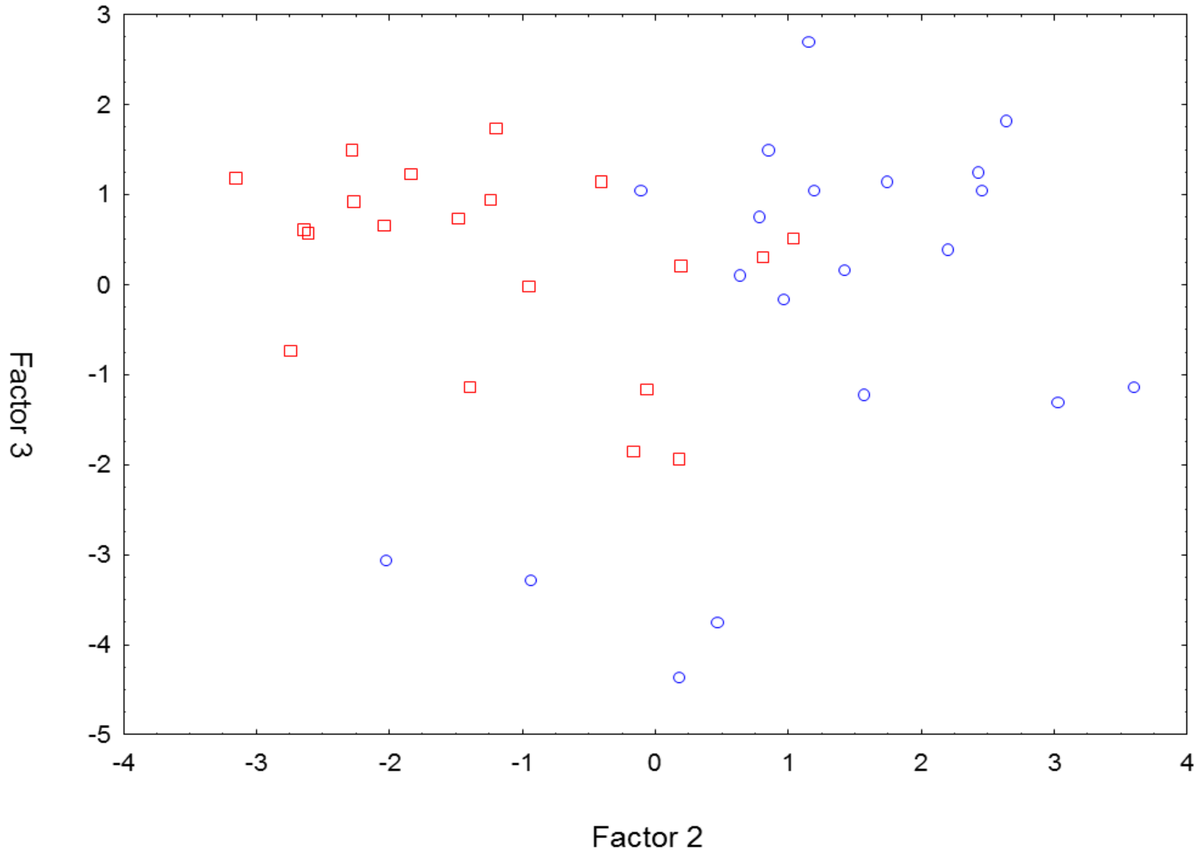
Legend: ○ - Lake Sibaya; □ - Lake Piti; ◊ - Mhlathuze river system; Δ - Nseleni (Coastal) river system;

Fig. 3.16. Ordination plot of the principal component analysis representing the second and third principal components of the morphology of *L. myaposae* from the different river systems.



Legend: ○ - Lake Sibaya; □ - Lake Piti; ◊ - Mhlathuze river system; △ - Nseleni (Coastal) river system;

Fig. 3.17. Ordination plot of the Canonical Variate Analysis representing the first two roots for the morphology of *L. myaposae* from different river systems.



Legend: ○ - Cunene river system; □ - Kwanza river system

Fig. 3.18. Ordination plot of the principal component analysis representing the second and third principal components of the morphology of *L. macrurus* from the different river systems.

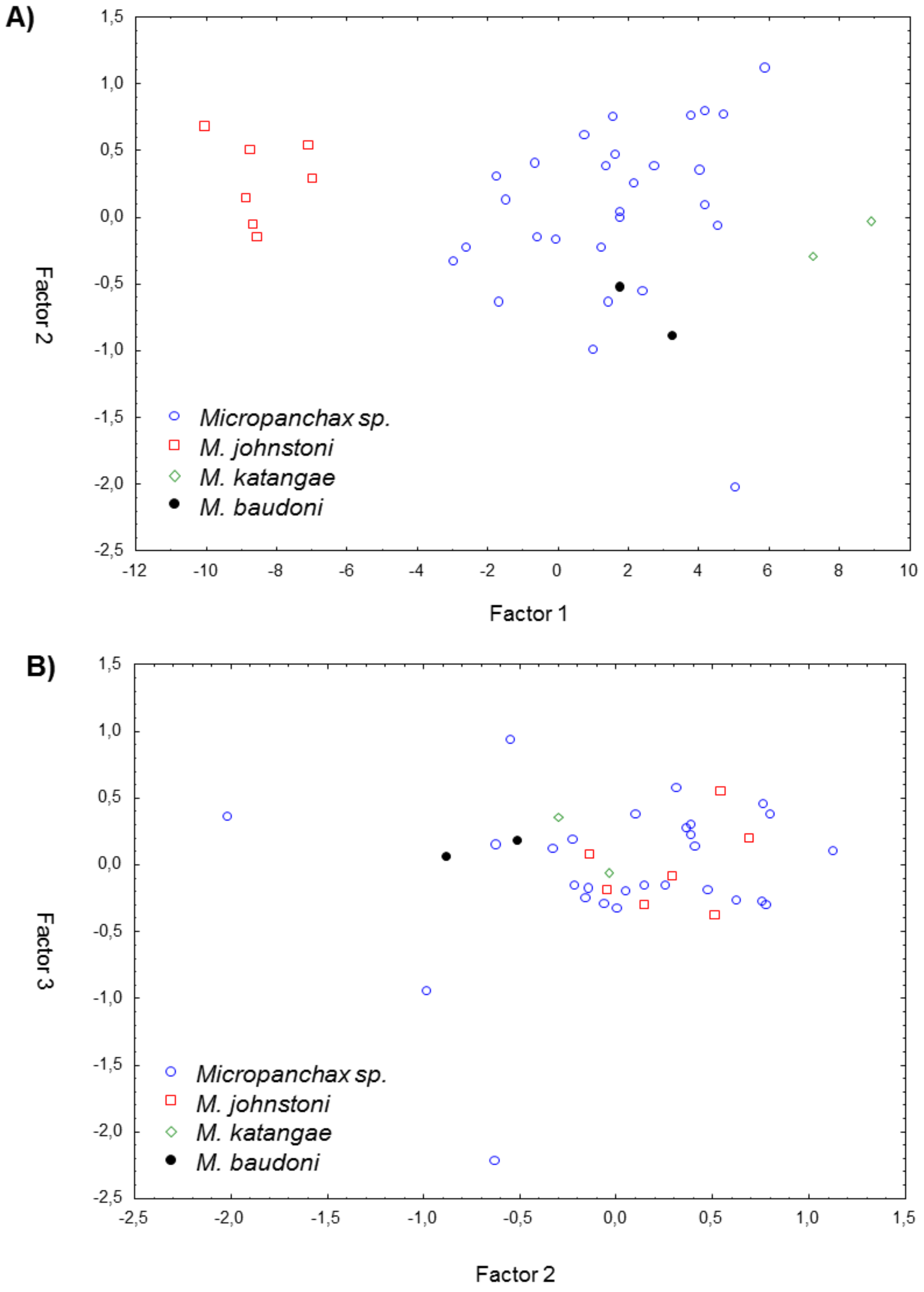


Fig. 3.19. A) Ordination plot of the principal component analysis representing the first two principal components and **B)** the ordination plot representing the second and third principal components of the morphology of *Micropanchax sp.*

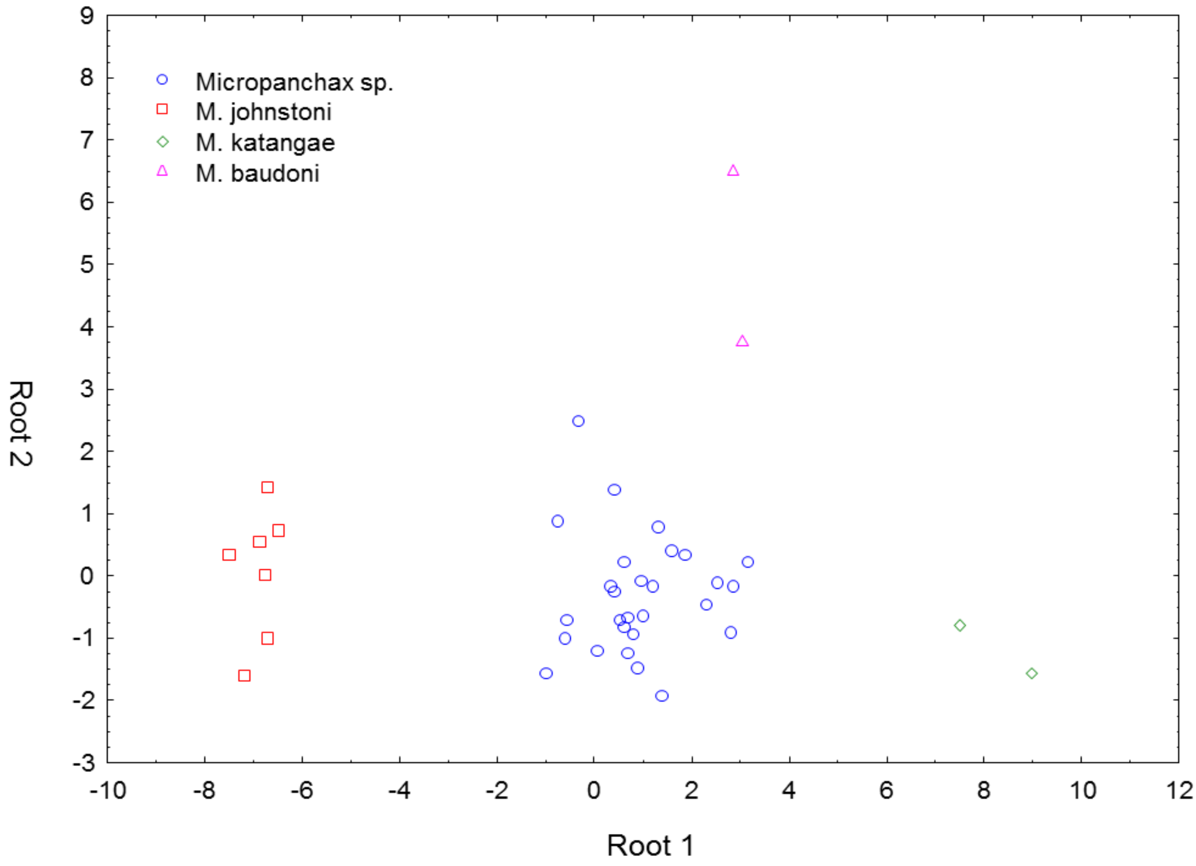


Fig. 3.20. Ordination plot of the Canonical Variate Analysis representing the first two roots for the morphology of *Micropanchax sp.*

4 Implications of the phylogenetic and morphometric analyses of the focal species of *Micropanchax* and *Lacustricola* (Procatopodidae: Procatopodinae) in southern Africa.

4.1. Introduction

The study conducted in this thesis aimed to investigate the genetic and morphometric variation among and within populations of five species of topminnow found in southern Africa. They are currently in two genera, *Micropanchax* and *Lacustricola*, and identified as *M. johnstoni*, *M. katangae*, *M. hutereaui*, *L. myaposae* and *L. macrurus*. Their morphology has been described, but their osteology and molecular taxonomy was not previously explored.

The combination of morphometric and molecular analyses can form a powerful tool to provide insight into the taxonomy of fish species, with the development of new methods and technology allowing for more results in a quicker and less expensive manner (Vrijenhoek 1998; Packer *et al.* 2009; Krück *et al.* 2013; Pereira *et al.* 2013). The rise in use of DNA markers has helped scientists to distinguish between populations of fish and in some cases find genetic diversity within the same population. The discovery of cryptic species and cases of natural hybridization has increased the popularity of DNA evidence, with a number of both mitochondrial and nuclear markers available (Krück *et al.* 2013; Pohl *et al.* 2015; Bragança *et al.* 2018). The advancements in photographic equipment and image analysis software have streamlined the process of morphometrics and has allowed for the analysis of large amounts of data in relatively short periods (Adams *et al.* 2004). The study presented in this thesis made use of a multi-gene phylogenetic analysis and the box-truss morphometric method to analyse variation between and within the different populations of the five target taxa.

4.2. Final discussion

The analyses revealed that the Procatopodidae in southern Africa had many unresolved taxonomic problems and revealed new cryptic species that needed to be described. The phylogenetic analysis provided a clear picture of the populations of

Micropanchax while the morphometric analysis showed that while some populations could be distinguished by the box truss analysis, many of the populations showed body shapes that overlapped in size and shape and often overlapped in scatterplots of ordinations. The use of multi-gene analysis in the phylogenetic study allowed the use of genes that have different substitution rates to test phylogenetic relationships at different historical levels (Hebert *et al.* 2003; Rubinoff 2006). The inclusion of a nuclear marker would normally aid in the detection of hybridization when compared to the mtDNA markers, but the 28S nuclear marker had an extremely slow rate of substitution. This led to little or no species-level variation but was able to provide evidence of relationships of the in-group species (Mindell & Honeycutt 1990).

4.2.1. Genera

The results of the phylogenetic and morphometric analyses showed that a clear distinction could be made between *Micropanchax* and *Aplocheilichthys*. While the morphometric analysis was able to show that *A. spilauchen* is monomorphic, the phylogenetic analysis indicated two possible cryptic species. The total molecular evidence analysis showed typically species-level genetic variation between specimens collected from rivers in southern Africa (two from the Kwanza River and one from the Lucala River, Angola) and one aquarium specimen from Japan.

The genus *Hypsopanchax* showed clear separation from the other genera in the morphometric analysis, with only three specimens misrecorded as *Micropanchax* in the classification matrix in Chapter 3, while the only species of *Hypsopanchax*, *H. platysternus*, used in the phylogenetic study formed part of the out-group and showed no direct relationship to *Micropanchax*.

The phylogenetic analysis revealed that the putative members of *Lacustricola* and *Micropanchax* in southern Africa could not be sorted into reciprocally monophyletic clades that corresponded to these genera. Since the analysis did not include the type species of either genus (*Micropanchax loati* and *L. pumilus*), it is not clear to which genus the southern African species should be assigned, but they should be placed in the same genus to avoid paraphyly. The morphometric analysis showed that the body shapes of both *Micropanchax* and *Lacustricola* were often interpreted as one or the other, which agreed with the phylogenetic analysis. The transfer of the two *Lacustricola* species to *Micropanchax* would lead to the species of *Lacustricola* distributions coherently limited to central Africa (Fricke *et al.* 2018; Froese & Pauly 2018).

4.2.2. Species

The phylogenetic and species delimitation analyses provided more insight into the relationships of the species and the morphometric analysis helped to make the identification of specimens easier while providing diagnostic features for the populations identified in the phylogenetic analysis.

4.2.2.1. Micropanchax

Micropanchax hutereaui formed two genetic clades, Clade A and B, which appeared next to each other on the tree, leaving the species monophyletic. The morphometric analysis of *M. hutereaui* could not distinguish between Clade A and B as Clade B specimens could not be matched to specimens used in the morphometric analysis. *Micropanchax hutereaui* showed genetic variation between the Okavango and Kwando rivers with the Okavango River specimens appearing in Clade A and Kwando River specimen in Clade B, both within Namibia. Namibian specimens from the Zambezi River were found in both clades. The Bayesian Generalized Mixed Yule-Coalescent (bGMYC) showed that while *M. hutereaui* showed genetic variation with the split into two clades, all the specimens of *M. hutereaui* were likely to belong to one species. The Bayesian Poisson Tree Processes analysis (bPTP) showed a different result. It grouped Clade A as one candidate species and split Clade B into two candidate species. The morphometric analysis showed that *M. hutereaui* had slight shape change between the Upper and Lower Zambezi River, which could be a result of both environmental and anthropogenic causes that may have led to the fragmentation of the gene pool, or phenotypic plasticity in response to conditions that vary over time (Engelbrecht & Mulder 1999). The addition of the Kariba and Cahora Bassa Dams to the Zambezi River and the construction of dykes in the lower reaches of the river have caused significant changes to the flow regime and flooding pattern of the river that can lead to phenotypic changes in fishes (Beilfuss & Brown 2010). A similar result was observed in *M. johnstoni* but not in *M. katangae*.

Micropanchax baudoni was considered a synonym of *M. hutereaui* (Fricke *et al.* 2018) before the analysis took place and the result of the morphometric analysis, which included the type material of *M. baudoni*, showed that there was little-to-no shape variation between representatives of the two names. The lack of molecular sequences assignable to *M. baudoni* meant that there was insufficient evidence to reinstate *M. baudoni* as a separate species and confirms that it should remain a synonym of *M. hutereaui* until further evidence can be collected.

The eight specimens found in the morphological study believed to possibly be a new species, currently referred to as the Pygmy Topminnow (Skelton 2001), would need to be included in future phylogenetic study to determine its relationship within *Micropanchax* and would need to be described when its true relationship is known. The Pygmy Topminnow was shown to be morphologically distinct from *M. hutereaui* with a shorter and deeper body shape and different live colouration. *Micropanchax hutereaui* was lighter coloured with a translucent yellowish body, commonly with blue reflections on the scales (Skelton 2001; Marshall 2011). The Pygmy Topminnow was darker in colour with a more light brown body colour; the postero-ventral scales have blue in them. Both species showed turquoise colour iris (Skelton 2001).

Micropanchax hutereaui with Clades A and B which suggested that there was slight genetic differentiation between specimens collected in Namibia and Zambia and between the Kwando and Okavango Rivers within the Caprivi Strip of Namibia. The specimens from the Kwando River showed a colouration similar to that of the Pygmy Topminnow but had some differences such as a dark iris with a white stripe in the bottom of the iris; the Kwando specimens had more blue over the operculum, with blue along the lateral line scales only (fig 4.1 A). The Okavango specimens showed a completely blue iris with a darker brown body and yellow, green and blue colouration on the operculum and lateral scales; the dorsal, anal and caudal fins showed yellow colouration without brown spots, and clear ends of the fins (Fig. 4.1 B).

Micropanchax katangae in the phylogenetic analysis showed low genetic variation of a wide portion of its distribution throughout southern Africa with all of the specimens tested falling in one clade except for two specimens from the Shinengeni Stream, Zambia, and Lake Curumbo, Angola. The species delimitation analyses, the bGMYC and the bPTP showed that all the *M. katangae* specimens are likely to belong to the same species with the only difference shown in the bPTP placing the two specemins, one from Lake Curumbo, Angola and one from the Shinengeni Stream, Zambia, that branched alone as two possible candidate species. It would be diificult to tell what the new candidate species are without the inclusion of material from the type locality to confirm *M. katangae* as a named species. The specimen from the Shinengeni Stream was included in the morphometric analysis but showed no distinction from the rest of the *M. katangae* specimens.

The morphometric analysis of *M. katangae* showed that there was slight change in morphology between specimens found in rivers and that of specimens found in lakes. The lake specimens often have less streamlined bodies with deeper body depths and shorter caudal lengths, as the need to overcome fast-flowing water are not as important (Langerhans *et al.* 2003). This phenotypic diversification could be the result of gene flow or phenotypic plasticity influenced by environmental selection pressures such as resource composition and lake specimens have adapted to a different environmental niche than river specimens (Langerhans *et al.* 2003).

Micropanchax katangae normally showed a deeper body than both *M. johnstoni* and *M. hutereaui*, while only reaching a longer caudal length than *M. hutereaui*. *Micropanchax katangae* was well known for the black stripe along the body that is most often recognizable as a zigzag shape (Fig. 4.1 C) (Skelton 2001).

The phylogenetic analysis of *M. johnstoni* revealed the presence of three clades, Clades C, D and E and showed *M. johnstoni* to be polyphyletic as Clade C grouped closely with *L. myaposae*; Clade D grouped with *L. macrurus* from Olushandja Dam and Clade E showed a close relationship to *M. katangae*. Clade C of *M. johnstoni* included specimens from near the type locality in the Shire River near Fort Johnston, Malawi (Fricke *et al.* 2018). In both of the species delimitation analyses, bGMYC and bPTP, Clade D and E were shown to be candidate species and were most like cryptic species with similar morphologies to *M. johnstoni*. The bGMYC suggested that Clade C, *L. myaposae* and *L. macrurus* clades were likely to be one species but the bPTP was not in agreement and split the three groups. The unidentified specimen *Micropanchax* sp from the Kwanza River was shown in the bGMYC to belong to Clade D but in the bPTP shown to be its own candidate species. This specimen would need further investigation and probable description of a new species.

The morphometric analysis revealed little geographical variation in the morphology of the specimens of *M. johnstoni*. The three clades identified in the phylogenetic analysis showed very little variation between them with Clade C and E showed similar morphologies with Clade D having differed slightly from the other two clades. The group of *M. johnstoni* specimens that could not be identified to belong to either of the clades, showed morphologies similar to that of Clade C and E. That Clade E showed a morphology closer to Clade C was an unexpected result as it showed the most genetic variation from the topotypes. Clade C contained the type material for *M. johnstoni*, providing a definitive identification. The presence of three clades in the phylogenetic and species delimitation analysis and the shared morphologies indicate the presence of cryptic species. This may have arisen due to convergent evolution of morphological features of the three clades. There is little geographical isolation of the three clades, which are often affected by similar or identical environmental selection pressures, predation pressure and anthropogenic disturbances (Díez-del-Molino *et al.* 2018; Østbye *et al.* 2018). Shared selective pressures can lead to the parallel development of the same phenotypic traits, which would lead to convergent morphology, making identification practically impossible without genetic analysis (Doolittle 1994; Tobler *et al.* 2011).

The live colouration was difficult to distinguish, most specimens showed blue or turquoise colours along the body of the fish and yellow in the dorsal, anal and caudal fins. The specimens caught in the Kwando River (Fig. 4.1 D) and Zambezi River (Fig. 4.1 E) (Namibia) showed a similar colouration that was darker than specimens in the Okavango River (Fig. 4.1 F) (Namibia). The Kwando and Zambezi Rivers showed a darker eye and the Zambezi specimens lacked yellow in the anal fin. The Okavango specimens showed brighter blue or turquoise colouration along the body with the upper iris displaying a turquoise stripe.

4.2.2.2. Lacustricola

The phylogenetic analysis revealed two populations of *L. macrurus* in the Kunene River System. One occurred in the Kunene River itself and the second in Olushandja Dam, both sites in northern Namibia. The two groups were shown to be separate species by the bGMYC and bPTP analyses with the specimens with specimens from The Kunene River itself likely to be the named species and the Olushandja Dam specimens, with the rest of Clade D, a new candidate species. Olushandja Dam is connected to the Kunene River System via canals that bring water into the dam and act as a pathway for fauna such as fishes to migrate to the dam (Pelicice *et al.* 2015). The canal that enters Olushandja Dam was observed to open above the dam’s surface, which could act as a barrier preventing fish from travelling back to the Kunene River and limiting the direction of gene flow, similar to how a waterfall may form a natural barrier to fish movements upstream (Crass 1969), eventually creating a genetically isolated population upstream.

The phylogenetic analysis resolved *L. myaposae* into two groups, one from the region surrounding Lake Sibaya in its own clade and the other from nearer to Kosi Bay, north of Lake Sibaya, in Clade C, with both groups occurring in KwaZulu-Natal, South Africa. Two specimens from the Hluhluwe River, South Africa, could not be identified in the field but were identified as *L. myaposae* by the Bayesian analysis. The specimens collected near Kosi Bay and lying within Clade C with topotypical specimens of *M. johnstoni* suggested that these specimens could have been misidentified as *M. johnstoni* and found in a region not previously understood to host *M. johnstoni* (Skelton 2001; Marshall 2011). This suggests that the distribution of *M. johnstoni* extends further south than the Limpopo River, entering KwaZulu-Natal in South Africa near the border of Mozambique. The species delimitation of the bGMYC suggested that both *M. johnstoni* from Clade C and *L. myaposae* could belong to the same species but the bPTP split the two groups with the analysis further splitting the *L. myaposae* specimens into two candidate species, one group containg two specimens from the Mseleni River and the other group with three specimens, two from the Hluhluwe River and one from Lake Mgobozeleni.

The morphometric analysis demonstrated that the body shape of *L. myaposae* was similar to that of *M. johnstoni*, which is unusual as *L. myaposae* is described to have a greater body depth than *M. johnstoni*, although not to the same extent as *M. katangae* (Skelton 2001; Marshall 2011). The overall morphometric similarities between *M. johnstoni*, *L. myaposae* and *L. macrurus* might have arisen through hybridization between the species in areas that the species co-occurred of sympatry, facilitated by their external egg fertilization (Scribner *et al.* 2000; Chiesa *et al.* 2013; Meraner *et al.* 2013). If hybridization had occurred, it could lead to the shared morphologies of the three species and the polyphyly seen in the phylogenetic analysis (Chiesa *et al.* 2013; Meraner *et al.* 2013).

Lacustricola myaposae demonstrated similar result to *M. katangae*, with slight difference in body shape seen between specimens caught in rivers and specimens caught in lakes. The morphometric analysis showed that specimens of *L. myaposae* found in coastal regions had slight morphological variations from the rest of the specimens. This result may be due to changing salinity levels which could cause small changes in the environment, resulting in new phenotypic expression and variations (Warren & Pardew 1998; Engelbrecht & Mulder

1999; Pelicice *et al.* 2015). *Lacustricola myaposae* showed colouration similar to *M. johnstoni* but had a deeper body and the fins of mature males distinctly showed a black colouration with yellow border and yellow spots in the black (Skelton 2001). *Lacustricola macrurus* was characterised by a light translucent brown body with a bright white upper iris and the lower iris a duller white. The lateral scales showed a light blue sheen with bright blue behind the operculum and the all the fins yellow except the pectorals (Skelton 2001). The specimens from Olushandja Dam showed colouration similar to *L. macrurus* in the iris of the eye and the yellow of the fins but showed a darker body with bright blue along the body below the lateral line. The dorsal and anal fins were rounder than in *L. macrurus* (Fig. 4.1 G).

4.3. Species and their distributions

The focus species of this study, *M. johnstoni*, *M. katangae*, *M. hutereaui*, *L. myaposae* and L. *macrurus* have shown the discovery of possible new species in the phylogenetic and morphometric analyses, which may alter understanding of their distributions.

***Haplochilus johnstoni* Günther, 1894** was described from near Fort Johnston, Malawi. The phylogenetic and morphometric analysis support the placement of the species in *Micropanchax* as suggested by Ghedotti (2000) in the family Procatopodidae suggested by Braganca *et al.* (2018), therefore the current status remained *Micropanchax johnstoni* (Procatopodidae: Procatopodinae) (Fricke *et al.* 2018). The phylogenetic analysis showed that *M. johnstoni* is not monophyletic, with the presence of cryptic species shown in the three clades formed by specimens assigned to this name.

The known distribution of *Micropanchax johnstoni* has altered slightly, while remaining the most widely distributed species in southern Africa. The species occurred in the Kunene, Kwando, Okavango, Zambezi, Pungwe, Buzi, Lower Congo and Limpopo Rivers, and could be found in Lake Rukwa and Lake Malawi (Skelton 2001; Marshall 2011; Froese & Pauly 2018). The distribution of *M. johnstoni* has been increased to include northern KwaZulu-Natal near to Kosi Bay. Clade D shown in the phylogenetic analysis was distributed across five rivers, the Cuito, Kwando, Okavango, Thamalakane and Zambezi Rivers in three countries, Angola, Namibia and Botswana. Clade E showed specimens distributed across three rivers, the Kafue, Okavango and Kwando Rivers and included the Boa-fé Flood plain in three countries, Angola, Namibia and Zambia. Clades C and D would have to be described in future to avoid confusion with *M. johnstoni sensu stricto*.

***Haplochilus katangae* Boulenger, 1912** was described from the Lumbumbashi River near Elisabethville, Katanga in the Democratic Republic of Congo. The results of the phylogenetic and morphometric analyses suggest that the genus and family placement was correct as there was no evidence to argue against this, therefore the current status remained *Micropanchax katangae* (Procatopodidae: Procatopodinae) (Ghedotti 2000, Braganca *et al.* 2018; Fricke *et al.* 2018). The conceived distribution of *M. katangae* did not change over the course of the study, with *M. katangae* mainly distributed across the top of southern Africa in the Kunene, Kwando, Okavango and Zambezi River Systems. Its distribution extends south into Mozambique and KwaZulu-Natal, South Africa, but not south of St Lucia (Wildekamp 1995; Skelton 2001; Marshall 2011). Future analysis will need to include topotypes of *M. katangae* to confirm the status of *M. katangae* from the rest of the apparent distribution.

The two specimens from Lake Curumbo, Angola and the Shinengeni Stream, Zambia identified as *M. katangae* in the field but showing genetic distinctions required further analysis but was believed to represent new species and would need to be described. The known distributions of both taxa were restricted to their points of collection, with new material required to perform descriptions because live colouration is required.

***Haplochilus hutereaui* Boulenger, 1913** was described from the upper Uelé River, Dungu Democratic Republic of Congo. The phylogenetic and morphometric analysis showed that species belonged in the same genus as *M. johnstoni* and *M. katangae,* showing a closer relationship to these species than to the genera found elsewhere in Africa. The absence of topotype material in the phylogenetic analysis and the shared morphology with the holotype and paratype material suggests that the current status remains *Micropanchax hutereaui* (Procatopodidae: Procatopodinae) (Fricke *et al.* 2018). *Micropanchax hutereaui* had a smaller distribution than *M. johnstoni* and *M. katangae*, occurring primarily in the Okavango and Zambezi Systems while being found in the lower Shire and the Pungwe-Buzi System and can be found in the Congo and Kwando Rivers (Skelton 2001; Marshall 2011; Froese & Pauly 2018).

Clade B showed a slight genetic difference between the Okavango and Kwando Rivers, with the colour variation seen in both the Kwando and Okavango Rivers occurred in the Zambezi River, which suggests one clade may be a new species. Currently the Pygmy Topminnow that requires describing had a distribution restricted to the Okavango Delta and upper Zambezi flood plains. The colour variation from the Kwando is very similar to the

Pygmy Topminnow; this relationship needs to be investigated further and, if the Kwando specimens are the Pygmy Topminnow, this would increase its distribution to include the Kwando/Linyanti River System.

***Haplochilus myaposae* Boulenger, 1908** was originally described from the Myaposa River, currently known as the Mposa River, Zululand, KwaZulu-Natal, South Africa. The results of the phylogenetic and morphometric analyses showed that the genus that the species is currently placed in is potentially incorrect and the species should be transferred to *Micropanchax*. This would lead to the status becoming *Micropanchax myaposae* (Procatopodidae: Procatopodinae). The apparent distribution of *L. myaposae* did not change over the course of the study, and includes Maputoland and KwaZulu-Natal, South Africa (Skelton 2001; Mmopelwa *et al.* 2009 Marshall 2011).

***Haplochilus macrurus* Boulenger, 1904** described from Lake Sarmento, Marimba in southern Angola showed a phylogenetic and taxonomic result similar to that of *L. myaposae*, forming clades between the species of *Micropanchax*, which suggests that this species may be in the wrong genus and should be transferred from *Lacustricola* to *Micropanchax*. This would lead to it becoming *Micropanchax macrurus* (Procatopodidae: Procatopodinae). In southern Africa *L. macrurus* had a distribution across the Kunene, Kwanza and Okavango River systems in Namibia and Angola (Skelton 2001). The central African distribution of the upper Kasai River and Congo River basin would need to be analysed to determine if the central and southern African taxa are the same species.

The new species found in Olushandja Dam in Northern Namibia showed a distribution restricted to this specific dam where they were collected. The transference of *L. myaposae* and *L. macrurus* to the genus *Micropanchax* would remove the genus *Lacustricola* from southern Africa, leaving it a restricted distribution in central Africa, while the *Micropanchax* would be more widely distributed throughout Africa.

4.4. Conclusion

The results of the study presented in this thesis showed the population structure of the southern African topminnows of the genus *Micropanchax* and how difficult it can be to identify the different populations based on the morphology of the fishes. The identification of new populations or species and the identification of cryptic species would play a role in conservation and management plans (Vrijenhoek 1998; Pusey & Arthington 2003; Strecker *et al.* 2011). The use of phylogenetic and morphometric analyses helped to resolve difficult taxonomic groups which provided a better understanding of gene flow. A better understanding of gene flow could lead conservation agencies implementing management plans that allow gene flow to occur in areas that it may be needed (Vrijenhoek 1998; Pusey & Arthington 2003; Strecker *et al.* 2011). Understanding gene flow and the detection of diversity within and between populations could provide extra information on how anthropogenic changes to riparian zones and flow regimes of rivers affect fish populations. This would allow management agencies to implement effective strategies for the conservation of the fish species, especially those that have restricted distributions or appear on the endangered species lists (Vrijenhoek 1998; Pusey & Arthington 2003; Strecker *et al.* 2011). Phylogenetic analyses identify species that would require more attention for conservation and identify geographical areas that would require specific management plans especially in freshwater ecosystems where upstream and downstream processes would have to be taken into account (Vrijenhoek 1998; Strecker *et al.* 2011).

The phylogenetic, species delimitation and morphometric analyses presented in this thesis showed the existence of cryptic species with Clades D and E showing morphology similar to that of *M. johnstoni*. The study showed the influence of anthropogenic changes in river systems with the implementation of dams seen in the slight changes found between the upper and lower Zambezi River and the influence of canals seen with *L. macrurus* in the Kunene River. The construction of these man-made features has had an impact on the distribution of the fishes and the gene flow between populations of the species. This demonstrates that future installations would need carefully considered plans on how to conserve fish species of this study and other fishes found in the same rivers. The two species of *Lacustricola* in the study, *L. myaposae* and *L. macrurus*, showed very close relationship to *M. johnstoni* and it has been suggested that the two species be transferred to the genus *Micropanchax*.

The phylogenetic and morphometric analysis demonstrated that *A. spilauchen* did not have a close relationship to the species of *Micropanchax* which agreed with Huber (1982) and Ghedotti (2000). *Aplocheilichthys spilauchen* was shown in the phylogenetic analysis to possibly represent two species. The small difference seen in the morphology of all of the species showed that there may be low gene flow between the different populations and that environmental conditions had an impact on the phenotype displayed by the fishes as most of the fishes were found to co-occur in the same biotopes of the rivers within the same rivers and were often caught together, and therefore experience the same overall environmental conditions. The slight changes seen in the morphometric analysis and the results of the box truss analysis indicated that body depth, position of the dorsal and anal fin and length of the tail were the body regions most variable to change.

Future studies of these fish species should include more topotype material for the phylogenetic analysis for *M. hutereaui* and *M. katangae* and could include specimens from other areas of the distributions of all of the species. The morphometric analysis should include more measurements, especially around the head and the caudal region of the fishes. Crucial measurements of the fin structure of the fishes could not be completed due to older specimens having damaged dorsal, anal and caudal fins. Live colouration could not be analysed to its full extent due to the delicate nature of the fish when stored for long periods of time, losing their colouration and becoming dull yellow with translucent fins. Colouration was an important identifying tool for these fishes. Overall, the phylogenetic and morphometric analyses were able to provide resolution to the taxonomy of the southern African topminnows and showed the presence of cryptic species in *M. johnstoni* populations and a number of possible new species that will need to be described in the near future.

4.5. References

ADAMS, D.C., RHOLF, F.J. & SLICE, D.E. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71(1)**: 5-16.

BEILFUSS, R. & BROWN, C. 2010. Assessing environmental flow requirements and trade-offs for the Lower Zambezi River and Delta, Mozambique. *International Journal of River Basin Management* **8(2)**: 127-138.

BOULENGER, G.A. 1904. Descriptions of new West-African freshwater fishes. *Annals and Magazine of Natural History* **14**: 16-20.

BOULENGER, G.A. 1908. On a collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum* **1**: 219-235.

BOULENGER, G.A. 1912. Description d'un poisson nouveau du genre *Haplochilus* provenant du Katanga. *Revue de Zoologie Africaine* **2**: 47-48.

BOULENGER, G.A. 1913. Sur une petite collection de poisons recueillis dans l'Uelé par la mission dirigée par M, hutereau. *Revue de Zoologie Africaine* **2**: 155-161.

BRAGANCA, P.H.N., AMORIM, P.F. & COSTA, W.J.E.M. 2018. Pantanodontidae (Teleostei, Cyprinodontiformes), the sister group to all other cyprinodontoid killifishes as inferred by molecular data. *Zoosystematics and Evolution* **94(1)**: 137-145.

CHIESA, S., FILONZI, L., VAGHI, M., PAPA, R. & MARZANO, F.N. 2013. Molecular barcoding of an atypical cyprinid population assessed by Cytochrome b gene sequencing. *Zoological Science* **30**: 408-413.

CRASS, R.S. 1969. The effects of land use on freshwater fish in South Africa, with particular reference to Natal. *Hydrobiologia* **34(1)**: 38-56.

DÍEZ-DEL-MOLINO, D., GARCÍA-BERTHOU, E., ARAGUAS, R.A., ALCARAZ, C., VIDAL, O., SANZ, N. & GARCÍA-MARÍN, J.L. 2018. Effects of water pollution and river fragmentation on population genetic structure of invasive mosquitofish. *Science of the Total Environment* **637**: 1372-1382.

DOOLITTLE, R.F. 1994. Convergent evolution: the need to be explicit. *Trends in Biochemical sciences* **19**: 15-18.

ENGELBRECHT, G.D. & MULDER, P.F.S. 1999. Extremely high genetic differentiation between two populations of the river goby, *Glossogobius callidus* (Smith, 1937). *Water SA* **25(1)**: 85-90.

FOWLER, H. W. 1916. Notes on the fishes of the orders Haplomi and Microcyprini. *Proceedings of the Academy of Natural Sciences, Philadelphia* **68**: 415-439.

FRICKE, R., ESCHMEYER, W.N. & VAN DER LAAN, R. (eds). 2018. Catalog of fishes: genera, species, references (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 26 November 2018.

FROESE, R. & PAULY, D. EDS. 2018. World Wide Web electronic publication (http://www.fishbase.org), version (02/2018). Electronic version accessed 26 November 2018.

GHEDOTTI, M.J. 2000. Phylogenetic analysis and taxonomy of the poecilioid fishes (Teleostei: Cyprinodontiformes). *Zoological Journal of the Linnean Society* **130**: 1-53.

GÜNTHER, A. 1894. Second report on the reptiles, batrachians, and fishes transmitted by Mr. H. H. Johnston, C. B., from British Central Africa. *Proceedings of the Zoological Society of London 1893* (pt 4): 616-628, Pls. 53-57.

HEBERT, P.D.N., CYWINSKA, A., BALL, S.L. & DEWAARD, J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* **270**: 313-321.

HUBER, J. H. 1982. *A review of the cypronodont fauna of the coastal plain in Rio Muni, Gabon, Congo, Cabinda, and Zaire, with taxonomic shifts in* Aphyosemion, Epiplatys and *West African Procatopodins.* British Killifish Association, London.

KRÜCK, N.C., TIBBETTS, I.R., WARD, R.D., JOHNSON, J.W., LOH, W.K.W. & OVENDEN, J.R. 2013. Multi-gene barcoding to discriminate sibling species within a morphologically difficult fish genus (*Sillago*). *Fisheries Research* **143**: 39-46.

LANGERHANS, R.B., LAYMAN, C.A., LANGERHANS, A.K. & DEWITT, T.J. 2003. Habitat-associated morphological divergence in two Neotropical fish species. *Biological Journal of the Linnean Society* **80**: 689-698.

LUCINDA, P.H. 2003. Family Poeciliidae. *Check list of the freshwater fishes of south and Central America*, pp. 555-581.

LUCINDA, P.H.F. & REIS, R.E. 2005. Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. *Neotropical Ichthyology* **3**: 1-60.

MARSHALL, B. 2011. *The fishes of Zimbabwe and their biology*. Smithiana Monographs. **3**: i-vi + 1-290.

MERANER, A., VENTURI, A., FICETOLA, G.F., ROSSI, S., CANDIOTTO, A. & GANDOLFI, A. 2013. Massive invasion of exotic *Barbus barbus* and introgressive hybridization with endemic *Barbus plebejus* in Northern Italy: where, how and why? *Molecular Ecology* **22**: 5295-5312.

MINDELL, D.P. & HONEYCUTT, M.R. 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Annual Review of Ecology and Systematics* **21**: 541-566.

MMOPELWA, G., MOSEPELE, K., MOSEPELE, B., MOLEELE, N. & NQWENYA, B. 2009. Environmental variability and the fishery dynamics of the Okavango Delta, Botswana: the case of subsistence fishing. *African Journal of Ecology* **47**: 119-127.

ØSTBYE, K., TAUGBØL, A., RAVINET, M., HARROD, C., PETTERSEN, R.A., BERNATCHEZ, L. & VØLLESTAD, L.A. 2018. Ongoing niche differentiation under high gene flow in a polymorphic brackish water threespine stickleback (*Gasterosteus aculeatus*) population. *BMC Evolutionary Biology* **18(1)**: 1-14.

PACKER, L., GIBBS, J., SHEFFIELD, C. & HANNER, R. 2009. DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources* **9**: 42-50.

PARENTI, L. R. 1981. A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History* **168**: 335-557.

PELCICE, F.M., POMPEU, P.S. & AGOSTINHO, A.A. 2015. Large reservoirs as ecological barriers to downstream movements of neotropical migratory fish. *Fish and Fisheries* **16**: 697-715.

PEREIRA, L.H.G., HANNER, R., FORESTI, F. & OLIVEIRA, C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics* **2013**: 14-20.

PUSEY, B.J. & ARTHINGTON, A.H. 2003. Importance of the riparian zone to the conservation and management of freshwater fish: a review. *Marine and Freshwater Research* **54**: 1-16.

RUBINOFF, D. 2006. Utility of mitochondrial DNA barcodes in species conservation. *Conservation Biology* **20**: 1026-1033.

SCRIBNER, K.T., PAGE, K.S.& BARTRON, M.L. 2000. Hybridization in freshwater fishes, a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* **10(3)**: 293-323.

SKELTON, P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town.

SKELTON, P.H. 2002. Changes to the scientific and common names of southern African freshwater fishes. *African Journal of Aquatic Science* **27**: 171-174.

STEENKAMP, M.K.J., ENGELBRECHT, G.D. & MULDER, P.F.S. 2001. Allozyme variation in a Johnston's topminnow, *Aplocheilichthys johnstoni*, population from the Zambezi River system. *Water SA* **27**: 53-55.

STRECKER, A.L., OLDEN, J.D., WHITTIER, J.B. & PAUKERT, C.P. 2011. Defining conservation priorities for freshwater fishes according to taxonomic, functional, and phylogenetic diversity. *Ecological Applications* **21(8)**: 3002-3013.

TOBLER, M., PALACIOS, M., CHAPMAN, L.J., MITROFANOV, I., BIERBACH, D., PLATH, M., ARIAS-RODRIGUEZ, L., GARCÍA DE LEÓN, F.J. & MATEOS, M. 2011. Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. *Evolution* **65**: 2213-2228.

VRIJENHOEK, R.C. 1998. Conservation genetics of freshwater fish. *Journal of Fish Biology* **53**: 394-412.

WARREN, JR, M.L. & PARDEW, M.G. 1998. Road crossings as barriers to small-stream fish movement. *Transactions of the American Fisheries Society* **127**: 637-644.

WILDEKAMP, R.H. 1995. A world of killies. Atlas of the oviparous cyprinodontiform fishes of the world. Volume II. The genera *Aplocheilichthys*, *Aplocheilus*, *Austrofundulus*, *Campellolebias*, *Crenichthys*, *Cualac*, *Cubanichthys*. *Cynolebias*, *Cynopoecilus* and *Cyprinodon*. *The American Killifish Association, Inc*. 384 p.

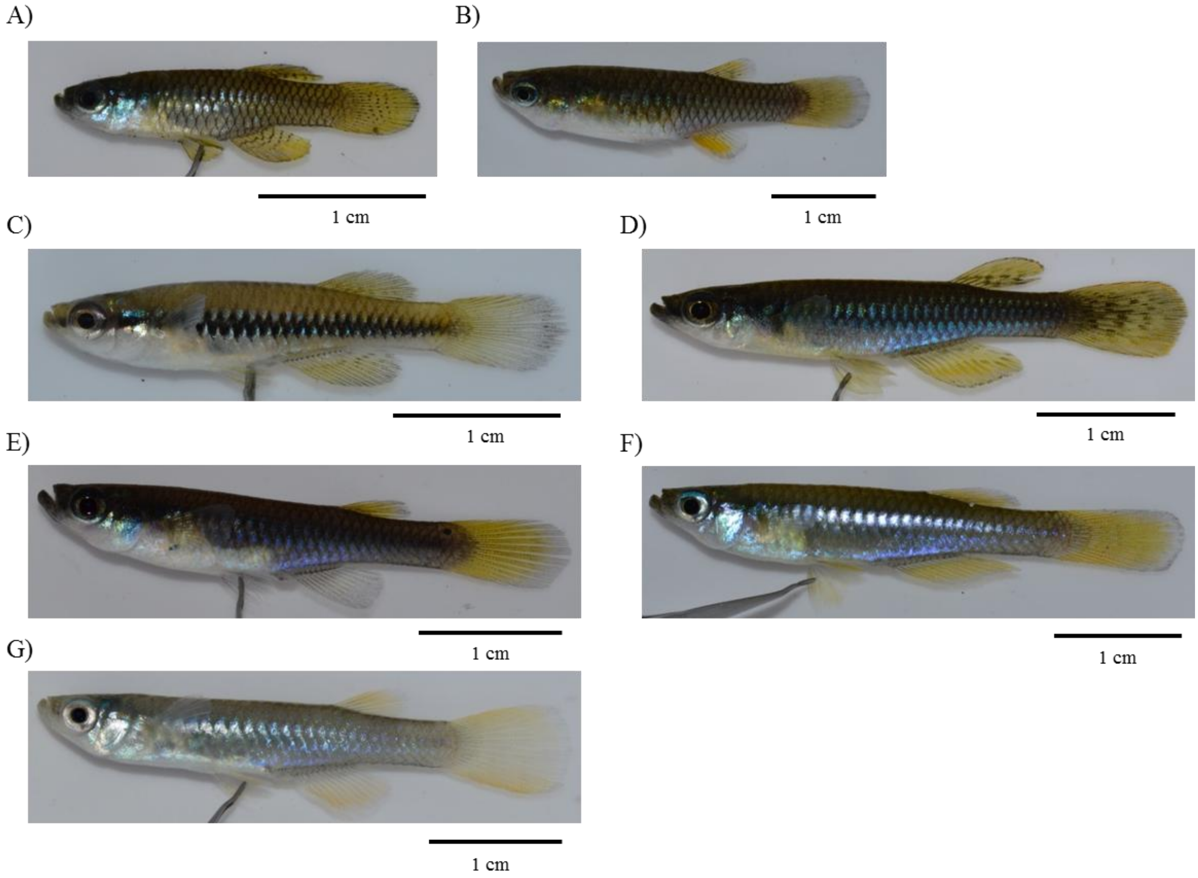


Fig. 4.1. Photographs of the live colouration of unique specimens identified by the phylogenetic and morphometric analyses, illustrating slight differences between certain river systems in Namibia. **A)** *Micropanchax hutereaui* from the Kwando River. **B)** *Micropanchax hutereaui* from the Okavango River. **C)** *Micropanchax katangae*. **D)** *Micropanchax johnstoni* from the Kwando River. **E)** *Micropanchax johnstoni* from the Zambezi River. **F)** *Micropanchax johnstoni* from the Okavango River. **G)** Specimens from Olushandja Dam which could be a new species.