Systematics and biogeography of *Mesobola brevianalis* (Boulenger, 1908) (Teleostei: Cyprinidae)

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

Master of Science in Zoology

at

Rhodes University

by

Megan Amy Riddin
BSc (*Rhodes 2011*)*; BSc Hons* (*African Vertebrate Biodiversity*) (*Rhodes 2012*)

February 2014

Abstract

The accurate identification of fish species, their life stages and their products enables the correct management of fisheries, research and conservation of distinctive populations for long-term survival and sustainability.

Mesobola brevianalis Boulenger, 1908, commonly known as the river sardine, is found in many river systems in southern Africa. Because it exhibits widely separated populations showing subtle differences, particularly in colour, it is thought that there may be cryptic species involved.

Standard phylogenetic techniques using three genetic markers (mitochondrial COI, nuclear protein RAG1 and nuclear ribosomal 28S rRNA), enabled the building of phylogenetic networks for *M. brevianalis* and some outgroup species. Consistent patterns of relationship were seen with 28S supporting monophyly. COI and RAG1 suggested that populations that are currently identified as *M. brevianalis* in fact represent several species. There was sufficiently strong support for the evolutionary independence of the *M. brevianalis* populations from the Rovuma, Kunene and Orange River Systems to consider them as independent species. The independence of the genus *Mesobola* was brought into question because *Engraulicypris sardella* and *Rastrineobola argentea* were placed within it phylogenetically.

Morphometric methods in the form of multivariate truss network analyses, were performed to locate morphological markers for populations. There was little to no variation among most of the populations synonymized under *M. brevianalis*. Furthermore, neither latitude nor longitude had an effect on the morphological characters that might be linked to functional evolution.

A molecular clock analysis of COI data was used to calibrate a paleobiogeographical model which entailed a divergence of lineages starting from an easterly reigning Paleo-Congo Basin, via a hypothetical Paleo-Kalahari Lake that was fragmented by a series of uplifts and drying events beginning ~65 million years ago.

Complete evidence supported the synonymisation of the genera *Engraulicypris* and *Mesobola*, the resurrection of the species name *gariepinus* for the Orange River System population, and the description of two new species: *E. ngalala* from the Rovuma River System and *E. howesi* from the Kunene River System.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I am extremely appreciative for all the help and support of the many people who helped to make this thesis possible. First, I would like to thank my supervisors, Martin Villet and Roger Bills, for their support and guidance throughout the study. I cannot thank Martin Villet enough for the endless guidance, support and encouragement, and for the many late nights and long hours he put into his supervision of my work. He was always available to answer my many, many questions. It is thanks to him that this study has been completed and for which I will be eternally grateful.

A huge thank you must go to Roger Bills for his support, constant help in the lab and especially for taking the time to travel throughout South Africa with me to collect fish in many crocodile-infested rivers.

My sincere thanks also go to Paul Skelton for always being prepared to offer guidance and advice throughout the study and to the SAIAB laboratory staff, particularly Nkosinathi Mazungula and Mark Lisher at SAIAB for helping with cataloguing, X-rays and general assistance in the lab. Thanks also to Bafo Konqobe and Vuyani Hanisi for their help with the loaning of materials.

I would like to thank James Maclain at the British Museum of Natural History for providing photographs, X-rays and cat-scans of type specimens when asked and Fenton P.D. Cotterill at the University of Stellenbosch for offering his knowledge and assistance regarding the paleobiogeographical history and molecular clock analysis. In particular, I am grateful for all Dr Alicia Timm's help in the genetics laboratory and for the many hours she put into teaching me how to use the equipment and helping to solve PCR problems. A special thank you must also be made to Hugo Retief for his help with generating maps and for his patience and encouragement throughout the study.

This study would not have been possible without the funding supplied by a number of people and organisations. Thank you to the National Research Foundation (NRF) for funding provided by the NRF Innovation Scholarships, the South African Institute for Aquatic Biodiversity (SAIAB) and Martin Villet for financial support.

Finally, the constant love, support and encouragement from my family were what motivated me to continue to reach for my goals; without my family's belief in me, I would not have come as far as I have.

1 Introduction

1.1 Background to study

Globally, fish biodiversity conservation, management and assessment rely heavily on correct taxonomic assessments and species identifications (Ward *et al*. 2009). Accurately identifying fish species, their life stages and their products has become an essential aspect of ichthyological science and conservation (Ward *et al.* 2005). Research enables the correct management of fisheries through research and conservation of distinctive populations for long term survival and sustainability (Ward *et al*. 2005, 2009).

The last 300 years have seen cryptic species ('two or more distinct species classified as a single species': Bickford *et al.*, 2006) posing a problem for identification, management and conservation (Bickford *et al*. 2006). The river sardine, *Mesobola brevianalis* (Boulenger, 1908), has been reported in several river systems across southern Africa (van der Horst 1934; Barnard 1943; Hay *et al*. 2008). Populations from different river systems show subtle differences in morphology and colouration that may indicate cryptic species. However, very little research has been done on this taxon, perhaps because of its highly underestimated commercial potential (Engelbrecht & Mulder 1999).

There are currently four species in the genus *Mesobola* (Froese & Pauly 2014) (Table 1.1). *Mesobola brevianalis* is important for various reasons particularly its great commercial potential as a southern African food source for game and predatory fish, including nembwe (*Serranochromis robustus*), silver catfish *(Schilbe intermedius*), tigerfish (*Hydrocynus vittatus*) and bass (*Micropterus* spp.) (Engelbrecht & Mulder 1999). The importance of this species further includes its role as bait for southern African subsistence fisherman and as an aesthetic aquarium species (Engelbrecht & Mulder 1999).

River sardines are a small, shoaling species that favours the upper stratum of open waters particularly in rivers and dams (Engelbrecht & Mulder 1999), where they breed in early summer and feed on planktonic crustaceans and insects (Hay *et al*. 2008).

They have been found and documented across many river systems in southern Africa including the Kunene, Orange, Okavango and Zambezi River systems (van der Horst 1934;

Barnard 1943; Hay *et al*. 2008). They have been recorded as endemic to the Orange River system; they were found to be widespread below the Augrabies Falls, a natural fish barrier, and are the most common species in the lower Orange River (Ramollo 2011). The river sardine has also been found in the eastern coastal rivers from the Limpopo River (Olivier *et al*. 2009) to the Mfolozi River (Hay *et al*. 2008). They are thought to be exotic to the Lower Zambezi River, since they were introduced into the Nyamombe River, which flows into the Mazowe River in Mozambique (Kadye 2008).

Mesobola brevianalis has undergone a number of taxonomic rediagnoses and synonymising of species within the population since its description by Boulenger (1908).

***Neobola brevianalis* Boulenger 1908** was described from a single specimen from the Mkuzi River in Zululand, South Africa (Table. 1.1), an opinion which was later reassessed by Boulenger (1911) to *Engraulicypris brevianalis* based on further specimens from the Dwaars River, Transvaal. A number of later authors agreed with Boulenger's (1911) placement within *Engrualicypris* (Gilchrist & Thompson 1913; Boulenger 1916; David 1936; Barnard 1943; Barnard 1947; Crass 1960; Farquharson 1962; Whitehead 1962; Jubb 1961; Jubb 1963; Bell-Cross 1965; Jubb & Farquharson 1965; Jubb 1967; Jubb & Gaigher 1971; Gaigher & Pott 1973; Bell-Cross 1976; Fowler 1977). The name reverted to *Neobola brevianalis* by Bruton & Kok (1980), with no explanation, and later supported by Howes (1980) and Lévêque and Daget (1984). Howes (1984) finally transferred the species to *Mesobola brevianalis* based on osteological characters and this remains its current taxonomic status (Skelton 2001; Eschmeyer 2014). Skelton & Whitfield (1989) found an abundance of this species in the Mkhuze River and they were found in small shoals in slack regions of the river during field work for this study.

Engraulicypris whitei* van der Horst 1934** was described from the Aapies River near Petronella and Hamanskraal in the Transvaal. This description was later supported by a number of authors (Barnard 1943; Poll 1945; Farquharson 1962; Whitehead 1962). This species was only considered as a junior synonym of *Engraulicypris brevianalis,** for unstated reasons, by Jubb (1963) and later supported by Jubb (1967), Fowler (1977), and then supported as a junior synonym for *Neobola brevianalis* by Lévêque and Daget (1984). This was finally transferred to *Mesobola brevianalis* (Howes 1984); this is the species' current status (Eschmeyer 2014).

Engraulicypris gariepinus **Barnard 1943** was described from the lower Orange River in South Africa (Table. 1.1) based predominantly on meristic characters. This was supported later by Farquharson (1962). The species was then recognized as a junior synonym of *Engraulicypris brevianalis* for unstated reasons by Jubb (1963) and supported by Jubb & Farquharson (1965), Jubb (1965) and Fowler (1977) and later by Lévêque and Daget (1984) as a junior synonym for *Neobola brevianalis*. It was subsequently transferred to *Mesobola brevianalis* (Howes 1984); this is its current taxonomic status (Eschmeyer 2014).

Howes (1984) created the genus *Mesobola* to house *M. brevianalis* because of unique external and osteological characters, based on cranium, jaw bone and jaw muscle morphology. Howes classified the genera of the African radiation within the subfamily Danioninae into three tribes based on their external and osteological features (Howes 1980, 1984). Similarly, related genera like *Neobola* (Vinciguerra 1895), *Rastrineobola* (Fowler 1936), *Engraulicypris* (Günther 1894) and *Chelaethiops* (Boulenger 1899) were found to be osteologically distinctive (Howes 1984). Howes (1980; 1984) placed the southern African Danioninae (Rasborinae) into two tribes, the members of which changed over a number of his studies. Howes (1980) placed *Engraulicypris*, *Opsariichthys*, *Zacco, Luciosoma* and *Barilius* within a monophyletic group known as the 'bariliine group' and *Neobola*, *Rastrineobola* and *Chelaethiops* within a monophyletic group, the 'neobolines'. Later, he (Howes 1984) moved *Engraulicypris, Leptocypris* and *Mesobola* into the neoboline group based on shared osteological characteristics and *Opsariichthys*, *Zacco*, *Opsaridium* and *Barilius* into a separate bariliine group (Howes 1984, Howes 1991). In more recent morphometric and molecular studies this cyprinid subfamily has been subdivided into tribes known as the Chedrini, Danionini and Rasborini (Tang *et al*. 2010), which greatly diverges from Howes' (1984) concepts. Tang *et al*. (2010) found that, based on molecular data, the Chedrini, also known as Bariliini, are sister to a Danionini-Rasborini clade and the African members are a monophyletic clade (Fig. 1.1). This is different from the three separate dispersals from Asia hypothesised by Howes (1984). *Mesobola*, *Chelaethiops* and *Neobola* did not form a monophyletic group, nor were they sister taxa (Tang *et al*. 2010). However, no specimens of *Rastrineobola* or *Engraulicypris* were included in the molecular study (Tang *et al*. 2010) and, as a result, the exact relationships of the genus *Mesobola* within the Danioninae could not be found (Fig. 1.1). Therefore, a study that includes these genera is needed to determine the phylogenetic position of *Mesobola*.

Neobola contains four species; *N. bottegoi* (Vinciguerra 1895) *N. fluviatilis* (Whitehead 1962), *N. nilotica* (Werner 1919) and *N. stellae* (Worthington 1932); *Chelaethiops* contains five species, *C. bibie* (Joannis 1835), *C. congicus* (Nichols & Griscom 1917), *C. rukwaensis* (Ricardo 1939), *C. elongatus* (Boulenger 1899) and *C. minutus* (Boulenger 1906); *Rastrineobola* contains one species; *R. argentea* (Pellegrin 1904); and *Engraulicypris* contains one species; *E. sardella* (Günther 1868) (Froese & Pauly 2014). All of these taxa would be suitable outgroups for phylogenetic analysis of the populations ascribed to *M. brevianalis*. The relationships of these genera are not clear, even after phylogenetic analysis of molecular data (Fig. 1.1), partly because some taxa were not represented and partly due to weak support (Tang *et al*. 2010). More distant relatives in the genera *Chelaethiops* and *Opsaridium* would be useful in securing the root of the entire phylogeny (Fig. 1.1).

A study is therefore required to assess the taxonomy of biogeographically distinct populations of *M. brevianalis.* It is important to use more than one method to assess fish stock structure as different methods can produce different results (Wessels 2009). Different methods would enable the cross-validation of results and maximize the likelihood of correctly identifying different stocks. A study including both morphometric and genetic methods will allow for a comprehensive study of this currently widely-distributed freshwater fish. Ward *et al*. (2005) suggested that a fish barcode database based on the sequence of the cytochrome oxidase (COI or cox1) gene was required for the identification of fish species with a high degree of confidence. The correct use of barcoding methods enabled the identification of seven species mistakenly classified as one species that was being highly exploited in the Amazon River (Ardura *et al*. 2010). This allowed for advances in resource conservation and sustainability of the new species. Performing genetic structure studies on populations to implement and maintain successful conservation and management of species is very important (Engelbrecht & Mulder 1999; Ward *et al*. 2005; Costa & Carvalho 2007; Swartz *et al*. 2008).

Phenotypic variation for the identification of fish species and the measurement of relationships between and among populations (Turan 1999), as seen in *M. brevianalis,* has been used for many decades (Cadrin 2000; Wessels 2009). The use of morphometry, particularly advancements in truss network systems (Strauss & Bookstein 1982) and morphometric analysis software, has enabled the identification of distinct morphological markers with which to identify species and taxa within the field (Wessels 2009).

1.1.1 Scope

This thesis will include *Mesobola brevianalis* and representatives of the allied genera *Engraulicypris*, *Neobola* and *Rastrineobola* as comparative material and outgroups such as *Raiamas* spp., *Chelaethiops* spp. and *Opsaridium* spp. Technically, the work will focus on a DNA-based assessment of the status of the populations, a morphometric analysis to aid their identification, a study on the paleobiogeographical movement of this group, and the rediagnosis and description of new populations if required.

1.1.2 Aims of Project

This thesis will investigate both genetic and morphometric variation within and among populations of *M. brevianalis* from different river systems in southern Africa to understand their relationships. The aims of the study are therefore to:

- Analyse data from three genes from specimens of each population of *Mesobola brevianalis* to estimate their taxonomic and biogeographic relationships.
- Use morphometrics to identify phenotypic markers for the genetic clades and genetically distinct populations.
- Form a paleobiogeographical model for speciation and divergence.
- Revise the taxonomy of the populations to reflect these findings, which may entail describing new species or reinstating previous synonymized names.

This study may allow the discovery and description of new species and the identification of populations that require independent management plans.

1.2 References

ARDURA, A., LINDE, A.R., MOREIRA, J.C., GARCIA-VAZQUEZ, E. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. *Biological Conservation*, **143**: 1438-1443.

BAILEY, R.G., MATTHES, H. 1971. A new species of Engraulicypris (Cyprinidae) from Tanzania, East Africa. *Revue de Zoologie et de Botanique Africaines*, **83(1-2)**: 79-83.

BARNARD, K.H. 1943. Revisions of the indigenous freshwater fishes of the SW Cape Region. *Annals of South African Museum*, **36**: 101-262.

BARNARD, K.H. 1947. *A pictorial guide to South African fishes, marine and freshwater*. Maskew Miller Ltd. London. pp. 58.

BELL-CROSS, G. 1965. Additions and amendments to the checklist of the fishes of Zambia. *Puku*, **3**:29-43.

BELL-CROSS, G. 1976. The fishes of Rhodesia. *National Museums and Monuments of Rhodesia Information Circular*. Harare. pp. 262.

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(3)**: 148-155.

BOULENGER, G.A. 1899. Matériaux pour la faune du Congo. Poissons nouveaux du Congo. Cinquième Partie. Cyprins, Silures, Cyprinodontes, Acanthoptérygiens. *Annales du Musée du Congo (Ser. Zoology),* **1(5)**: 40-47.

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 the collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum*, **1(3)**: 219-235.

BOULENGER, G.D. 1911. *Catalogue of the fresh-water fishes of Africa in the British Museum (Natural History)*. London. v. 2: 529.

BOULENGER, G.A. 1916. *Catalogue of the fresh-water fishes of Africa in the British Museum (National History)*. London. v. 4: i-xxvii + 1-392.

BRUTON, M.N., KOK, H.M. 1980. *The freshwater fishes of Maputaland*. In Bruton, M.N., Cooper, K.H. 1980. Studies on the ecology of Maputaland. Rhodes University, Grahamstown and Natal Branch of the Wildlife Society of South Africa, Durban, South Africa. pp. 121.

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

COSTA, F.O., CARVALHO, G.R. 2007. The Barcode of Life Initiative: synopsis and prospective societal impacts of DNA barcoding of Fish. *Genomics, Society and Policy*, **3**: 29-40.

CRASS, R.S. 1960. Notes on the freshwater fishes of Natal with descriptions of four new species. *Annals of the Natal Museum*, **14(3)**: 405-458.

DAVID, L. 1936. Die Von Dr. Fritz Haas auf der schomburgk-Afrika-expedition (1931-32) gesammelten sãœsswasser-fische. *Senckenbergiana biologica*, **18(1-2)**: 72-74.

ENGELBRECHT, G.D., MULDER, P.F.S. 1999. Allozyme variation in the river sardine, *Mesobola brevianalis* (Pisces, Cyprinidae). *Water South Africa*, **25**: 293-296.

ESCHMEYER, W. N. (ed). *Catalog of Fishes*. California Academy of Sciences (http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 2014.

FARQUHARSON, F.L. 1962. The distribution of Cyprinids in South Africa. *Annals of Cape Provincial Museums*, **2**: 23-251.

FOWLER, H.W. 1936. Zoological results of the George Vanderbilt African Expedition of 1934. Part III, The fresh water fishes. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **88**: 243-335.

FOWLER, H.W. 1977. A catalog of world fishes. *Quarterly Journal of the Taiwan Museum*, **30**: 88.

FROESE, R., PAULY, D. (eds). *FishBase*. World Wide Web electronic publication. (www.fishbase.org). Electronic version accessed 2014.

GAIGHER, I.G., POTT, R.M.C. 1973. Distribution of fishes in southern Africa. *South African Journal of Sciences*, **69(1)**: 25-27.

GILCHRIST, J.D.F., THOMPSON, W.W. 1913. The freshwater fishes of South Africa. *Annals of the South African Museum*, **7(5)**: 321-463.

GÜNTHER, A. 1868. *Catalogue of the fishes in the British Museum. Catalogue of the Physostomi, Containing the Families Heteropygii, Cyprinidae, Gonorhynchidae, Hyodontidae, Osteoglossidae, Clupeidae, Chirocentridae, Alepocephalidae,* *Notopteridae, Halosauridae, in the Collection of the British Museum, vol 7*. Taylor and Francis. London. p. 292.

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*. **4**: 616-628.

HAY, C.J., NÆSJE, T.F., THORSTAD, E.B. 2008. Fish populations, gill net catches and gill net selectivity in the Kunene River, Namibia. *NINA Report,* **325**: 98.

HOWES, G.J. 1980. The anatomy, phylogeny and classification of bariliine cyprinid fishes. *Bulletin of British Museum of Natural History (Zoology)*, **37(3)**: 129-198.

HOWES, G.J. 1984. A review of the anatomy, taxonomy, phylogeny and biogeography of the African neoboline cyprinid fishes. *Bulletin of the British Museum of Natural History (Zoology)*, **47(3)**: 151-185.

HOWES, G.J. 1991. *Systematics and biogeography: an overview*. In: Winfield, I.J., Nelson, J.S. (Eds.), Cyprinid Fishes: Systematics, Biology and Exploitation. Chapman and Hall. London. pp. 1–33.

JOANNIS, L. 1835. Observations sur les poissons du Nil, et description de plusieurs espèces nouvelles. *Magasin de Zoologie 1835*, **5**: 53.

JUBB, R.A. 1961. *An illustrated guide to the freshwater fishes of the Zambezi River, Lake Kariba, Pungwe, Sabi, Lundi, and Limpopo Rivers*. Stuart Manning Ltd. Bulawayo. pp. 107.

JUBB, R.A. 1963. A revised list of the freshwater fishes of southern Africa. *Annals of Cape Provincial Museums*, **3**: 5-39.

JUBB, R.A. 1965. Freshwater fishes of the Cape Province. *Annals of Cape Provincial Museums*, **4**: 1-72.

JUBB, R.A. 1967. *Freshwater fishes of southern Africa.* Gothic Printed Company Limited. Cape Town. South Africa. pp.127.

JUBB, R.A., FARQUHARSON, F.L. 1965. The freshwater fishes of the Orange River drainage basin. *South African Journal of Science*, **61(3)**: 118-125.

JUBB, R.A., GAIGHER, I.G. 1971. Checklist of the fishes of Botswana. *Arnoldia, Rhodesia*, **5(7)**: 1-22.

KADYE, W.T. 2008. The application of a Fish Assemblage Integrity Index (FAII) in a Southern African river system. *Water SA*, **34(1)**: 25-32.

LÉVÊQUE, C., DAGET, J. 1984. *Cyprinidae*. In: Daget, J., Gosse, J.P., van den Audenaerde, D.F. (eds). Check List of the Freshwater Fishes of Africa. MRAC, Tervuren & ORSTOM. Paris, France. pp 217–342.

NICHOLS, J.T., GRISCOM, L. 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(25)**: 653-756.

OLIVIER, P.A.S., LUUS-POWELL, W.J., SAAYMAN, J.E. 2009. Report on some monogenean and clinostomid infestations of freshwater fish and waterbird hosts in Middle Letaba Dam, Limpopo Province, South Africa. *Onderstepoort Journal of Veterinary Research*, **76**: 187-199.

PELLEGRIN, J. 1904. Diagnoses préliminaires de Poissons nouveaux du Lac Victoria, recueillis par M. Alluaud. *Bulletin de la Société Zoologique de France*, **29(8)**: 184-186.

POLL, M. 1945. Description de cinq espãšces nouvelles de Cyprinidae de Congo belge appartenant aux genres *Barbus* et *Engraulicypris*. *Revue de Zoologie et de botanique africaines*, **38(3-4)**: 298-311.

RAMOLLO, P.P. 2011. Freshwater fish abundance and distribution in the Orange River, South Africa. *Journal of Fisheries International,* **6(1)**: 13-17.

RICARDO, C.K. 1939. The fishes of Lake Rukwa. *The Journal of the Linnaean Society of London (Zoology),* **40 (275)**: 625-657.

SKELTON, P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town, South Africa. pp. 119-120.

SKELTON, P.H., WHITFIELD, A.K. 1989. Distribution and diversity of Mkhuze swamp fishes during a summer flood. *Southern African Journal of Aquatic Sciences*, **15(1)**: 50-66.

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

SWARTZ, E.R., MWALE, M., HANNER, R. 2008. A role for barcoding in the study of African fish diversity and conservation. *South African Journal of Science*, **104**: 293-298.

TANG, K.L., AGNEW, M.K., HIRT, M.V., SADO, T., SCHNEIDER, L.M., FREYHOF, J., SULAIMAN, Z., SWARTZ, E., VIDTHAYANON, C., MIYA, M., SAITOH, K., SIMONS, A.M., WOOD, R.M., MAYDEN, R.L. 2010. Systematics of the subfamily

Danioninae (Teleostei: Cypriniformes: Cyprinidae). *Molecular Phylogenetics and Evolution*, **57**: 189-214.

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

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

VINCIGUERRA, D. 1895. Esplorazione del Giuba e dei suoi affluent compiuta dal Cap. V. Bottego durante gli anni 1892-93 sotto gli auspicii della Società geografica Italiana. III. Pesci. *Annali del Museo Civico di Storia Naturale di Genova (Serie 2),* **15**: 21-60.

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 of Biological Sciences*, **360**: 1847-1857.

WARD, R.D., HANNER, R., HEBERT, P.D.N. 2009. The campaign to DNA barcode all fishes, FISH-BOL (review paper). *Journal of Fish Biology*, **74**: 329-356.

WERNER, F. 1919. Wissenschaftliche Ergebnisse der mit Unterstützung der Kaiserlichen Akademie der Wissenschaften in Wien aus der Erbschaft Treitl von F. Werner unternommenen zoologischen Expedition nach dem Anglo-Ägyptischen Sudan (Kordofan) 1914. IV. *Bearbeitung der Fische, Amphibien und Reptilien. Denkschriften, Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschafliche Klasse*, **96**: 437-452.

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

WHITEHEAD, P.J.P. 1962. Two new river fishes from eastern Kenya. *Annali del Museo Civico di Storia Naturale 'Giacomo Dorai'*, **73**: 98-108.

WORTHINGTON, E.B. 1932. Scientific results of the Cambridge Expedition to the East African lakes, 1930 (fishes other than Cichlidae). *The Journal of the Linnean Society of London. Zoology*, **38(258)**: 121-134.

Table 1.1: Species described under the genus *Mesobola* (Froese & Pauly 2014).

Scientific name	Author	Valid name	English name
Mesobola bredoi	(Poll, 1945)	*Mesobola bredoi*	
Mesobola brevianalis *=M. gariepinus* *=M. whitei*	(Boulenger, 1908) (Barnard, 1943) van der Horst (1934)	*Mesobola brevianalis*	River sardine
Mesobola moeruensis	(Boulenger, 1915)	*Mesobola moeruensis*	
Mesobola spinifer	(Bailey & Matthes, 1971)	*Mesobola spinifer*	Malagarasi sardine

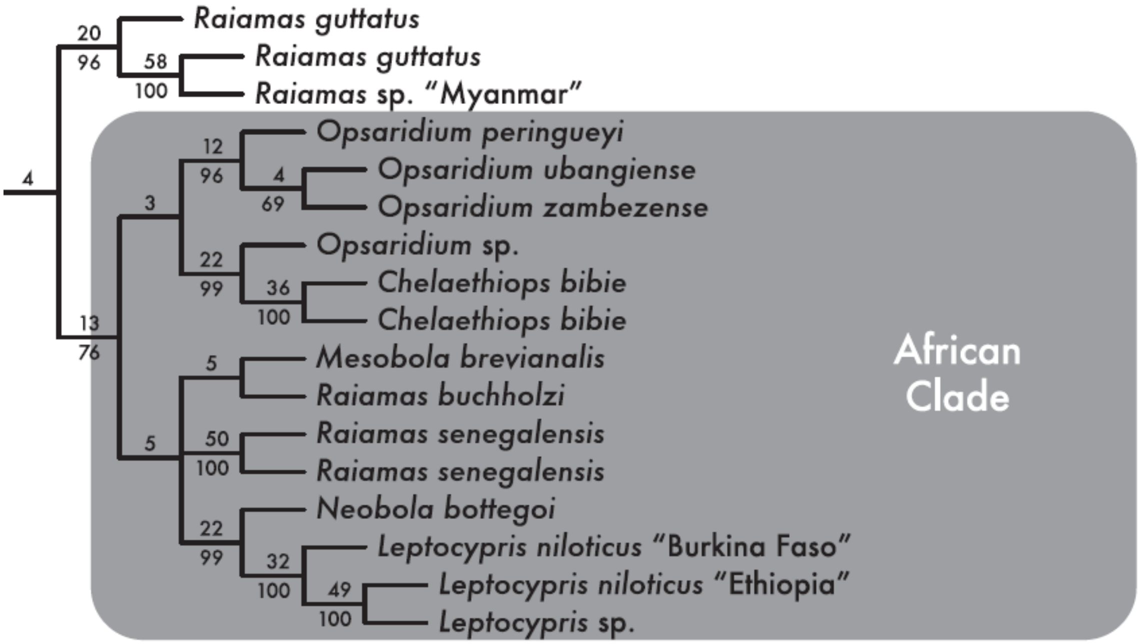


Fig. 1.1. A portion of Fig. 1b from Tang *et al*. (2010) indicating the phylogenetic relationships of *Mesobola* based on a total of 4117 bp of DNA sequence data from the mitochondrial genes cytochrome b (CytB) and cytochrome c oxidase 1 (COI), and the nuclear genes opsin (Opsin) and recombination activating gene 1 (RAG1). Bremer support values are shown above the relevant branches and bootstrap values below them. The phylogenetic position of *Mesobola* is not well supported relative to the rest of the clade.

2 A phylogenetic study of *Mesobola brevianalis* (Teleostei: Cyprinidae) and its relatives within the southern African Chedrini.

2.1 Introduction

The identification and classification of fish species has traditionally been largely based on external morphological features (Teletchea 2009). Although populations of *M. brevianalis* in different river systems may appear to have slight physical differences, particularly in colouration, it is not known which populations might differ sufficiently to be considered a separate species or management unit. Unfortunately there is no uncontested guideline for what makes a species a species, and there are at least seven different schools of thought on the definition of what a species is (Mayden 1997; Hausdorf 2011).

Howes (1980, 1984) was one of the first authors to classify the group of fish that includes *Mesobola* on anatomical and osteological features. Howes (1980) considered the genera *Opsariichthys*, *Zacco*, *Luciosoma*, *Barilius* and *Engraulicypris* (then including *M. brevianalis*) to possess a group of synapomorphic characters that united them in a monophyletic clade which he informally named ‘the bariliine group’. Later, based on the identification of further osteological and external morphological characters, he placed the genera *Neobola, Chelaethiops*, *Mesobola* and *Rastrineobola* into a monophyletic group informally named the neobolines, while *Engraulicypris, Leptocypris, Raiamas* and *Opsaridium* were placed in the bariliine group (Howes 1984). He considered that *Engraulicypris* contained a single species, *E. sardella* (Howes 1980), which led to the transfer of *E. brevianalis* into the genus *Neobola* (Lévêque & Daget 1984) and then into a new genus, *Mesobola* (Howes 1984). Howes (1984), further postulated that the species currently within Africa dispersed in three separate events into three separate clades named the *Opsaridium-Barilius*, monophyletic *Raiamas sensu lato* and neoboline groups (*Chelaethiops*, *Mesobola*, *Neobola* and *Rastrineobola*) including *Engraulicypris* and *Leptocypris*.

Many populations of fish that look morphologically similar have proved to be genetically independent, which has important implications for their management and conservation (Krück *et al*. 2013). Methods involving DNA-based identification have been developed to address this problem, predominantly in the last 15 years (Teletchea 2009). The use of molecular tools has become an essential part of a number of disciplines including those surrounding systematics, taxonomy and conservation (Rubinoff 2006; Hajibabaei *et al*. 2007). These tools were recently used on the fishes related to *Mesobola,* with results finding aspects of Howes' (1984) assumptions and postulated relationships to be incorrect; the Neoboline group including *Chelaethiops*, *Neobola* and *Mesobola* did not form a monophyletic clade (Tang *et al*. 2010). Furthermore, Tang *et al.* (2010), found that *Chelaethiops* and *Neobola* did not appear to be sister taxa, and rather *Leptocypris* and *Neobola's* relationship placed them together within a clade (Tang *et al*. 2010). However, not all the close relatives of *Mesobola* were included in that study, particularly the genera *Rastrineobola* and *Engraulicypris,* so a second aim of this chapter is to present and assess molecular data of those genera and species.

The standardized use of the nucleotide sequence of the mitochondrial Cytochrome Oxidase 1 gene (COI) (a protein-coding mitochondrial gene), also referred to as 'cox1', for identification was first proposed by Hebert *et al.* (2003) as a cost-effective, reliable and relatively accessible molecular technique known as 'barcoding'. The Fish Barcode of Life Initiative is a project dedicated to establishing a standardized reference sequence library for all fish species in the world (FISH-BOL, www.fishbol.org) (Krück *et al.* 2013). In the initial application of the initiative, COI barcodes were found to separate 98% of marine fish (Krück *et al.* 2013). It was also proposed that the COI marker would be able to produce valuable information on molecular evolution and diversification (Hebert *et al*. 2003; Galtier *et al*. 2009), although this claim has since been considerably toned down. DNA barcoding as a method for species discovery has been criticized for its over-simplification of the biological identification problem, as using only a single mitochondrial marker gene may limit the ability to decipher genetic species identification (Krück *et al*. 2013). Later, studies found that a multi-gene approach is essential for identifying problematic fish species (Krück *et al*. 2013), and that it is essential to choose genetic markers that are suited to a particular study (Sunnucks 2000; Rubinoff 2006).

For this reason, a multi-gene approach using COI, Recombination Activating Gene 1 (RAG1) (a nuclear protein-coding gene) and 28S ribosomal RNA (28S) (a nuclear ribosomal gene) has been adopted to identify the relationships of *Mesobola brevianalis*. These markers were selected for their different mutation rates (Krück *et al*. 2013), which will provide different temporal windows into the organisms' ancestries, and because together they can indicate nuclear-cytoplasmic mismatches (e.g. horizontal gene transfer) and serve individually as independent phylogenetic cross-validation for the other genes. The COI marker was selected for its prior success in identifying other species and its significance for the Barcode of Life initiative (Hebert *et al*. 2003). COI has also been identified as a gene that evolves rapidly (Brown *et al*. 1979) and, therefore, can aid in refined identification of many fish species (Ward *et al*. 2005). To complement COI, RAG1 was selected for its slower divergence rate (Mauro *et al*. 2009) and 28S, which mutates even slower, was chosen to probe the deepest levels of divergence in this study (Hillis & Dixon 1991). Engelbrecht & Mulder (1999) performed the first study on the genetic structure of a river sardine population and found low variability within the population from the Olifants River in the Limpopo System. It was hypothesized that this low variability was the result of cycles of drought and rainfall which resulted in the formation of bottlenecks and founder effects (Engelbrecht & Mulder 1999).

The phylogenetic study within this thesis will use molecular data to identify the relationships among the populations currently identified as *Mesobola brevianalis*, and also to examine the relationship of *Mesobola* to other genera.

2.2 Materials and methods

2.2.1 Materials

Specimens of *Mesobola brevianalis* were collected from ten countries within Africa: Botswana, Central African Republic, Malawi, Mozambique, Namibia, South Africa, Swaziland, Tanzania, Uganda and Zambia. Twelve river systems were sampled: Congo, Kunene, Limpopo, Luapula, Mbuluzi, Nile, Nkomati, Orange, Pongolo, Rovuma, St Lucia and Zambezi (Fig. 2.1). The fish were caught by various methods including seine netting, electrofishing, gill netting and ichthyocide.

Fish that were selected for genetic analysis in the field were first anaesthetized by being placed into clove oil. When possible, photographs were taken of the left side of the fish at this stage to record its live colouration. If a tissue sample was not immediately removed for DNA analysis, the specimens were placed in propanol or ethanol and each sample of specimens was placed into a container with a waterproof label bearing the date, sample number, location, details of the capture and preservation, the sample number and the specimen number (Table 2.1).

In either the field or the laboratory, muscle tissue from the end point of the dorsal fin to the beginning of the caudal fin (tail fin), or the entire caudal fin, was removed and placed into 95% ethanol in a separate microcentrifuge tube and kept at temperatures of below freezing. The rest of the specimens was then placed directly into 10% formalin in the field and then transferred into 70% ethanol for long-term preservation once in the lab. The DNA tissue and the whole specimens were stored at the South African Institute for Aquatic Biodiversity (SAIAB), Grahamstown.

2.2.2 Deoxyribonucleic Acid (DNA) Sequencing

Prior to DNA extraction, each tissue sample was dried completely before being placed in a new microcentrifuge tube. DNA was extracted using the DNeasy® blood and tissue kit (Qiagen, Valencia, CA) and the NucleoSpin® Tissue kit (Machery-Nagel GmbH & Co. KG) following the standard protocol for animal tissue isolation. The incubation period was 12 h to allow for complete tissue digestion. The final dilution step was performed with 50 µl, rather than 200 µl, nuclease-free distilled H_20 during extraction with the DNeasy® blood and tissue kit to allow for a higher concentration of DNA material. Otherwise, standard protocol was followed.

The concentration and purity of each DNA sample was determined by using a NanoDrop 2000 Spectrophotometer. The DNA concentration, A260, A280, 260/280 and 260/230 values were documented to ensure the DNA was of a sufficiently high concentration and purity.

Three genes were selected for amplification: 28S (nuclear ribosomal gene), COI (mitochondrial protein-coding gene) and RAG1 (nuclear protein-coding gene). The primers used for the 28S PCR were 28S-2F and 28S-2R; for the COI PCR the primers used were LCOI490 and HCO2198; and for the RAG1 PCR, the primers RAG1F1 and RAG1R2 were used with the use of RAG1R3 with problematic samples. The primer names and sequences used were 28S-2F-5’ CTT GAA ACA CGG ACC AAG GA 3’; 28S-5’2R-GTA TAG GCA

CGA CGC TTC AG 3'; LCOI490-5' GGT CAA ATC ATA AAG ATA TTG G 3'; HCO2198-5' TAA ACT TCA GGG TGA CCA AAA AAT CA 3' (Folmer *et al*. 1994); RAG1F1- 5'CTG AGC TGC AGT CAG TAC CAT AAG ATG T 3'; RAG1R2- 5'TGA GCC TCC ATG AAC TTC TGA AGR TAY TT 3'; RAG1R3-5' GTC TTG TGS AGG TAG TTG GT 3' (López *et al*. 2004).

For amplification of the 28S and RAG1 genes, reactions were conducted in a total volume of 25 µl and included 0.5 µl DNA (100 ng), 1µl forward primer (10 pmol), 1µl reverse primer (10 pmol), 12.5 µl Master Mix (2x) (Promega) and 10 µl dH_2O. For amplification of the COI gene, reactions included 2-4 µl $MgCl_2$ (25 mM) (Promega). Amplifications were conducted in a Thermo Hybaid PX2 PCR machine.

The 28S gene has been found to be a helpful marker for phylogenetic analysis (Zhao *et al*. 2012). The primers 28S-2F and 28S-2R amplify an average of 800 basepairs. The PCR conditions for 28S amplification had a denaturation temperature of 95°C for a period of 5 min for a single cycle followed by an annealing of 95 °C for 45 s, 58 °C for 45 s and 72 °C for 2 min, repeated for 35 cycles. The final elongation stage was 72 °C for 5 min.

The LCOI490 and HCO2198 primer set amplifies 658 base pair fragments of the COI gene within a wide range of vertebrate species (Folmer *et al*. 1994). The PCR conditions for this gene fragment were 94 °C for 1 min, 45 °C for 1.5 min, 72 °C for 1.5 min, annealing of 94 °C for 1 min, 50 °C for 1.5 min and 72 °C for 1 min for 40 cycles and a final elongation stage at 72 °C for 5 min.

The primers RAG1F1 and RAG1R2 target the region between the nucleotide position 2215 and 3772 (1557bp) of the *O. mykiss* RAG1 sequence accessioned in GenBank (U15663) (López *et al*. 2004). The RAG1R3 primer was used where RAG1R2 gave unsatisfactory results as it is an internal primer and allowed for binding of a different, shorter region of the fragment of the DNA sequence. The PCR programme included a denaturation single cycle of 94 °C for 1 min, annealing of 94 °C for 1 min, 52 °C for 1 min and 72 °C for 2 min for 40 cycles and a final elongation cycle of 72 °C for 10 min.

Each of the PCR products was electrophoresed on a 1% agarose gel at 80 V for 30 min. Capillary electrophoresis was conducted by Macrogen using both the forward and reverse primers.

Nucleotide sequences were edited using BioEdit sequence alignment software (Hall 2004). The forward and reverse sequences were aligned using ClustalX multiple sequence alignment (Larkin *et al*. 2007) within BioEdit to form consensus sequences (Appendix A). All of the sequences for a marker, including outgroups, were aligned using ClustalX (Larkin *et al*. 2007) and saved in a nexus format.

2.2.3 Phylogenetic analysis

The outgroups for this phylogenetic study were chosen on the basis of the work of Howes (1980; 1984) and Tang *et al*. (2010). A total of nine outgroups were chosen within the Cyprinidae and the respective sequences were either generated from tissues or downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) (Table 2.2). They were aligned with the RAG1 and COI samples tested.

DAMBE (Data Analysis in Molecular Biology and Evolution) (Xia & Xie 2001) was used to create saturation plots (transition versus transversion graphs) for both the COI and the RAG1 aligned sequences to test whether mutational saturation had been reached or whether the entire data set could be used for confident analysis.

The aligned RAG1 and COI sequences with the respective GenBank aligned sequences were saved as Phylip files (.phy) and opened in Splitstree (Huson 1998) for Neighbour-Net analysis using K2P distances. The outgroups were chosen according to the initial relationship seen in Splitstree (Table 2.2).

MrModel Test (Nylander 2004) was used to access the model of best fit for the sequences using the Akaike Information Criterion weightings (Akaike1973). The model of best fit for the COI gene was calculated as TrN+I+G, while the model of best fit for the RAG1 gene was TrNef + G. The model of best fit was then used to build a Bayesian inference tree using MrBayes (Huelsenbeck & Ronquist 2001) and using a total of 1000 pseudoreplicate generations. Once completed, 20% burn-in was removed manually from the total data set. The Bayesian trees were viewed using TreeView (Page 1996) and refined further.

2.3 Results

Once sequencing and editing were complete, ~700 clear bases were received for the 28S gene; the COI gene returned ~660 clear bases; and the RAG1 gene returned ~1200 clear bases for analysis. For a number of the RAG1 and COI sequences, the alignment of the forward and reverse sequences was difficult as there were some degenerate bases within the sequences or the sequences were entirely unusable as fragments had not been amplified. A number of sequences for these genes had to be abandoned.

2.3.1 28S rRNA

The sequences generated for the 28S gene showed no more than five base differences even relative to the outgroup genera. No African danionine species' 28S sequences were available on GenBank. Although this gene shows that the species were all very closely related, it was not included in further analyses.

2.3.2 COI

The saturation plot for the COI dataset (Fig. 2.2.A) showed that saturation had not been reached and therefore the COI sequences were likely to be informative at all phylogenetic levels represented in the sample.

The COI Neighbour-Net trees generated in Splitstree showed similar topologies to trees generated by Bayesian inference analysis. There was a division of the species into two separate radiations (Fig. 2.3.A). The Congo River (Central African Republic) samples were included in a clade with one of the outgroup taxa, *Chelaethiops congicus*, with which they share a particularly close relationship (Fig. 2.3.A). The other species grouped within this subnetwork included *Chelaethiops bibie*, *Neobola bottegoi*, *Raiamas* species and the *Opsaridium* species (Fig. 2.3.A). The second group included species from further south in Africa: *Engraulicypris sardella*, *Rastrineobola argentea* and the *Mesobola brevianalis* populations (Fig. 2.3.A).

The Neighbour-Net analysis of COI showed a clearer view of the second group, which included *M. brevianalis* samples from Mozambique, Namibia, South Africa and Swaziland (Fig. 2.3.B). The *M. brevianalis* Kunene System (Namibia) samples showed an early origin (Fig. 2.3.B). The *M. brevianalis* Lucheringo and Rovuma River (Mozambique) samples showed a similar separate branching and, although they appeared on separate branches, they showed a very strong relationship that could place these two populations in a single clade (Fig. 2.3.B). The *M. brevianalis* Orange System specimens showed a distinct independence representative of a clade arising slightly later than the Kunene samples (Fig. 2.3.B). The final *M. brevianalis* group is the eastern South African material which ranges from St Lucia in the south to the Limpopo system in the north (Fig. 2.3.B).

R. argentea and *E. sardella* were on separate branches nested among the *M. brevianalis* populations (Fig. 2.3.B). *R. argentea* appeared to show a weakly supported relationship to *M. brevianalis* from the Orange, Kunene and eastern South African rivers, while *E. sardella* showed a well-supported relationship with *M. brevianalis* from Mozambique (Fig. 2.3.B). The inclusion of these groups within this branch places the monophyly of *Mesobola* in question (Fig. 2.3.B).

The relationships within the eastern South African *M. brevianalis* specimens for the COI gene can be seen in Fig. 2.3.C. The specimens from the western Limpopo form a tight clade (Fig. 2.3.C). The Pongolo system specimens were also grouped together in a tight clade but it is only moderately supported and attached with short pathways to the St Lucia, eastern Limpopo and Nkomati System populations (Fig. 2.3.C).

MrModelTest indicated that the model of best fit for the COI gene data was the TrN+I+G model with an AIC (Akaike 1973) score of 6627.79.

The Bayesian tree rooted on *Neobola bottegoi* placed the outgroups on a separate branch to the *M. brevianalis* samples with an early branching of *Chelaethiops*, then *Opsaridium* and finally *Raiamas* (Fig. 2.4). *Chelaethiops congicus* showed a close relationship with the Congo River (Central African Republic) samples with a posterior probability of 98.57% (Fig. 2.4). *Opsaridium* specimens were grouped together in an early branch with *Raiamas* and a strong posterior probability weighting of 86.12% (Fig. 2.4). The Kunene River *M. brevianalis* specimens were grouped together with a posterior probability of 100% and showed the first divergence from the other *Mesobola* species (Fig. 2.4). A four-way split within the population then occurred, placing the *M. brevianalis* Rovuma (with *E. sardella* from Lake Malawi), Orange, eastern South Africa and *Rastrineobola* samples on their own branch (Fig. 2.4). The Orange River specimens were grouped together with a posterior probability of 100% as were the *R. argentea* specimens (Fig. 2.4). *E. sardella* was grouped with the *M. brevianalis* specimens from the Rovuma System with posterior probability support of 73.41% (Fig. 2.4). The support for the Lake Chiuta and Rovuma *M. brevianalis* clade was 100% (Fig. 2.4). The *M. brevianalis* eastern South African specimens sampled were grouped together with specimens of *M. brevianalis* from the species' type locality (posterior probability = 100%) (Fig. 2.4). The Nkomati and Pongolo River populations were placed in their own clades, but the posterior probabilities are low at 52.44% and 57.76% respectively (Fig. 2.4). The western Limpopo population lay on a separate clade with a high posterior probability value of 91.85% within the eastern South African specimens (Fig. 2.4). The *M. brevianalis* Kunene System (Namibia), Nkomati System (Swaziland) and Orange System (Namibia) samples also formed distinct clades (Fig. 2.4).

2.3.3 RAG1

The saturation plot produced using the RAG1 sequence also showed that saturation had not been reached (Fig. 2.2.B).

The Neighbour-Net analysis of the RAG1 data using K2P distances showed a similar topology to the COI analyses. The Congo River (Central African Republic) samples were closely related to *Neobola bottegoi, Raiamas salmolucius, Opsaridium* species and other *Chelaethiops* species (Fig. 2.5.A). The *E. sardella* species branch first into a clade and then the *M. brevianalis* Rovuma samples (Fig. 2.5.A). The next populations to become distinct are the *M. brevianalis* specimens from the Orange (Namibia) and Kunene (Namibia) Systems and *R. argentea* (Uganda) (Fig. 2.5.A). The final group in the *Mesobola* clade consists of the specimens from eastern South Africa which lay with specimens from the type locality (Fig. 2.5.A). The relationships of the *M. brevianalis* populations can be seen more clearly in Fig. 2.5.B. Once again, the *M. brevianalis* Rovuma specimens made up of Mozambique (Lucheringo and Rovuma Rivers) and Malawi (Lake Chiuta) specimens are grouped within a defined clade that is sister to *E. sardella* from Lake Malawi (Fig. 2.5.B). The *M. brevianalis* specimens from the Orange and Kunene Systems are grouped with *R. argentea* (Fig. 2.5.B).

The populations of *M. brevianalis* from eastern South Africa form a distinct clade (Fig. 2.5.B).

A MrModelTest calculated that the model of best fit for the RAG1 gene was the TrNef+G with an AIC (Akaike 1973) weighting of 4894.13.

The Bayesian tree showed distinct relationships among the river system samples (Fig. 2.6). As expected, the *Chelaethiops* species were closely related to one another and are the closest relation to the Congo River samples, which are classified as *Chelaethiops* species (Fig. 2.6). There is a four-way split within the specimens broken into: two clades of *Chelaethiops;* a clade of *Opsaridium* + *Raiamas*; and the final clade including all *M. brevianalis*, *E. sardella* and *R. argentea* (Fig. 2.6). There is strong posterior probability support (99.43%) for three further groupings involving *E. sardella* from Lake Malawi*, M. brevianalis* from the Rovuma System, and a final group made up of *M. brevianalis* from eastern South Africa and the Kunene and Orange Rivers and *R. argentea* from Lake Victoria (Fig. 2.6). There is a strong relationship between the Orange River *M. brevianalis*, Kunene River *M. brevianalis* and *R. argentea* (Uganda) with a posterior probability of 96.48% (Fig. 2.6). The Orange River samples formed a clade with a moderate support value of 64.7% (Fig. 2.6). The *M. brevianalis* populations from eastern South Africa are grouped with moderate posterior probability support (Fig. 2.6). The Olifants River (eastern South Africa) population of *M. brevianalis* appeared to be distinct with regard to RAG1 but the relationship is not strongly supported (Fig. 2.6).

2.3.4 Total molecular evidence

The Neighbor-Net view of the total evidence analysis, including RAG1 and COI, showed a distinct separation of the samples into two sister clades, one including *Chelaethiops*, *Neobola*, *Raiamas* and *Opsaridium*, and the other including *M. brevianalis*, *R. argentea* and *E. sardella* (Fig. 2.7.A). Although not every specimen could be included in the total evidence analysis, each river system and taxon was represented at least once (Fig. 2.7.A). The Congo River samples aligned with *Chelaethiops* (Fig. 2.7.A). Once again these *Chelaethiops* species were aligned with the outgroup taxa including *O. splendens, O. peringueyi, O. microcephalus, R. salmolucius* and *N. bottegoi* (Fig. 2.7.A). Within the *Mesobola* clade, once again, six distinct sub-clades were formed (Fig. 2.7.A). The *M. brevianalis* Kunene System population was the first to separate, closely followed by the *M. brevianalis* Orange System population and then the *R. argentea* population (Fig. 2.7.A). The *M. brevianalis* Rovuma System specimens were closely related to *E. sardella,* as seen previously (Fig. 2.7.A). *M. brevianalis* specimens from the eastern South African systems form a strong cluster with some internal structure (Fig. 2.7.B). The western Limpopo specimens appear on a distinct but short branch with strong bootstrap support (Fig. 2.7.B). The specimens from the Pongolo and Mbuluzi systems form their own weakly supported clades (bootstrap value = 47.2%) (Fig. 2.7.B). The Nkomati system specimens, which included *M. brevianalis* from the Crocodile, Sabie and Mlumati Rivers, appear to group together with low to moderate support (Fig. 2.7.B). Finally, the eastern Limpopo and St Lucia system *M. brevianalis* formed a group that was less well defined than in other Neighbour-Net and Bayesian analyses with a weak relationship to one another and the Olifants specimens with a weaker grouping within (Fig. 2.7.B).

Bayesian analysis of the concatenated RAG1 and COI sequences (Fig. 2.8) showed very similar relationships to the Neighbour-Net analysis. The outgroups which included *Neobola*, *Chelaethiops*, *Opsaridium* and *Raiamas* were all grouped together with rather moderate posterior probability of 61.71% to the *M. brevianalis*, *Engraulicypris* and *Rastrineobola* populations (Fig. 2.8). The relationship of *E. sardella* and *R. argentea* and all of the populations identified as *M. brevianalis* was very strong, with a posterior probability of 99.91% (Fig. 2.8). Evidence of the same relationships were seen within the Neighbour-Net analysis, with a five-way split in the tree separating the populations into *M. brevianalis* Rovuma + *E. sardella*, *M. brevianalis* Kunene, *M. brevianalis* Orange, eastern South African *M. brevianalis* and *R. argentea* (Fig. 2.8). The grouping of *M. brevianalis* from the Rovuma and Lucheringo Rivers and Lake Chiuta (all within the Rovuma System) showed a very strong relationship to one another (Fig. 2.8). The *M. brevianalis* Rovuma specimens were then closely related to *E. sardella* with a slightly above average posterior probability of 76.28% (Fig. 2.8). The Kunene River *M. brevianalis* population was defined by a posterior probability of 99.85% (Fig. 2.8), and the Orange River *M. brevianalis* population had 100% support (Fig. 2.8). The *R. argentea* specimens formed a branch with perfect posterior probability support (Fig. 2.8). The final group was a very distinct clade of all of the *M. brevianalis* populations from the St Lucia system to the Limpopo system (Fig. 2.8). The posterior probability value for the clade was 100%, while subgrouping within this cluster showed lower support (Fig. 2.8). The western Limpopo specimens (Laphalala, Martin's Drift and North-West) were also confidently grouped (Fig. 2.8). The Pongolo, Olifants and Nkomati System populations were each distinct but weakly supported (Fig. 2.8). The eastern Limpopo specimens (Albasini Dam & Olifants River) were placed with the St Lucia population once again, showing no particular relationships with the Olifants River specimens grouped within their own clade with weak support (Fig. 2.8).

2.4 Discussion

The genetic information analysed in this study provided a clearer perspective of the southern African populations placed under the name *M. brevianalis* and successfully retrieved all of the other known species that were included.

Analyses of the individual genes and the total evidence showed consistent patterns of relationship. The 28S rRNA gene supported the general monophyly of the southern African danionines, which was unsurprising due to its slower mutation rate compared to that of mt-rRNA (Mindell & Honeycutt 1990). It is apparent that the 28S gene has not had sufficient time to evolve species-specific differences as the RAG1 and COI genes did. The COI gene evolves unusually rapidly, mutating five-to-ten times faster than single copy nuclear DNA (Brown *et al*. 1979). This evolution time has been estimated as occurring at 0.02 substitutions per base pair per million years for vertebrates (Brown *et al*. 1979). The high rate of evolution of mtDNA can be attributed to either high mutation rates, high fixation rates, or both of these (Brown *et al*. 1979).

Surprisingly, the results support much of Boulenger's (1911) original assessment of *sardella*, *brevianalis*, *argenteus* and *bottegoi* within a single genus, *Engraulicypris*. These results, however, contradict much of Howes' (1984) informal subdivision of the African danionines into neobolines and bariliines. Of his neobolines, *Chelaethiops* grouped with *Neobola* as expected but *Mesobola* and *Rastrineobola* were aligned with bariliine taxa. Of his bariliines, *Engraulicypris sardella* was placed as expected, but *Raiamas* and *Opsaridium* were grouped with *Neobola* (Howes 1984). Tang *et al*. (2010) found similar disparities, confirming the analysis presented here and suggesting that morphological, anatomical and osteological characters are more subject to homoplasious evolution than molecular characters (Teletchea 2009). A multi-locus approach provides a more reliable means of identifying lineages and relationships and resolving 'troublesome evolutionary relationships' (Crow *et al*. 2004; Chen *et al*. 2008; Tang *et al*. 2010; Krück *et al*. 2013). The support for a multi-locus approach has been seen directly in a study performed on hexagrammid fish in which combined multiple loci reduced the number of parsimonious trees to one and increased the bootstrap node values to highly significant values (Crow *et al*. 2004).

The COI (mitochondrial) and RAG1 (nuclear) gene trees placed the samples from the Congo River (Central African Republic) on a separate branch with a close relationship to *Chelaethiops* (Boulenger 1899) (*C. bibie* and particularly *C. congicus*) and more generally to the genera *Opsaridium*, *Neobola* and *Raiamas*. The very close relationship of these specimens to *Chelaethiops congicus* (Nichols & Griscom 1917) makes it highly probable that the specimens are of the same species, found within the Congo River system, particularly Lake Tanganyika (Nichols & Griscom 1917). The biological details of *Chelaethiops* and *Chelaethiops congicus* can be found in FishBase (Froese & Pauly 2014).

The results also strongly suggest that the populations currently identified as *M. brevianalis* represent several unrecognized species. One source of evidence is the strong support for the distinctiveness of populations such as those in the Kunene, Orange, and Rovuma Rivers. A second source of evidence is the inclusion of *R. argentea* and *E. sardella* within the *M. brevianalis* clade. If these populations are sufficiently distinctive to be recognized as species, then the well-supported branches around them would seem to deserve similar recognition.

Barnard (1943) described *Engraulicypris gariepinus* using meristics from the lower reaches of the river system below the Augrabies Falls (28°35′29″S 20°20′27″E) on the basis of three diagnostic morphological characteristics. This species was later synonymized with *E. brevianalis* by Jubb (1963) and finally into *M. brevianalis* by Howes (1984) which is its state presently. The phylogenetic results obtained within this study support Barnard's (1943) opinion and therefore it is suggested that the species name *M. gariepinus* should be resurrected. The distribution of the samples and those measured by Barnard (1943) were from regions below the Augrabies Falls as this species is only found below the falls and not within the upper Orange River (Skelton 1986; Skelton & Tweddle 1991; Ramollo 2011). Waterfalls often prevent upstream movement of fish and therefore can genetically isolate populations within a river (Crass 1969). This may be the reason for the isolation of this population on a separate evolutionary path from other populations of *M. brevianalis*.

Most of the well-supported clades have sister groups that are also geographical neighbours, which lends independent support to the phylogenetic analysis. However, the clade containing

R. argentea and the Orange River *M. brevianalis* is comparatively poorly supported. This poses an interesting question for evolutionary biogeography because *R. argentea* is found only in Lake Victoria, thousands of kilometres from the lower Orange River. This question will be revisited in Chapters 4 and 5.

Both *Engraulicypris* Günther, 1894 and *Rastrineobola* Fowler, 1936 are currently monotypic genera. The inclusion of *E. sardella* and *R. argentea* amongst the populations identified as *M. brevianalis* was consistent in all analyses and makes *Mesobola* Howes, 1984 a paraphyletic genus. The clade can be made monophyletic by synonymizing the genera. *Engraulicypris* is the oldest name, and therefore has priority, and the others may then be treated as junior synonyms. This would support Boulenger's (1911) initial inclusion of *sardella, argenteus* (*argentea*) and *brevianalis* within the genus *Engraulicypris*. The details of the nomenclature of *Rastrineobola*, *Mesobola* and *Engraulicypris* can be found in FishBase (Froese & Pauly 2014), and is dealt with again in Chapter 5.

Engraulicypris whitei was described from the Aapies River within the western Limpopo System (van der Horst 1934) but later synonymized with *E. brevianalis* (Boulenger 1908) by Jubb (1963) and finally as a junior synonym of *M. brevianalis* (Howes 1984). Specimens from the western Limpopo System populations of *M. brevianalis* were placed on strongly supported branches but they always lay within the eastern South African populations for both the Bayesian and Neighbour-Net analyses. These other populations showed less clear support for monophyly. The Neighbour-Net analysis suggests a short period of genetic coherence before divergence set in amongst these populations. Although differentiated, these populations are insufficiently genetically unique to be considered new species. Since this clade includes specimens from the type locality (Mkhuze River) of *M. brevianalis* (Boulenger 1908), it is suggested that the population should be placed under the species name *M. brevianalis*, and that *E. whitei* should be treated as its junior synonym as it stands today. This nomenclatural matter is considered again in Chapter 5.

The relative isolation of the *whitei* population may be due to isolation from the other populations by either distance or barriers (Engelbrecht & Mulder 1999). It may also be due to a genetic bottleneck caused by the repeated drying out of large sections of the upper Limpopo, which is often reduced to isolated pools in the dry seasons (Alemaw *et al.* 2013). Habitat fragmentation or barriers occurring within a system can limit gene flow resulting in reduced genetic diversity and increased genetic variation among small populations (Wofford *et al*. 2005). Alemaw *et al*. (2013) found the western regions of the Limpopo drainage system were more susceptible to drought than other regions of the system. Severe drought events in Limpopo could have resulted in the isolation of fish populations in the western regions resulting in the divergence within a single river system. This issue is discussed further in Chapter 4.

The low variation seen among the other eastern South African populations may be as a result of the proximity of the systems which may connect through flooding (Alemaw *et al*. 2013), or may be a result of inter-basin transfer which has been identified as an effective short-term measure for tackling water supply problems (Mukheibir & Sparks 2003). During field work for this study, a sugar cane farmer in KwaZulu-Natal spoke of irrigation arrangements that involved pumping water from Jozini Dam (also known as the Pongolopoort Dam; Pongolo System) via a number of holding dams to the Mkhuze River (St Lucia System). This extensive inter-basin transfer is part of in a large-scale irrigation scheme (Walmsley & Davies 1991) that could promote very recent gene flow between otherwise isolated river systems, creating low genetic variation.

Unfortunately, the RAG1 and COI genes of the Zambezi River samples could not be amplified. Certain genes may have very particular species-specific conditions for the PCR to succeed and this was true particularly of the RAG1 gene (Tang *et al*. 2010). The tissue samples from the Luapala River in Zambia did not show adequate DNA concentrations and purification after extraction and these samples, therefore, did not allow for genetic analysis. This low concentration can possibly be attributed to poor preservation in the field. Before about 2008, a number of different preservatives were used including formalin. The current successful method involves placing the vouchered tissue into 95% ethanol and then into a freezer. Prior to the adoption of this method, fluid-preserved museum samples were initially fixed in formalin and later transferred into alcohol (Palero *et al.* 2010). Formalin can directly modify DNA, which hinders the PCR (Palero *et al*. 2010).

2.5 Conclusion

It is apparent that *M. brevianalis* contains cryptic species, these being the populations from the Rovuma, Orange and Kunene River Systems. Finding three lineages within *M. brevianalis* brings into question whether they bear distinct morphological characters to further set them apart as unique taxa. The inclusion of the monotypic genera *Rastrineobola* and *Engraulicypris* suggests that these genera require morphological rediagnosis too. These goals will be addressed in the next chapter.

2.6 References

AKAIKE, H. 1973. *Information theory and an extension of the maximum likelihood principle*. In: Petrov BN, Csaki F, (eds). Second International Symposium on Information Theory. Budapest. Hungary. Akademiai Kiado. 267–281.

ALEMAW, B.F., KILESHYE-ONEMA, J.M., LOVE, D. 2013. Regional drought severity assessment at a basin scale in the Limpopo Drainage System. *Journal of Water Resource and Protection*, **5**: 1110-1116.

BARNARD, K.H. 1943. Revisions of the indigenous freshwater fishes of the SW Cape Region. *Annals of South African Museum*, **36**: 101-262.

BOULENGER, G.A. 1908. On the collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum*, **1(3)**: 219-235.

BOULENGER, G.D. 1911. *Catalogue of the fresh-water fishes of Africa in the British Museum (Natural History)*. London. v. 2: 529.

BROWN, W.M., GEORGE, M., WILSON, A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences*, **76(4)**: 1967-1971.

CHEN, W.J., MIYA, M., SAITOH, K., MAYDEN, R.L. 2008. Phylogenetic utility of two existing and four novel nuclear gene loci in reconstructing Tree of Life of ray-finned fishes: The order Cypriniformes (*Ostariophysi*) as a case study. *Gene*, **423**: 125-134.

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.

CROW, K.D., KANAMOTO, Z., BERNARDI, G. 2004. Molecular phylogeny of the hexagrammid fishes using a multi-locus approach. Molecular *Phylogenetics and Evolution*, **32**:986-997.

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.

FOLMER, O., BLACK, M., HOEH, W., LUTZ, R., VRIJNHOEK, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology Biotechnology*, **3**: 294-297.

FOWLER, H.W. 1936. Zoological results of the George Vanderbilt African Expedition of 1934. Part III, The fresh water fishes. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **88**: 243-335.

FROESE, R., PAULY, D. (eds). *FishBase*. World Wide Web electronic publication. (www.fishbase.org). Electronic version accessed 2014.

GALTIER, N., NABHOLZ, B., GLÉMIN, S., HURST, G.D.D. 2009. Mitochondrial DNA as a marker for molecular diversity: a reappraisal. *Molecular Ecology*, **18**: 4541-4550.

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*. **4**: 616-628.

HAJIBABAEI, M., SINGER, G.A.C., HEBERT, P.D.N., HICKEY, D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and populations genetics. *TRENDS in Genetics*, **23(4)**: 167-172.

HALL, T. 2004. *BioEdit v. 7.0.1*. Department of Microbiology, North Carolina State University. (www.mbio.ncsu.edu/BioEdit/bioedit.html). Electronic version accessed 2012.

HAUSDORF, B. 2011. Progress toward a general species concept. *Evolution*, **65 (4)**: 923-931.

HEBERT, P.D.N, CYWINSKA, A., BALL, S.L., DEWAARD, J.R. 2003. Biological identifications through DNA Barcodes. *Proceedings of the Royal Society of London Biological Sciences*, **270(1512)**: 313-321.

HILLIS, D.M., DIXON, M.T. 1991. Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. *The Quarterly Review of Biology*, **66(4)**: 411-453.

HOWES, G.J. 1980. The anatomy, phylogeny and classification of bariliine Cyprinid fish. *Bulletin of the British Museum of Natural History*, **37(3)**: 129-198.

HOWES, G.J. 1984. A review of the anatomy, taxonomy, phylogeny and biogeography of the African neoboline cyprinid fishes. *Bulletin of the British Museum of Natural History,* **47(3)**: 151-185.

HUELSENBECK, J.P., RONQUIST, F. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, **17(8)**: 754–755.

HUSON, D.H. 1998. Splitstree a program for analyzing and visualizing evolutionary data. *Bioinformatics*, **14(1)**: 68-73.

JUBB, R.A. 1963. A revised list of the freshwater fishes of southern Africa. *Annals of Cape Provincial Museums*, **3**: 5-39.

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.

LARKIN, M.A., BLACKSHIELDS, G., BROWN, N.P., CHENNA, R., MCGETTIGAN, P.A., MCWILLIAN, 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(21)**: 2947-2948.

LÉVÊQUE, C., DAGET, J. 1984. *Cyprinidae*. In: Daget, J., Gosse, J.P., van den Audenaerde, D.F. (eds). Check List of the Freshwater Fishes of Africa. MRAC, Tervuren & Orstom. Paris. France. pp 217–342.

ĹOPEZ, J.A., CHEN, W.J., ORTI, G. 2004. Esociform Phylogeny. *Copeia*, **3**: 449-464.

MAURO, D.S., GOWER, D.J., MASSINGHAM, T., WILKINSON, M., ZARDOYA, R., COTTON, J.A. 2009. Experimental design in Caecilian systematics: phylogenetic information of mitochondrial genomes and nuclear Rag1. *Systematic Biology*, **58(4)**: 425-438.

MAYDEN, R.L. 1997. *A hierarchy of species concepts: the denouement in the saga of the species problem*. In: Claridge, M.F., Dawah, H.A. and Wilson, M.R. (eds). Species: The Units of Biodiversity. Chapman and Hall, London. pp. 381-424.

MINDELL, D.P., HONEYCUTT, M.R. 1990. Ribosomal RNA in Vertebrates: Evolution and Phylogenetic Applications. *Annual Review of Ecology and Systematics*, **21**:541-566.

MUKHEIBIR, P., SPARKS, D. 2003. *Water resource management and climate change in South Africa: Visions, driving factors and sustainable development indicators. Report for Phase 1 of the Sustainable Development and Climate Change Project*. Energy & Development Research Centre. University of Cape Town. Cape Town. South Africa.

NICHOLS, J.T., GRISCOM, L. 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(25)**: 653-756.

NYLANDER, J.A.A. 2004. *MrModelTest 2.0.* Program distributed by the author. Evolutionary Biology Centre, Uppsala University. (http://www.ebc.uu.se/systzoo/staff/nylander.html). Electronic version accessed 2013.

PAGE, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Bioscience*, **12**:357–358.

PALERO, F., HALL, S., CLARK, P.F., JOHNSTON, D., MACKENZIE-DODDS, J., THATJE, S. 2010. DNA extraction from formalin-fixed tissue: new light from the deep sea. *Scientia Marina*, **74(3)**: 465-470.

RAMOLLO, P.P. 2011. Freshwater Fish Abundance and Distribution in the Orange River, South Africa. *Journal of Fisheries International*, **6(1)**: 13-17.

RUBINOFF, D. 2006. Utility of mitochondrial DNA barcodes in species conservation. *Conservation Biology*, **20(4)**: 1026-1033.

SKELTON, P.H. 1986. *Fish of the Orange-Vaal System*. In: Davies, B.R., Walker, K.F. (eds.) *The Ecology of the River Systems*. Dr W. Junk Publishers. Dordrecht. Netherlands. pp. 143-161.

SKELTON, P.H., TWEDDLE, D. 1991. *Cyprinids of Africa*. In: Winfield, I.J., Nelson, J.S. The biology of cyprinid fishes. Chapman & Hall, London, England. pp. 211-239.

SUNNUCKS, P. 2000. Efficient genetic markers for population biology. *Ecology and Evolution*, **15(5)**: 199-203.

TANG, K.L., AGNEW, M.K., HIRT, M.V., SADO, T., SCHNEIDER, L.M., FREYHOF, J., SULAIMAN, Z., SWARTZ, E., VIDTHAYANON, C., MIYA, M., SAITOH, K., SIMONS, A.M., WOOD, R.M., MAYDEN, R.L. 2010. Systematics of the subfamily Danioninae (Teleostei: Cypriniformes: Cyprinidae). *Molecular Phylogenetics and Evolution,* **57**: 189-214.

TELETCHEA, F. 2009. Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries*, **19(3)**: 265-293.

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

WALMSLEY, R.D., DAVIES, B.R. 1991. An overview of water for environmental management. *Water SA*, **17(1)**: 67-76.

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 of Biological Sciences*, **360**: 1847-1857.

WOFFORD, J.E.B., GRESSWELL, R.E., BANKS, M.A. 2005. Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecological Applications,* **15(2)**: 628-637.

XIA, X., XIE, Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity*, **92(4)**: 371-373.

ZHAO, Y.E., WU, L.P., WU, L., XU, Y., WANG, Z.H., LIU, W.Y. 2012. Sequencing for complete rDNA sequences (18S, ITS1, 5.8S, ITS2, and 28S rDNA) of *Demodex* and phylogenetic analysis of Acari based on 18S and 28S rDNA. *Parasitology Research*, **111**: 2109-2114.

Table 2.1. Sample catalogue numbers and locality information.

Catalogue Number	***Full Name***	***Start Date***	***River***	***Latitude1***	***Longitude1***	***Country***	***System***	***Event Number***
Uncatalogued	*Chelaethiops congicus*	2010/09/10	Malagarasi	-5.4164	30.0628	Tanzania	Congo	MGSI 06
76771	*Chelaethiops congicus*	2005/09/23	Lwela	-11.5261	29.1511	Zambia	Luapula	JPF-05-011
77701	*Chelaethiops sp.*	2006/02/24	Congo	5.86	20.73	CAR	Congo	JPF-06-005
Uncatalogued	*Engraulicypris sardella*	2013/03/10	Lake Malawi	-14.1200	34.9346	Malawi	Zambezi	DT13-M066
78759	*Mesobola brevianalis*	2006/08/21	Kunene	-17.41	14.22	Namibia	Kunene	ES 06 A 54
78685	*Mesobola brevianalis*	2006/09/20	Luvhuvhu	-22.90	30.70	South Africa	Limpopo	IRB-06-22
101196	*Mesobola brevianalis*	2013/04/10	Martin's Drift	-22.9978	27.9408	Botswana	Limpopo	RB13-Limpopo1
190710	*Mesobola brevianalis*	2013/06/02	Olifants	-24.1851	30.8233	South Africa	Limpopo	RB13-Mes26
187259	*Mesobola brevianalis*	2012/03/20	North-West	-25.6456	26.4289	South Africa	Limpopo	KW12-AT403
187259	*Mesobola brevianalis*	2012/03/21	North-West	-25.6456	26.4289	South Africa	Limpopo	KW12-AT410
82589	*Mesobola brevianalis*	2009/02/05	Albasini Dam	-23.1033	30.1217	South Africa	Limpopo	83
190657	*Mesobola brevianalis*	2013/05/30	White Mbuluzi	-26.1683	31.8790	Swaziland	Mbuluzi	RB13-Mes19
66145	*Mesobola brevianalis*	2002/07/09	Mlumati	-25.76	31.44	Swaziland	Nkomati	S7
190621	*Mesobola brevianalis*	2013/05/30	Mlumati	-25.6840	31.565	South Africa	Nkomati	RB13-Mes21
190665	*Mesobola brevianalis*	2013/05/32	Sabie	-25.0192	31.206	South Africa	Nkomati	RB13-Mes23
78805	*Mesobola brevianalis*	2006/10/18	Orange	-28.8681	18.61	Namibia	Orange	IRB-06-01
78822	*Mesobola brevianalis*	2006/10/20	Orange	-28.6886	17.56	Namibia	Orange	IRB-06-03
78831	*Mesobola brevianalis*	2006/10/21	Orange	-28.7472	17.6089	Namibia	Orange	IRB-06-04
66270	*Mesobola brevianalis*	2003/03/31	Usutu	-26.8575	31.9078	Swaziland	Pongolo	S30
Uncatalogued	*Mesobola brevianalis*	2013/05/27	Pongolo	-27.3458	31.7499	South Africa	Pongolo	RB13-Mes06
Uncatalogued	*Mesobola brevianalis*	2013/05/29	Usutu	-26.8578	31.9076	Swaziland	Pongolo	RB13-Mes16
191029	*Mesobola brevianalis*	2013/03/15	Lake Chiuta	-14.9139	36.022778	Malawi	Rovuma	DT13-M100
73944	*Mesobola brevianalis*	2003/08/19	Rovuma	-12.5964	36.9356	Mozambique	Rovuma	N22
74087	*Mesobola brevianalis*	2003/08/25	Lucheringo	-11.8156	36.2208	Mozambique	Rovuma	N39
190719	*Mesobola brevianalis*	2013/05/26	Hluhluwe	-28.149	32.2823	South Africa	St Lucia	RB13-Mes04

Catalogue Number	***Full Name***	***Start Date***	***River***	***Latitude1***	***Longitude1***	***Country***	***System***	***Event Number***
190643	*Mesobola brevianalis*	2013/05/27	Mkhuze	-27.6069	32.0446	South Africa	St Lucia	RB13-Mes05
88674	*Mesobola brevianalis*	2011/03/09	Mkhuze	-27.6652	32.3021	South Africa	St Lucia	PM09A211
88674	*Mesobola brevianalis*	2011/03/10	Mkhuze	-27.6652	32.3021	South Africa	St Lucia	PM09A214
66270	*Mesobola brevianalis*	2003/03/31	Usutu	-26.86	31.91	Swaziland	Pongolo	S30
78627	*Mesobola brevianalis*	2006/09/17	Laphalala	-23.5794	28.1156	South Africa	Limpopo	IRB169
190670	*Mesobola brevianalis*	2013/05/31	Crocodile	-25.5255	31.3295	South Africa	Nkomati	RB13-Mes22
190682	*Mesobola brevianalis*	2013/05/28	Pongolo	-27.4028	31.7001	South Africa	Pongolo	RB13-Mes11
190635	*Mesobola brevianalis*	2013/05/29	Mtindzekwa	-26.7486	31.8369	Swaziland	Pongolo	RB13-Mes18
190676	*Mesobola brevianalis*	2013/05/26	Mfolozi	-28.3856	32.0385	South Africa	St Lucia	RB13-Mes02
73944	*Mesobola* cf. *brevianalis*	2003/08/19	Rovuma	-12.60	36.94	Mozambique	Rovuma	N22
74087	*Mesobola* cf. *brevianalis*	2003/08/25	Lucheringo	-11.82	36.22	Mozambique	Rovuma	N39
76771	*Mesobola* sp.	2005/09/23	Lwela	-11.53	29.15	Zambia	Luapula	JPF-05-011
Uncatalogued	*Opsaridium* 'bartail'	2009/02/02	Malagarasi	-5.1783	30.0742	Tanzania	Congo	MGSI 02
Uncatalogued	*Opsaridium microcephalus*	2013/03/11	Lake Malawi	-14.1201	34.9346	Malawi	Zambezi	DT13-M022
190654	*Opsaridium peringueyi*	2013/06/01	Olifants River	-24.3796	30.6656	South Africa	Limpopo	RB13-Mes24
190656	*Opsaridium peringueyi*	2013/05/29	White Mbuluzi	-26.1683	31.8790	Swaziland	Mbuluzi	RB13-Mes19
Uncatalogued	*Opsaridium peringueyi*	2013/05/30	Crocodile	-25.5254	31.3295	South Africa	Nkomati	RB13-Mes22
190666	*Opsaridium peringueyi*	2013/05/31	Sabie	-25.0192	31.205	South Africa	Nkomati	RB13-Mes23
190607	*Opsaridium peringueyi*	2013/05/28	Pongolo	-27.4170	31.6253	South Africa	Pongolo	RB13-Mes08
190605	*Opsaridium peringueyi*	2013/05/28	Pongolo	-27.3846	31.6673	South Africa	Pongolo	RB13-Mes12
Uncatalogued	*Opsaridium peringueyi*	2013/05/29	Usutu	-26.8578	31.9076	Swaziland	Pongolo	RB13-Mes16
Uncatalogued	*Opsaridium splendens*	2010/09/15	Malagarasi	-5.31722	30.1728	Tanzania	Congo	MGSI 13
Uncatalogued	*Raiamas salmolucius*	2010/09/15	Malagarasi	-5.4231	30.2444	Tanzania	Congo	MGSI 12
188008	*Rastrineobola argentea*		Lake Victoria	0.0229	32.4902	Uganda	Nile	RB12-Misc081

Table 2.2. Outgroups used for phylogenetic analyses of the COI and RAG1 gene sequences.

Taxon	**Catalog no**	**Source**	**RAG1**	**COI**
Mesobola brevianalis	CTOLO3497	GenBank		HM224176
Mesobola brevianalis	CTOL03497	GenBank	HM224054	
Engraulicypris sardella	n/a	GenBank	JX197015	JX196997
Engraulicypris sardella	t-057-5634	GenBank		HM418189
Rastrineobola argentea	n/a	GenBank	JX197020	
Rastrineobola argentea	t-057-5648	GenBank		HM418191
Neobola bottegoi	CTOL02623	GenBank	HM224056	HM224178
Chelaethiops elongates	n/a	GenBank	JX197014	
Chelaethiops congicus	t-072-7111	GenBank		HM418156
Chelaethiops bibie	CTOL02657	GenBank	HM224023	HM224141
Opsaridium peringueyi	CTOL03496	GenBank	HM224072	HM224192
Opsaridium ubangiense	CTOL03214	GenBank	HM224073	HM224193
Raiamas salmolucius	n/a	GenBank	JX197019.1	JX197004.1

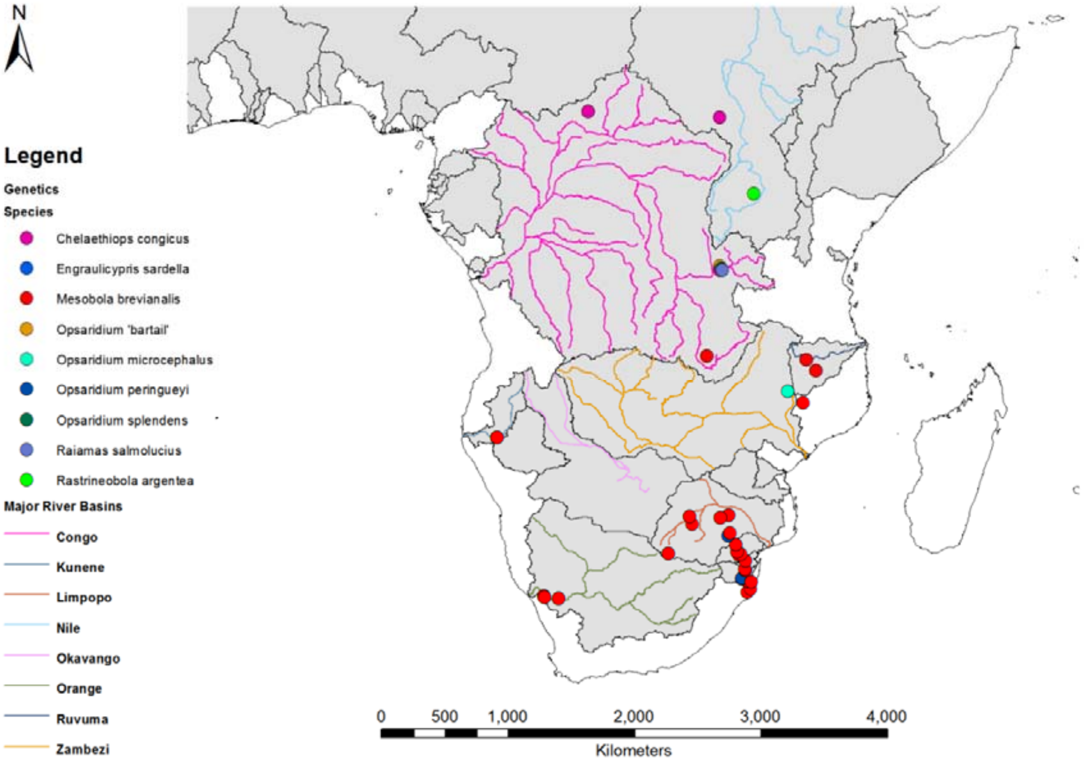


Fig. 2.1. Study site map showing genetic sample sites within twelve river systems in southern Africa. (Map generated by H. Retief, Rhodes University, Grahamstown).

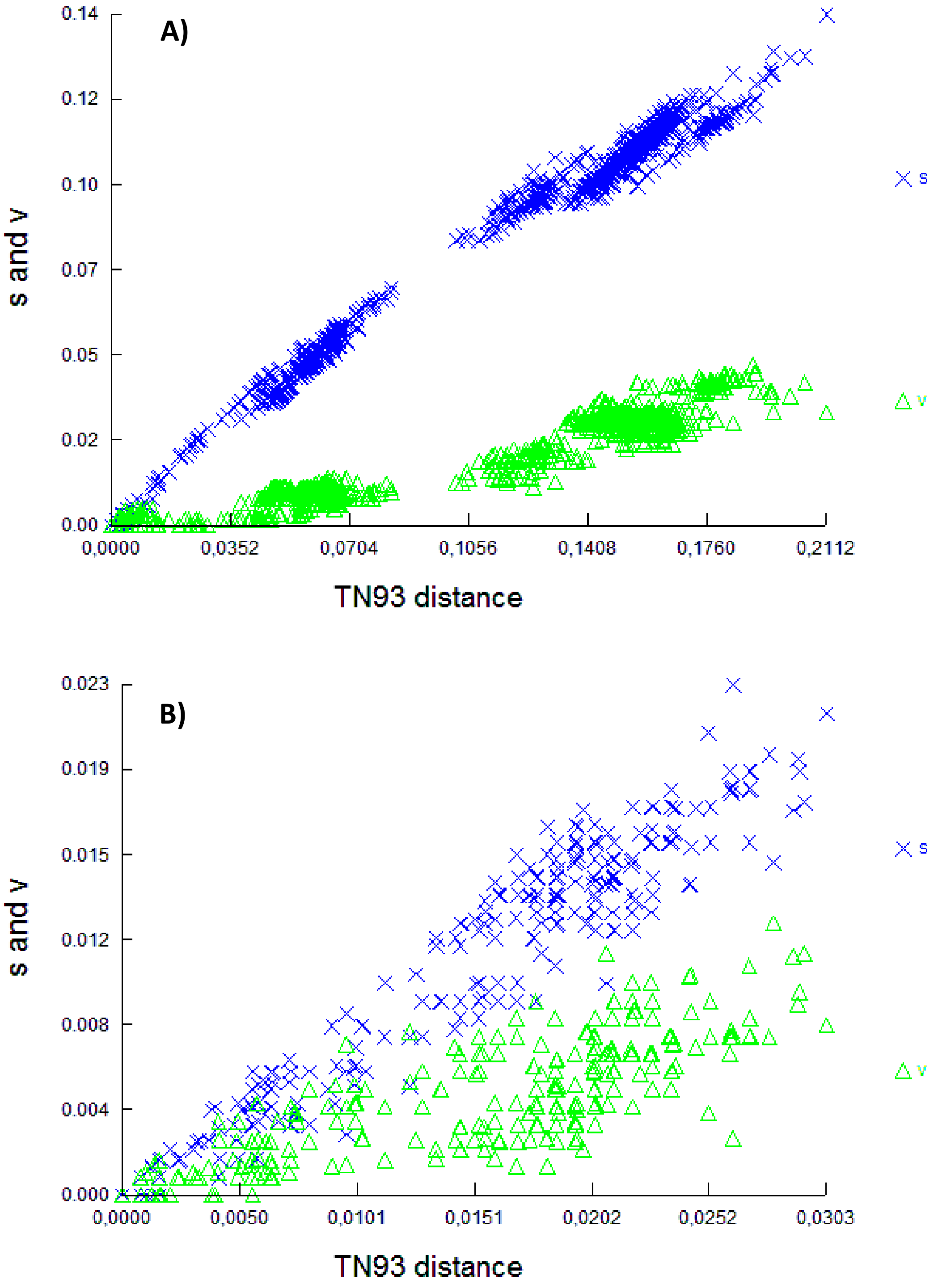


Fig. 2.2. Saturation plots using the TN93 model of best fit for **A)** COI and **B)** RAG1.

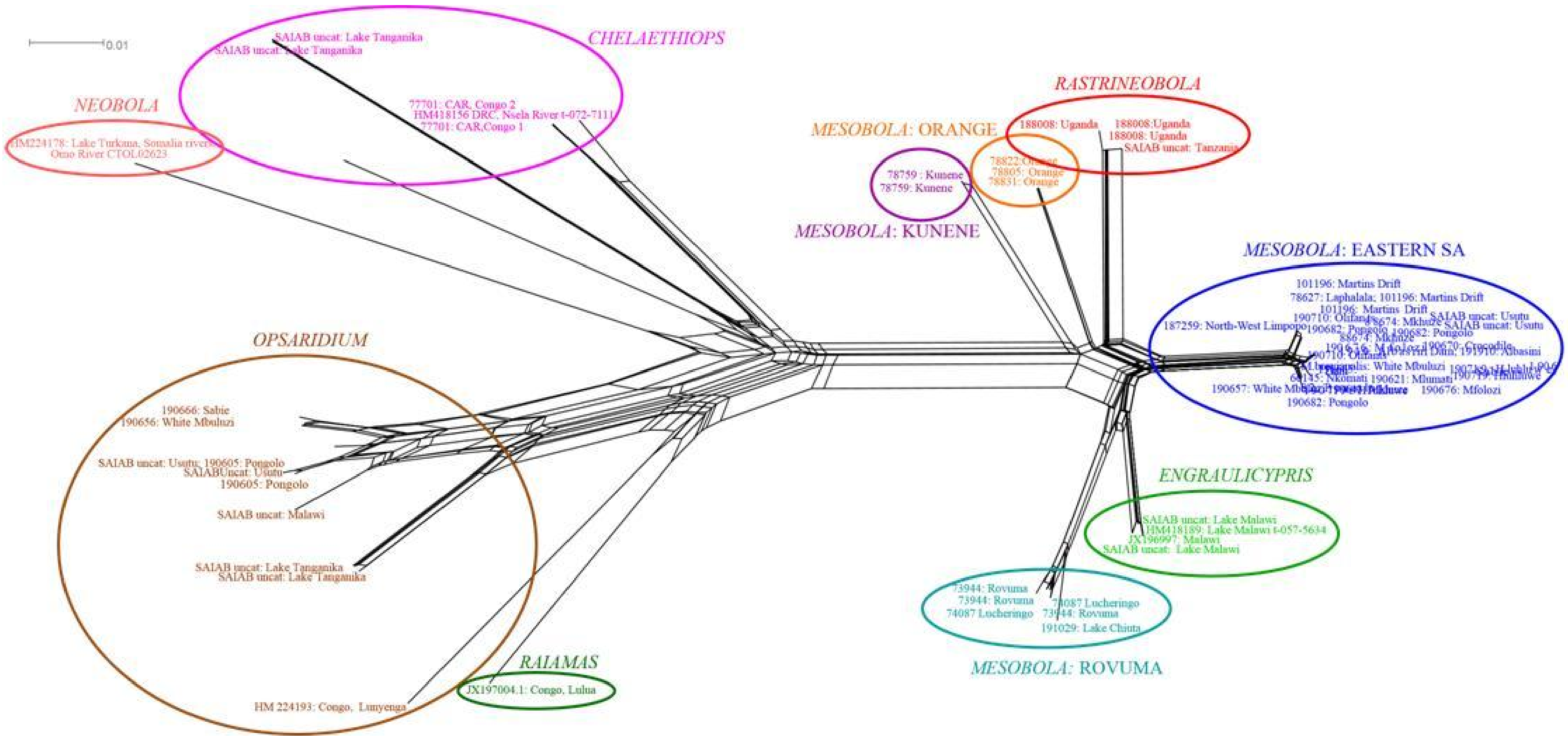


Fig. 2.3.A. Neighbour-Net diagram of the analysis of the COI gene with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

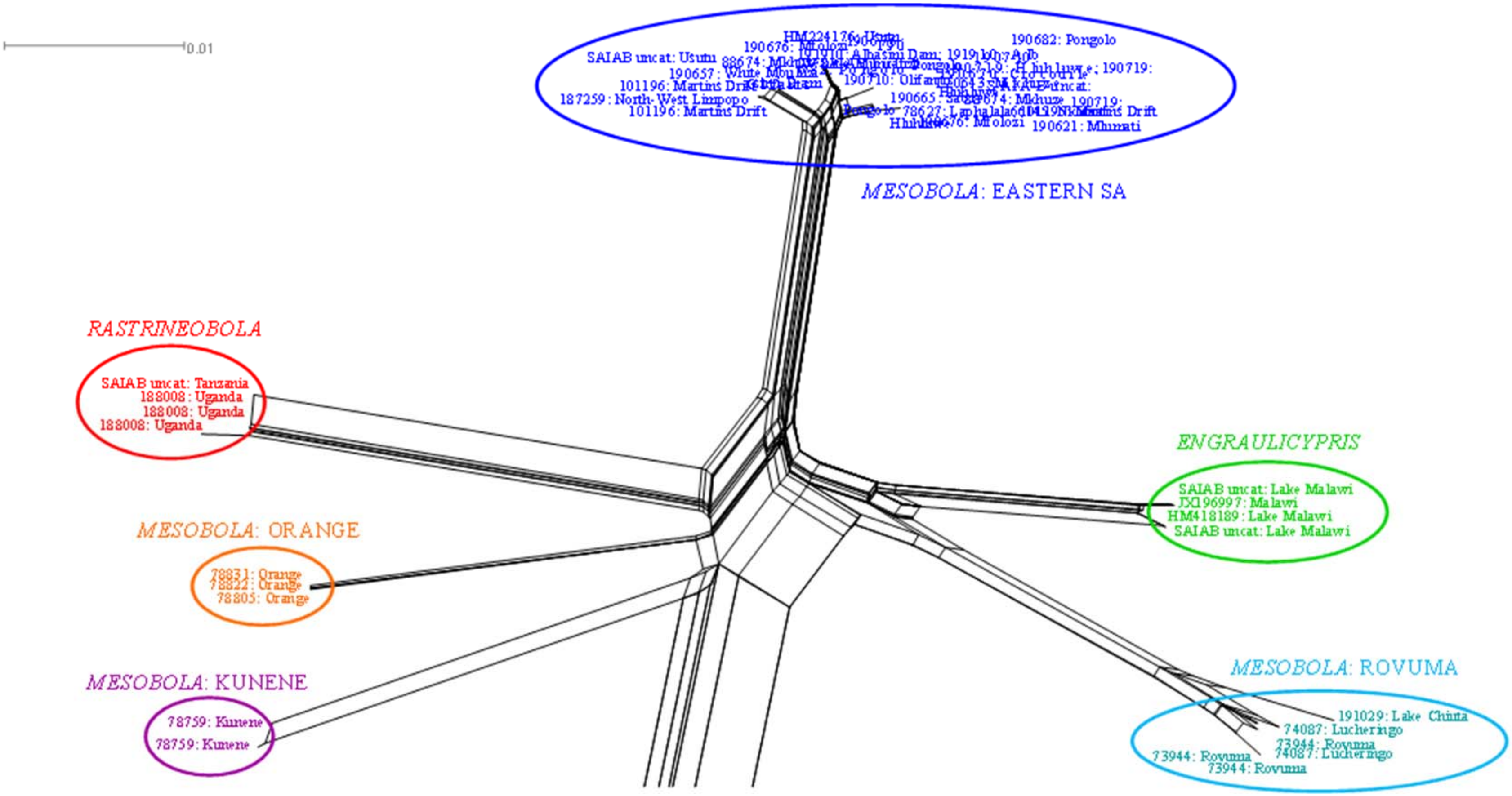


Fig. 2.3.B. Neighbour-Net diagram of the analysis showing the '*Mesobola*' clade for the COI gene with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

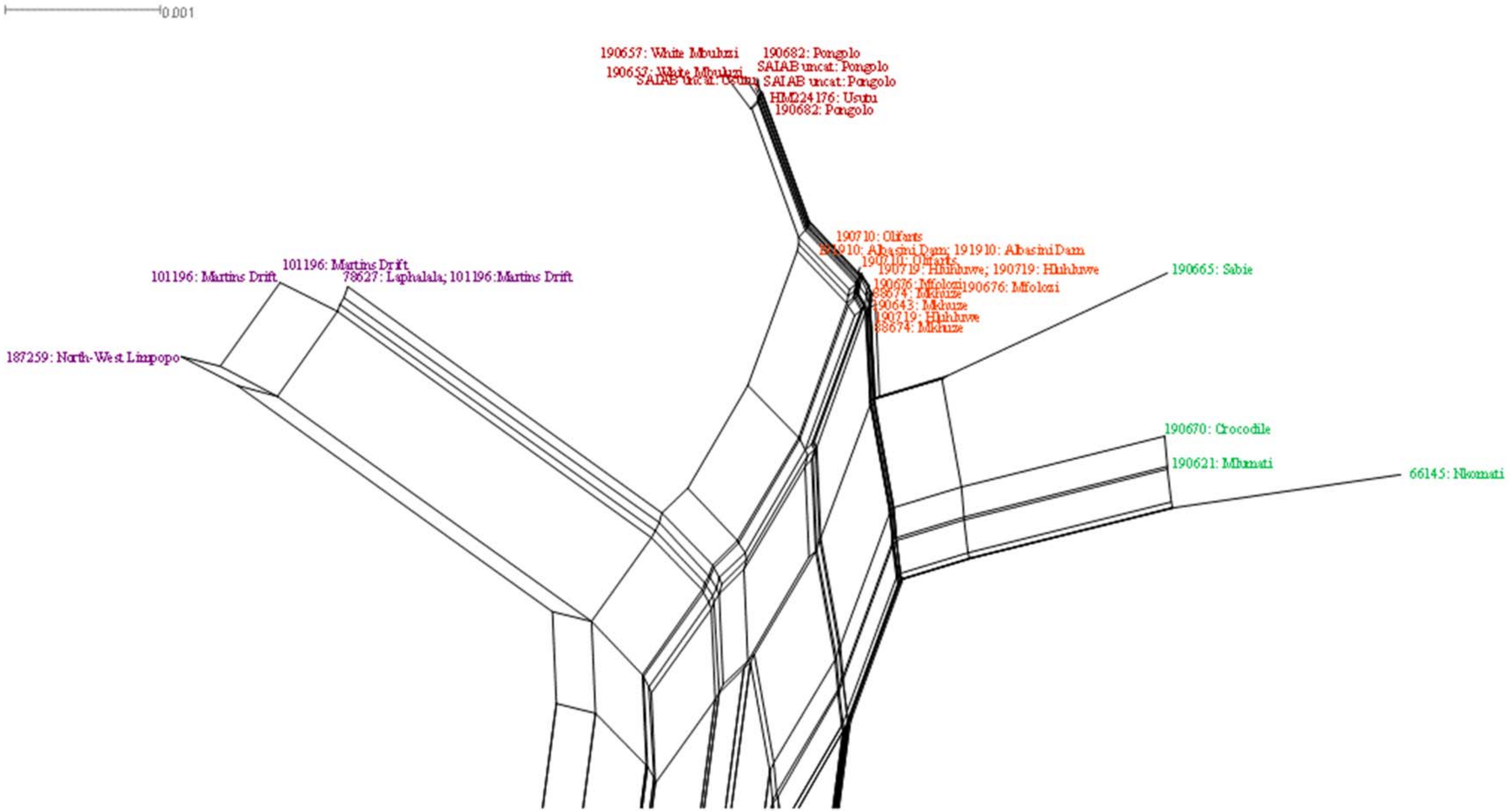


Fig. 2.3.C. Neighbour-Net analyses of the eastern South African populations of *M. brevianalis* for COI data with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

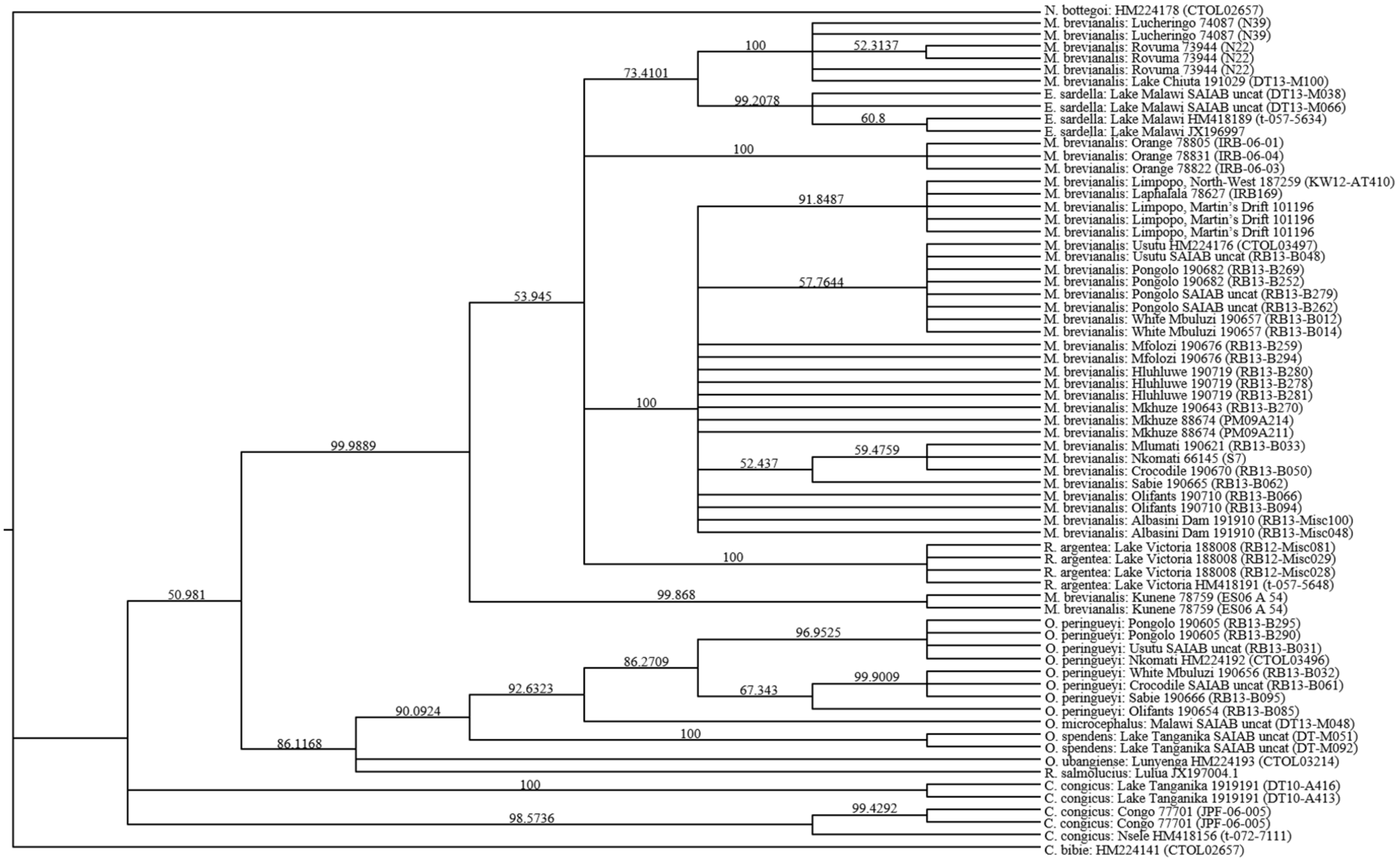


Fig. 2.4. Bayesian COI tree built using TrN+I+G model (outgroup = *Neobola bottegoi*). Posterior probability values were attained using a heuristic tree search using 1000 replicates. Numbers following species name represent GenBank accession and SAIAB catalogue numbers.

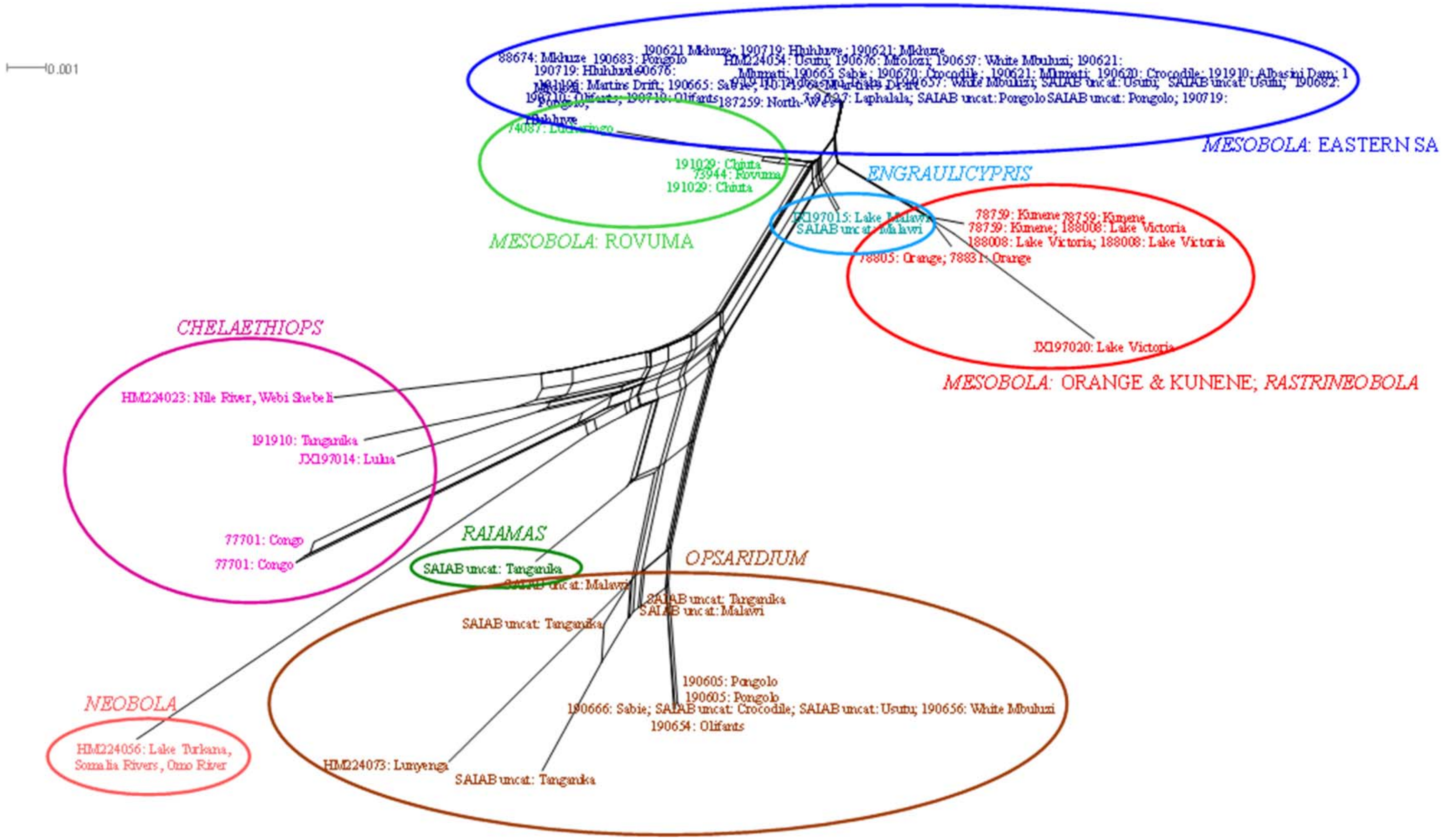


Fig. 2.5.A. Neighbour-net relationship of the family Cyprinidae for the RAG1 gene with a distance of K2P. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

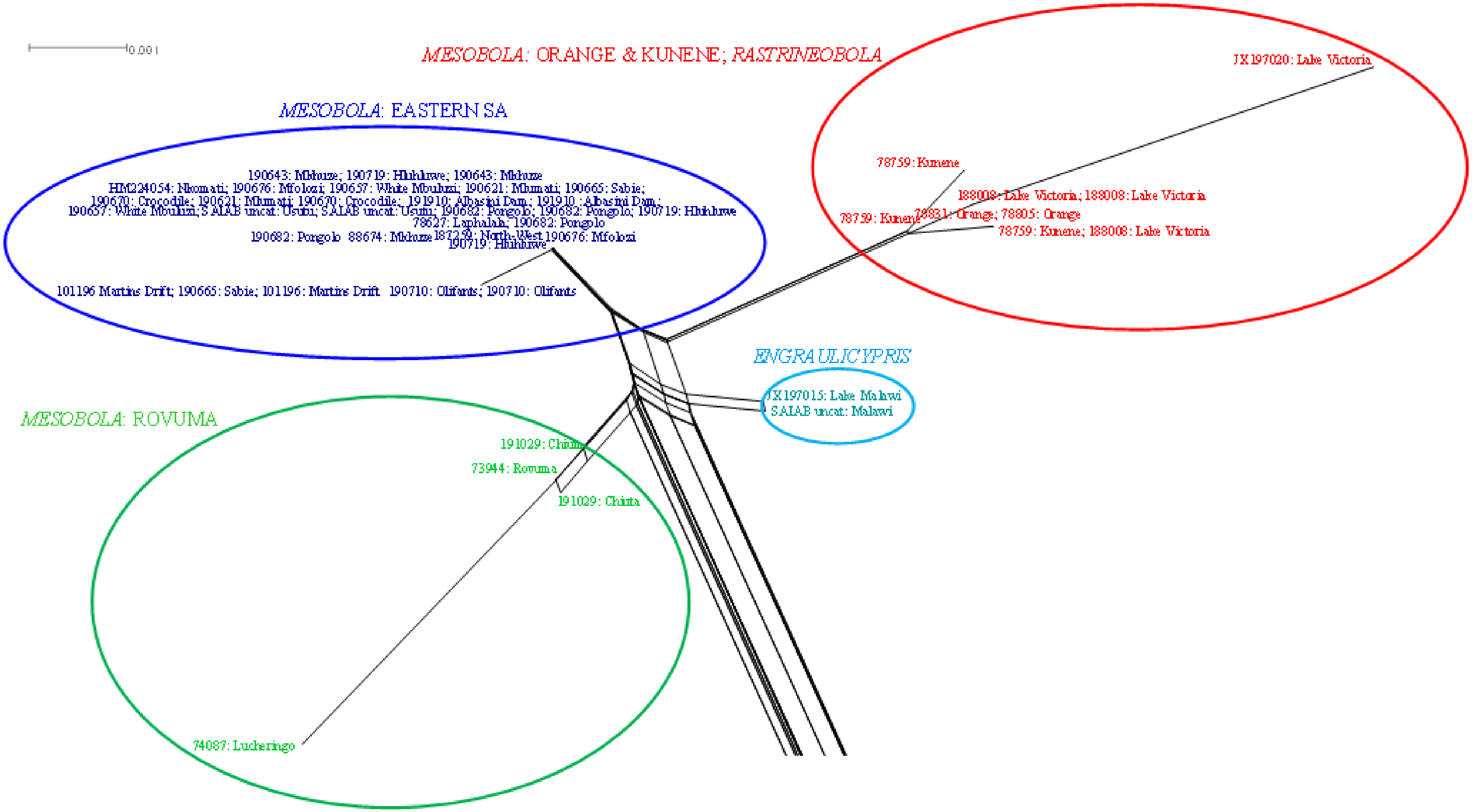


Fig. 2.5.B. NeighbourNet diagram showing the 'Mesobola' clade for the RAG1 gene with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

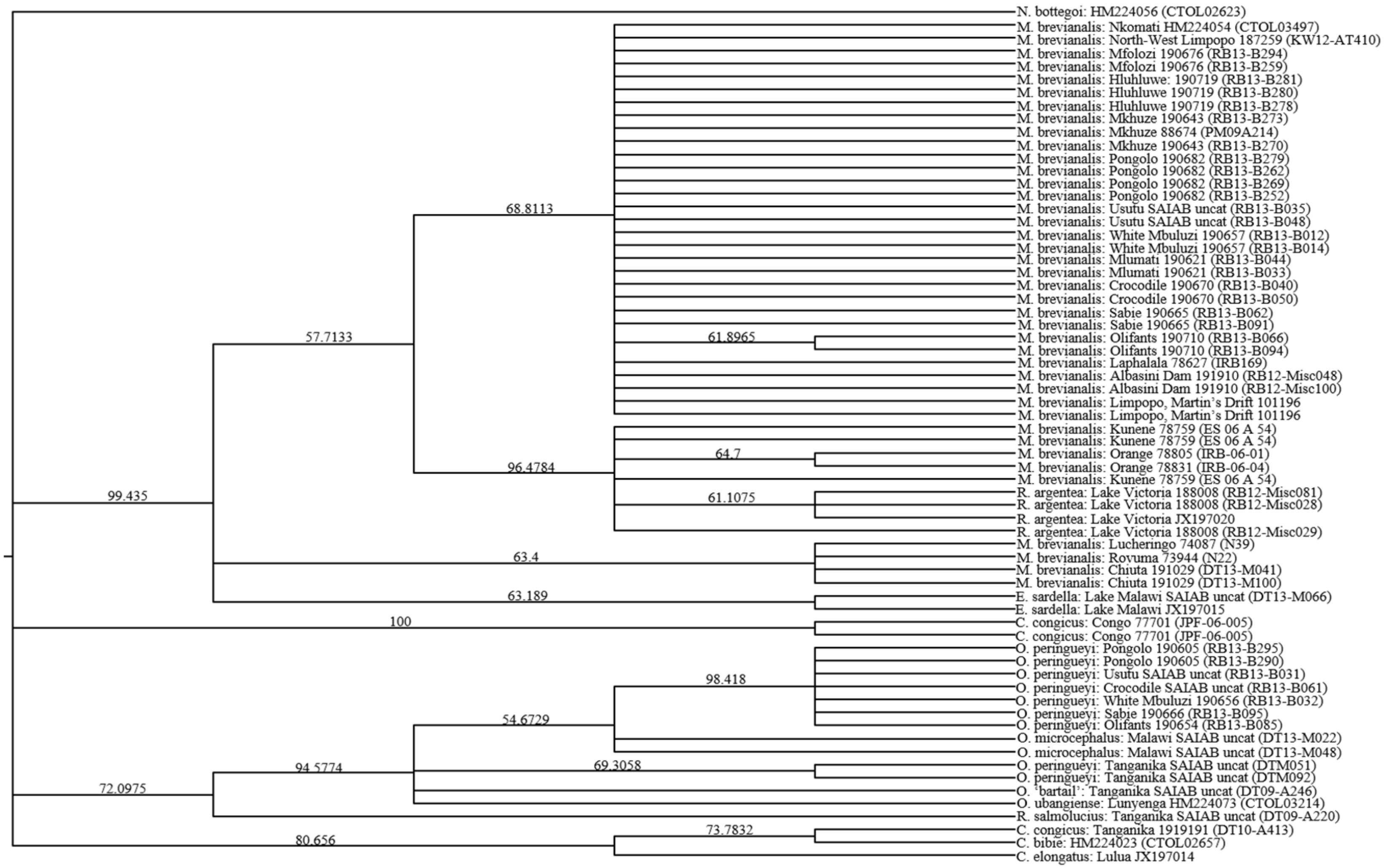


Fig. 2.6. Bayesian RAG1 tree built using TrNef+G model (outgroup = *Neobola bottegoi*). Posterior probability values were attained using a heuristic tree search using 1000 replicates. Numbers following species name represent GenBank accession and SAIAB catalogue numbers.

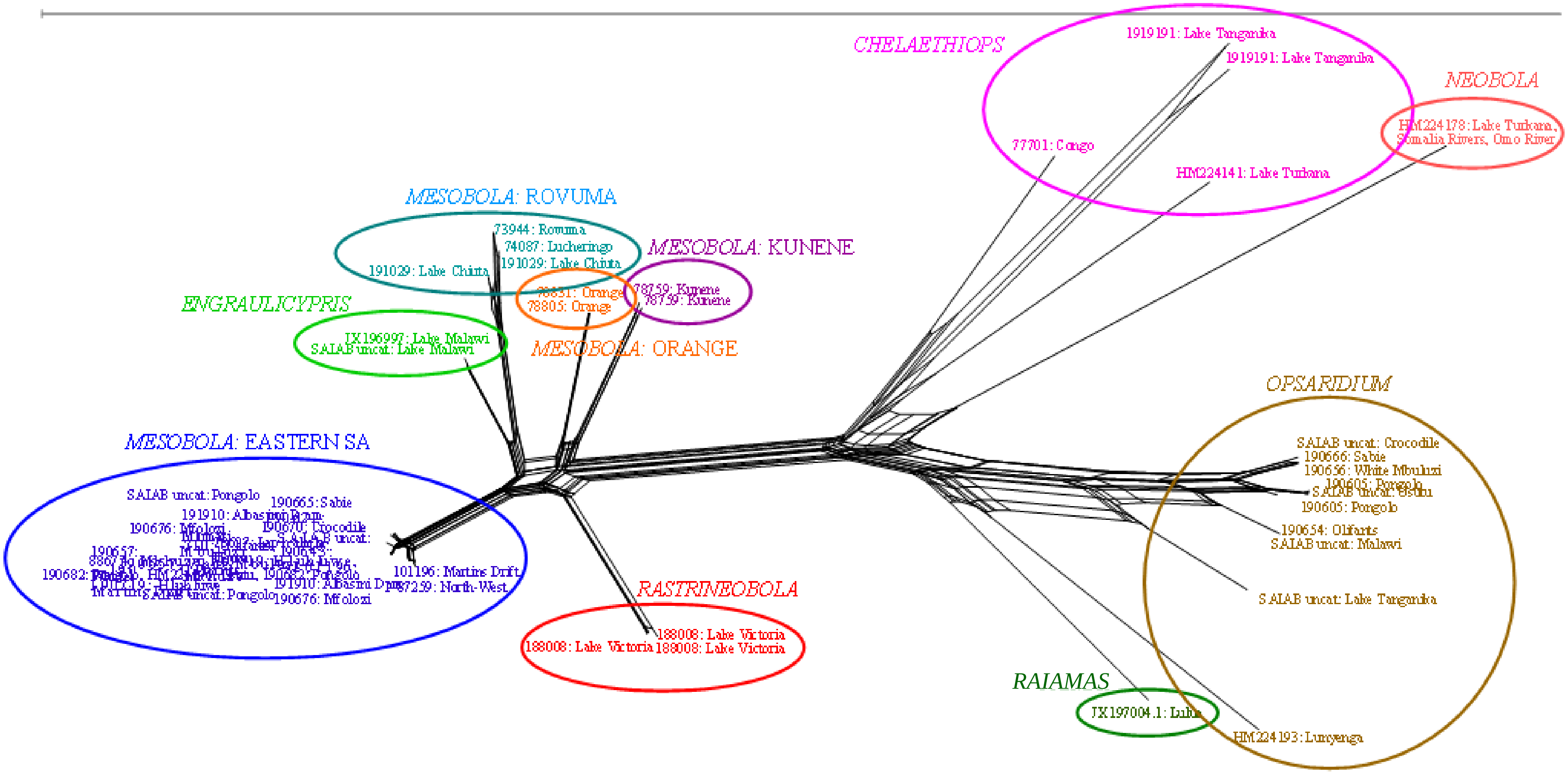


Fig. 2.7.A. Neighbour-Net diagram of the concatenated COI and RAG1 markers with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

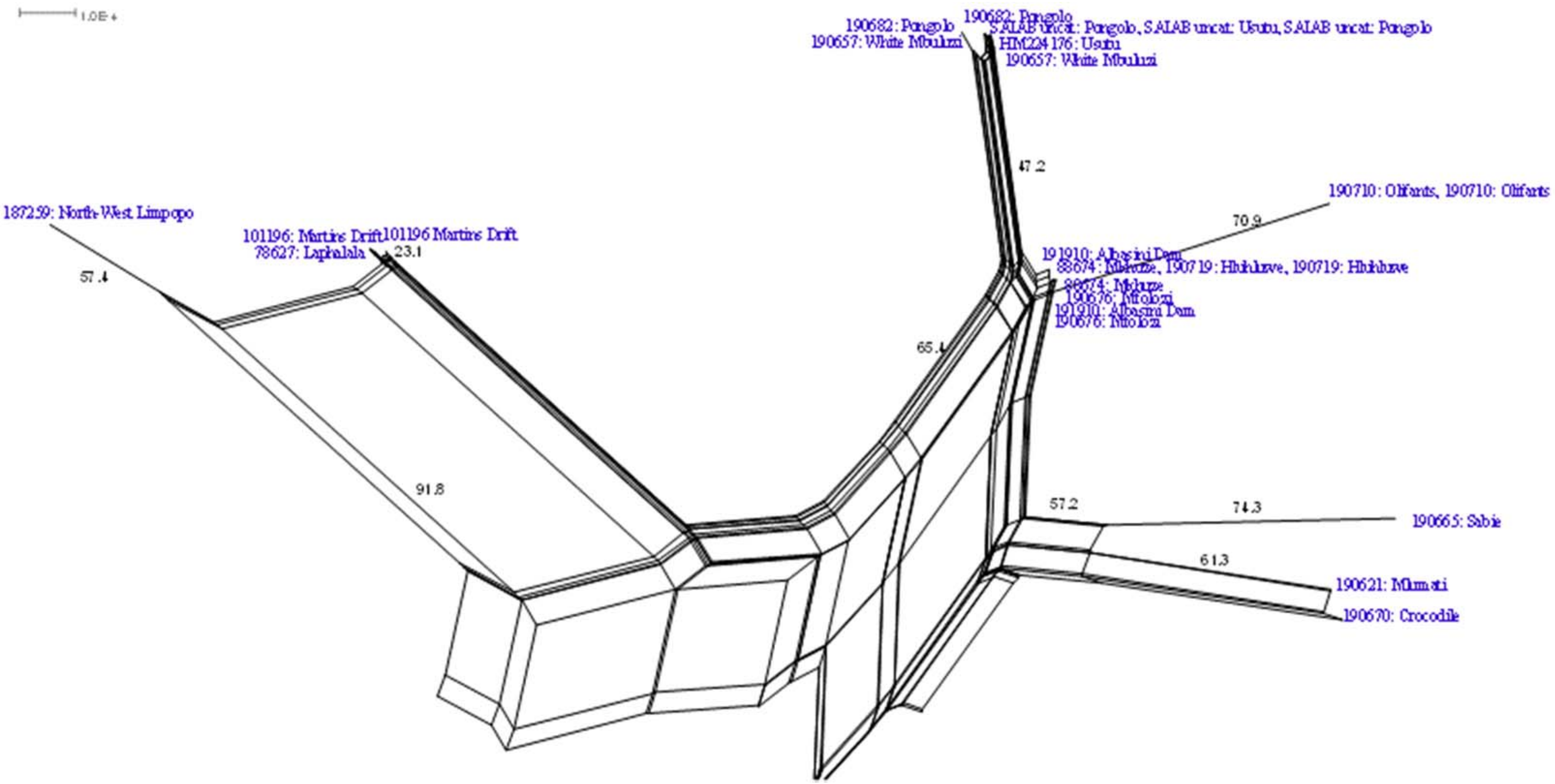


Fig. 2.7.B. Neighbour-Net diagram of the concatenated COI and RAG1 markers showing relationships of South Africa populations of *M. brevianalis* eastern South Africa with K2P distances. Numbers before location represent GenBank accession and SAIAB catalogue numbers.

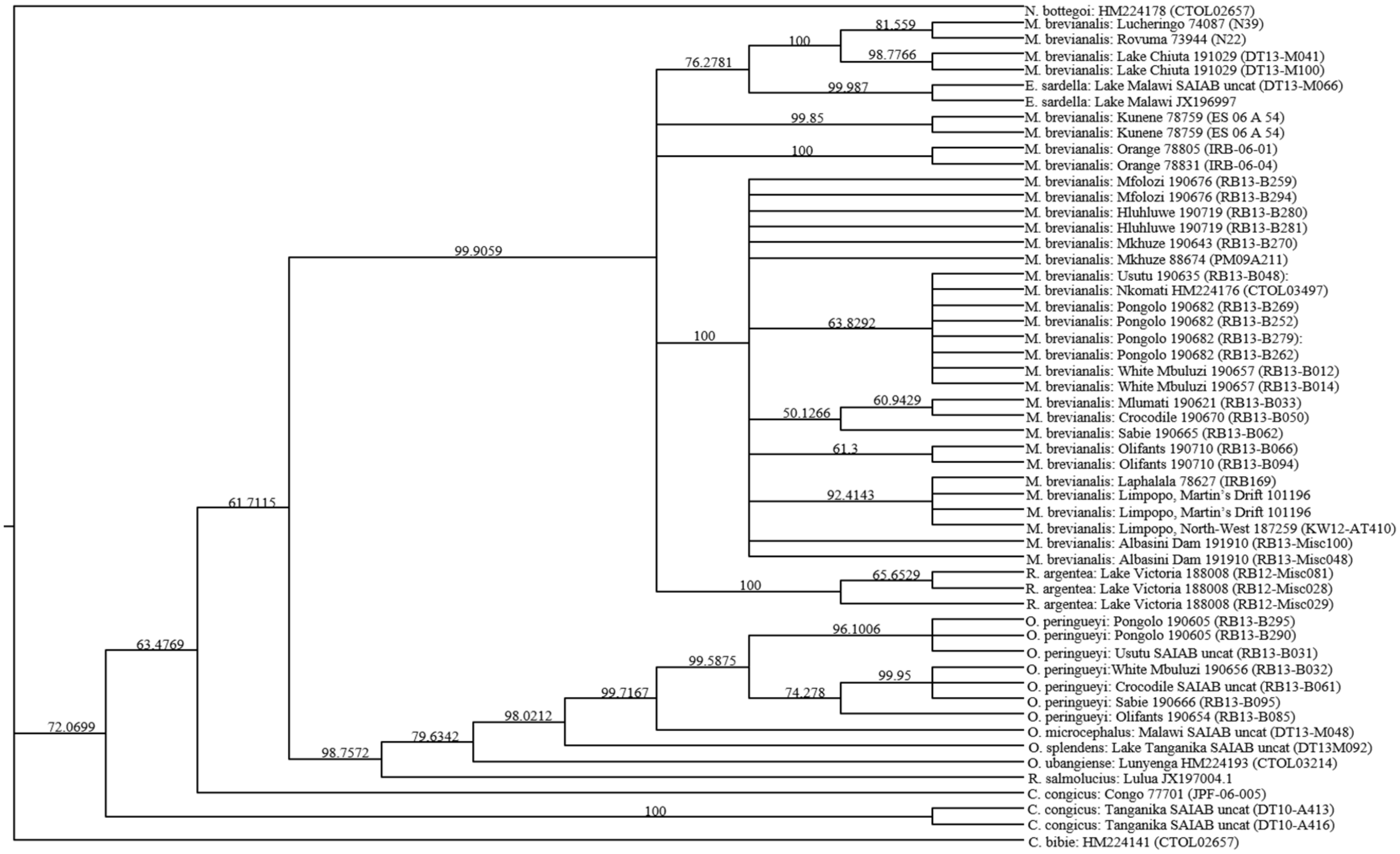


Fig. 2.8. Bayesian concatenated COI and RAG1 tree built using TrN+I+G and TrNef+G models respectively (outgroup = *Neobola bottegoi*). Posterior probability values were obtained using a heuristic tree search using 1000 replicates. Numbers following species name represent GenBank accession and SAIAB catalogue numbers.

3 A morphometric study of *Mesobola brevianalis* (Teleostei: Cyprinidae).

3.1 Introduction

The use of phenotypic variation in the form of morphological and meristic differences has been used for many decades for the identification of fish species (Cadrin 2000; Wessels 2009). The past few decades have allowed for the identification of species and the measuring of relationships among and between taxa (Turan 1999). Morphometrics, particularly multivariate morphometrics, can enable the identification of species and morphologically distinct populations (Turan 1999). The rapid development of morphometric methods has rendered it a cost-effective and widely-used method for identifying species (Sajina *et al.* 2011).

Morphometrics is based on characteristics that show ontogenetic changes in body shape and form (Gould 1966) and are often able to show the effects of allometry (Gould 1966; Alberch *et al.* 1979). Homologous landmarks are essential to morphometric measurements which involve "shared developmental features" of taxa such as attachment points of fin membranes and other such regions (Cadrin 2000). Numerous studies have used morphometric variation to discriminate between discrete fish stocks of different species (Wessels 2009). In fisheries science, the method of assessment of a single or many fish species based on differences in body shape or condition, has become standardized when assessing a population, or populations (Fitzgerald *et al*. 2002).

A caveat with the use of this method of species identification is that organisms' morphometric characters may alter due to a number of factors independent of phylogeny (Turan 1999). Phenotypic characteristics are subject to changes in the physiology and behavior of the taxa due to environmental variation and adaptation which then manifests as morphological change (Turan 1999). This can often lead to phenotypic adaptations being interpreted as genetic change or vice versa, resulting in the misalignment of taxa and species (Turan 1999). The use of integrative approaches and a suite of morphological characters, however, has been an argument in the championing of morphometrics as its own scientific method for identification (Will & Rubinoff 2004).

A study of *Ptychadena* (Anura) found that integrative approaches using molecular phylogeny, morphology and bioacoustics was most successful when addressing taxonomic problems (Dehling & Sinsch 2013). Dehling and Sinsch (2013) found that five species of *Ptychadena* could be identified efficiently through the use of quantitative and qualitative morphometrics. Furthermore, in Greece, 80% of red mullet (*Mullus barbatus*) samples could be assigned to their seven regions correctly using multivariate analyses of morphometric variation (Mamuris *et al*. 1998).

From a multivariate perspective, morphology would have two independent components: the shape of the organism and the size of the organism (Bookstein *et al*. 1985). As most variability within the multivariate characteristics is due to size (Junquera & Perez-Gandaras 1993), eliminating this component enhances our understanding of the effects of shape within statistics. Principal Component Analyses (PCA) aid in the elimination of the effect of size and are effective at detecting outliers within a population (Cadrin 2005).

One of the greatest advances in the traditional methods of morphometric measurements was Strauss and Bookstein's (1982) creation of the box-truss network method, which exhibits greater discriminating power and more effective information abstraction on shape characteristics than had been possible before (Strauss & Bookstein 1982). This method also enables greater accuracy during the classification and identification of species (Strauss & Bookstein 1982). A number of studies utilized this method successfully, including Cavalcanti *et al*. (1999) who found that the truss system was able to supply the most accurate information for ecomorphological studies of serranid fishes. The box-truss network method expands the potential of morphological methods due to its cost-effectiveness, simple data collection and the abundance of new analytical tools (Cadrin & Friedland 1999). The use of a multidimensional approach to quantify shape that the truss network provides, allows for the evaluation of shape differences between populations and allometry within populations (Fitzgerald *et al*. 2002). Fitzgerald *et al*. (2002), stated in a study on yellow perch (*Perca flavescens*) that truss analysis methods have the potential to provide a cheap, accurate and precise method for assessment and quantification of fish condition.

Furthermore, the development of image analysis systems, often computer-aided, has allowed for the advancement of morphometric methods (Cadrin 2000). It has been proven that a greater sample size can be analysed with similar resources and often in a shorter time when using imaging software systems for morphometric analyses (Cadrin 2000).

Morphology provides a practical and simple means of identification that according to Will & Rubinoff (2004) cannot be replaced by DNA barcoding. However, with the rapid developments in DNA and morphometrics it is now possible to use multiple methods to identify species correctly (Stepien & Kocher 1997; Feulner *et al*. 2007; McGee 2013). Although *M. brevianalis* specimens from differing biogeographical locations show only slight morphological variation, morphometrics is required to verify the phylogeny through distinctive morphological features to further qualify unique populations. This method also enables the discovery of characters that may be used to identify new taxa in the field. Many studies have yielded further results with regard to morphology in biogeographically-isolated species populations which included latitudinal, longitudinal and barrier effects, often linking morphology to functionality and environmental adaptation in comparison to phylogeny (Calvacanti *et al.* 1999; Blanck & Lamouroux 2007; Kelly *et al*. 2013).

The study in this chapter will apply standard morphological techniques and associated analysis software, in order to identify morphological markers that will enable the identification of the unique populations discovered by the molecular phylogenetics in Chapter 2. The effect of functionality and habitat will also be studied within this chapter by comparing latitudinal and longitudinal effects on morphological characteristics.

3.2 Materials and methods

The study covered eleven countries: Botswana, Central African Republic, Ethiopia, Malawi, Mozambique, Namibia, South Africa, Swaziland, Tanzania, Uganda and Zambia. Thirteen river systems were sampled: Chambeshi, Congo, Kunene, Limpopo, Mbuluzi, Ngwavuma, Nile, Nkomati, Orange, Pongolo, St Lucia, Rovuma and Zambezi Rivers (Fig. 3.1).

Fish were caught using seine netting, electrofishing, gill netting and ichthyocide and were placed into either 10% formalin, propanol or 95% ethanol, depending on time of capture. Most were placed directly into 10% formalin, alive, in the field and then transferred into 70% ethanol once in the lab for preservation. Each sample group of specimens was then placed into a container with a waterproof label bearing the date, sample number, location and details of the capture and preservation (Table 3.1). The samples were catalogued and stored at the South African Institute for Aquatic Biodiversity (SAIAB) from where they were accessed when required.

Samples were removed from storage and a single specimen was placed into a black- or white-based container (to allow for contrast) filled with 70% ethanol. Photographs were taken of the left side of each specimen, for consistency, using a Canon 550D SLR camera (18.1 megapixels) and 50 mm fixed macro lens (Plate 3.1). A 40 mm scale bar was included in each photograph to calibrate measurements (Plate 3.1). Each fish was then labeled with waterproof paper bearing its specimen number and photograph number, and placed in individual vials for any analysis required later. The fish were then arranged into groups (number, site, and date) and placed in labelled jars. Two to nine specimens of each of *C. congicus*, *E. sardella*, *R. argentea*, *O. peringueyi*, *R. salmolucius, M. spinifer* and *N. bottegoi* were photographed (Plate 3.2) and included in the analyses for comparison with the *M. brevianalis* specimens.

The photographs, including those of available holotypes and syntypes, were organized by their respective rivers and specimen numbers. Each image was imported into the imaging solution software, analySIS docu (Soft Imaging Systems: http://www.soft-imaging.net/). A digital approach to measurement was chosen as this allows for a complete record of each specimen's body shape to be kept for any required repeat measurements or consultations (Cadrin & Friedland, 1999). The development of imaging analysis software has allowed the expansion of the morphometric methods available for fish stock assessments (Cadrin & Friedland 1999).

Ten landmarks were defined on the body of the fish (Fig. 3.2.A from A-K; Table 3.2). Truss network landmarks are required to be areas of strong skeletal support that are known to not fluctuate with allometry (Fitzgerald *et al*. 2002) and are therefore the landmarks that will be most resistant to ontogenetic, temporal or non-taxonomic sources of variation. However, studies have shown that species-specific allometry exists, leaving analyses of some species vulnerable to changing character portions (Fitzgerald *et al*. 2002). This must always be considered when testing fish stocks.

A total of 27 measurements were made on each fish, including 21 truss network measurements, using the ten landmarks (Strauss & Bookstein 1982) (Fig. 3.2.B) and six further standard linear measurements (Fig. 3.2.A, Table 3.3). The measurements were entered into a single Excel spreadsheet with the specimen number, country, river, river system, and nomenclatural status (e.g. holotype, syntype). A scatter plot was run to ensure there were no outliers due to transcription errors. The logarithm of each of the measurements was then calculated to accommodate allometric variation in the ordinations.

Statistica (Statsoft 1994) was used to analyse the data. A principle component analysis was used to determine whether there was a pattern in the morphological variation that corresponded to the different river systems. A discriminant function analysis was performed to pinpoint diagnostic measurements of species or genera. Scatter plots were used to evaluate relationships between the ordination variables and geographical proxy variables, specifically latitude and longitude.

3.3 Results

3.3.1 Principal component analysis

Factor 1 summarised 88.89% of the variance in the data. A scree plot showed that Factors 1, 2 and 3 summarised practically all of the variation. Plotting Factor 1 against Factor 2 showed variation among some of the species tested (Fig. 3.3). *M. brevianalis* was scattered over a wide area of the plot; because of its dominant sample size, the sample lay in the centre of the factor space (Fig. 3.3). A few *M. brevianalis* specimens were grouped at the positive side of Factor 1 (Fig. 3.3) which is indicative of small (A-K caudal length) and narrow bodies (C-I truss) and small narrow tails (E-G truss). *Mesobola spinifer* and the *M. brevianalis* population from the Kunene River fell in the same region as the type locality *M. brevianalis* (Fig. 3.3), as did *R. argentea* (Fig. 3.3).

M. gariepinus (Orange System *M. brevianalis*) also showed similar grouping to *M. brevianalis,* but had predominantly more positive scores on Factor 2 (Fig. 3.3). The eigenvector of Factor 2 indicated that it exhibited a shorter (D-E truss), narrower (D-F truss) caudal peduncle and a longer dorsal fin (C-D truss). It also exhibited differences along Factor 1, scoring slightly more negatively (Fig. 3.3), because of its standard length, distance from pectoral fin to dorsal fin origin (C-I truss) and length of caudal peduncle (E-G truss).

There were two distinct groups in the Rovuma System *M. brevianalis,* which were both placed in a similar position to *M. brevianalis* type on Factor 2 (Fig. 3.3). The Rovuma specimens had two distinct groups on Factor 1, one scoring negatively and the other positively (Fig. 3.3). This shows a distinct difference in the size of these two samples as

Factor 1 is based predominantly on standard length, the depth of the body (C-I truss) and the depth of the caudal peduncle (E-G truss). The specimen grouping in the positive direction had shorter and shallower bodies and thinner caudal peduncles (Fig. 3.3). This may be due to difference in size between fish from the Rovuma System and Lake Chiuta, perhaps as an artifact of sampling shoals of different maturity.

Engraulicypris sardella showed a tight grouping within the negative regions of Factors 1 and 2 and was quite distinct (Fig. 3.3). The *E. sardella* specimens should then be identified based on a distinctive standard length, deep body and caudal peduncle, and elongated caudal regions and long dorsal fins.

Opsaridium peringueyi and *Raiamas salmolucius* appeared to have similar morphology with long bodies and deep caudal peduncles and bodies, but *R. salmolucius* was widely distributed on the plot (Fig. 3.3). Specimens of *Chelaethiops congicus* were grouped tightly within the negative region of Factor 1 and the neutral region of Factor 2 (Fig. 3.3).

The diagram depicting the principal component analysis Factor 2 showed a number of measurements with strong and distinct weightings as to where the differences lay amongst the species (Fig. 3.6.A). The tail length (D-E truss) and depth (D-F truss, D-G truss) both had strong weightings in the negative axis (Fig. 3.6.A). The length of the dorsal fin, width of the dorsal fin (C-D truss) and the depth of the trunk of the body (D-G truss) can be seen by the diagram to have strong weightings within the positive axis for the principal component analysis (Fig. 3.6.A). The body length and the length of the head and front regions did not appear to have strong weightings in terms of Factor 2 among the species (Fig. 3.6.A).

3.3.2 Discriminant function analysis

The discriminant function analysis for the species' morphological variables produced a scatter plot very similar to that of the PCA (Fig. 3.4). *M. brevianalis* type once again appeared in the middle of both axes because it dominated the sample (Fig. 3.4). The *M. spinifer*, *M. brevianalis* Kunene and *M. gariepinus* specimens plotted in the same regions as *M. brevianalis,* but they were slightly more scattered (Fig. 3.4). The Rovuma System specimens of *M. brevianalis* formed a single group centered on axis 2 but plotted high on the positive side of axis 1 (Fig. 3.4) which is evidence of specimens' individual characteristics with regard to short standard length, deep body depths (B-I truss; C-H truss) and short mid-regions (C-I truss). The Kunene population of *M. brevianalis* showed a similar relationship to the east coast *M. brevianalis* specimens as they were plotted on the positive side of axis 1, but they were higher in the positive side of axis 2 (Fig. 3.4), therefore exhibiting longer bodies, shorter heads and mid-regions (A-B truss; B-C truss), and narrower caudal peduncles (E-F truss).

N. bottegoi and *R. argentea* also plotted similarly to *M. brevianalis*, showing very similar body sizes and shapes. However, *R. argentea* lay slightly more to the positive side of axis 1 (Fig. 3.4). *R. argentea* appears to have unique characteristics with regard to standard length, shorter mid-regions between the head and fins (C-I truss) and a slightly deeper head base (B-I truss) and body (C-H truss).

R. salmolucius showed a scattered distribution on the ordination plot, but they were situated on Root 1 similarly to the *C. congicus* specimens (Fig. 3.4). *C. congicus* was slightly further along the positive end of Root 2 in comparison to *R. salmolucius* (Fig. 3.4). The specimens of *R. salmolucius* and *C. congicus* show morphological differences predominantly with regard to standard length, long mid-regions and shallow bodies as they are placed in the negative for the Root 1 axis according to the correlation coefficients (Fig. 3.4). *C. congicus* bears morphological characteristics expressed on Root 2 that show greater difference in terms of their standard length, a shorter head and mid-region leading to the beginning of the dorsal fin and a narrower caudal region compared to *R. salmolucius* (Fig. 3.4).

O. peringueyi and *E. sardella* are seen to be on the opposite ends of the Root 1 axis, indicating opposing characteristics (Fig. 3.4), with *E. sardella* exhibiting a tight group of specimens' morphological characters that are highly different in terms of standard length and head and trunk and body depth while *O. peringueyi* has characteristics that lie predominantly in its standard length, head and trunk length and a body depth (Fig. 3.4). On the Root 2 axis both *E. sardella* and *O. peringueyi* appear to share a strong relationship with specimens that have their morphology characterized by differing standard length head and trunk length, and tail depth (Fig. 3.4).

The discriminant function analysis for the species within the bariliines, which included the populations currently named *M. brevianalis*, showed strong distinctions among the species (Table 3.4). The *M. gariepinus* (Orange River specimens synonymized with *M. brevianalis*), *E. sardella* and *R. argentea* specimens were all identified with 100 % correct placement (Table 3.4). The *M. brevianalis* type specimens were 96.2% correctly placed except for seven specimens that were misidentified as *M. whitei* (western Limpopo specimens synonymized with *M. brevianalis*), two specimens that were identified as Rovuma System *M. brevianalis* and four specimens misidentified under *M. brevianalis* Kunene System population (Table 3.4). The Rovuma System *M. brevianalis* specimens were all identified within the correct species grouping except for four specimens which were misidentified under the *M. brevianalis* specimens and one specimen under *M. whitei* (Table 3.4). The species with the most misidentifications was *M. whitei* which could be seen to have over half the specimens correctly identified within the species and the others misidentified within the *M. brevianalis* cluster (Table 3.4). It can be seen that there were small numbers of some of the species tested including *M. gariepinus*, *E. sardella* and *R. argentea* (Table 3.4).

The discriminant function analysis for the species included within the bariliines showed some differences among the genetically distinct species (Fig. 3.5). The *M. brevianalis* specimens were aligned slightly on the positive of Root 1 axis and the negative of Root 2 (Fig. 3.5). The plot of *M. brevianalis* shows morphology with slight differences to the other species, being seen predominantly with body depth (C-H truss) and head (B-I truss) and length of body and trunk (C-I truss) (Fig. 3.5) according to the coefficients for the canonical variables. The *M. whitei* specimens showed similar body shapes and sizes to the type *M. brevianalis* with a more negative position along the Root 2 axis (Fig. 3.5), showing slightly longer specimens with deeper bodies (C-H truss; D-H truss). *Rastrineobola argentea* showed a very similar plot to type *M. brevianalis* (Fig. 3.5) indicating these specimens are very similar in body shape. The Kunene *M. brevianalis* specimens showed very similar groupings to both type *M. brevianalis* and *M. whitei,* but they can be seen to have a slightly more negative positioning on the Root 2 (Fig. 3.5) axis showing specimens with slightly smaller body lengths and slightly elongated heads (A-B truss) and trunks (B-C truss).

The Rovuma System *M. brevianalis* specimens have a very distinctive body form in terms of Root 1 as they are grouped on the negative axis, but they show a similar plot to type *M. brevianalis* on the Root 2 axis (Fig. 3.5). The Rovuma System specimens could be seen to bear morphological differences in standard length and trunk length (C-I truss), and body trunk (C-H truss) and head depth (B-I truss) according to the coefficients. *M. gariepinus* and *E. sardella* specimens each showed very different plots which would show distinct morphology for these two species (Fig. 3.5). *M. gariepinus* was plotted in the positive of both

Root 1 and Root 2 (Fig. 3.5), which indicates a species with unique morphology with regard to head (B-I truss) and trunk depth (C-H truss), standard length and head (A-B truss) and trunk length (B-C trunk). The *E. sardella* specimens plotted in the opposite position to *M. gariepinus* as the latter is in the deep negative of both Root 1 and Root 2 (Fig. 3.5). This shows that *E. sardella* has specimens with morphology differing in the length of their body trunk (B-C truss; C-I truss), length of head (A-B truss) and standard length (Fig. 3.5).

The diagram depicting the discriminant function analysis Root 1 for the species within the bariliines showed that there were a number of truss measurements that had stronger weightings (Fig. 3.6.B). The trunk depth (C-I truss) can be seen to be deeper for the negative axis in addition to body length and head to jaw length (I-J truss) (Fig. 3.6.B). The depth of the head (B-I truss) and the depth of the trunk (C-H truss) could be seen to be deeper in the positive axes and in the length of trunk (H-I truss) and the orbit (eye length) (Fig. 3.6.B).

The diagram depicting the discriminant function analysis Root 2 for the species within the Bariliini tribe showed stronger weightings for the coefficients of the canonical variables than could be seen in Root 1 (Fig. 3.6.C). The caudal length can be seen to have a strong weighting in the positive axis in addition to the depth of many parts of the body including the head (B-J truss) and trunk (C-H truss; D-H truss) (Fig. 3.6.C). The length of the head (A-B truss), trunk (B-C truss), the depth of the caudal peduncle (G-H truss) and the length between the pelvic and anal fin origin (G-H truss) can be seen to have strong weightings in the negative axis of the coefficients (Fig. 3.6.C).

3.3.3 Geographical gradients

The plot of geographical latitude against the principal component Factor 2 did not show a particularly strong spatial pattern in morphology (Fig. 3.7). The *E. sardella* specimens showed a distinctive relationship as they were placed in the far negative of the axis, apart from the other species (Fig. 3.7). The *M. gariepinus* specimens appeared to also share a distinct relationship with latitude in comparison to the other specimens as they were placed in the far positive plot of the Factor 2 axis (Fig. 3.7). The other species including type *M. brevianalis*, *M. whitei*, *M. brevianalis* Rovuma and Kunene all appeared to overlap in their morphology, showing very little variation in terms of latitudinal effects (Fig. 3.7). The *R. argentea* specimens also can be seen to overlap with these species showing very little variation in their morphology in terms of latitude, but they show a tight grouping within the plot (Fig. 3.7).

The plot of longitude against the Factor 2 principal component analysis could be seen to show more variation in relation to the morphology of the species (Fig. 3.8). *M. brevianalis* type, *M. whitei*, *M .brevianalis* Rovuma and Kunene once again showed large overlap and very little variation in terms of morphology with regard to longitude (Fig. 3.8). *M. gariepinus* occupied a distinct position on the positive side of Factor 2 against latitude. This was again evident in *E. sardella* which was plotted within the positive of Factor 2 axis (Fig. 3.8). The *R. argentea* tight grouping within the plot, overlapping those of the other low variation species, is mirrored in the longitude plot as was seen in the latitude plot (Fig. 3.8).

3.3.4 Discriminant function analysis of genera

The discriminant function analysis run on the genera that included all outgroups and ingroups showed a very strong discrimination of the genera (Table 3.5). All of the specimens were correctly identified to their various genera with 100% success except for a single *R. salmolucius* specimen that was misidentified as *O. peringueyi* and a single *M. brevianalis* specimen that was misidentified as *Rastrineobola* (Table 3.5).

The discriminant function analysis for the relationships among the genera showed some distinct groups when Root 1 was plotted against Root 2 (Fig. 3.9). The *Mesobola* species were grouped tightly in the neutral position with slight overlap into both the negative and positive regions for Root 1 and Root 2 (Fig. 3.9).

Rastrineobola resolved very similarly to the *Mesobola* specimens (Fig. 3.9), but they were plotted slightly more in the positive of Root 1, representative of morphological distinctions including the distance from head to dorsal fin origin (B-C truss), caudal fin to anal fin origin (F-G truss) and pectoral to pelvic fin origin (H-I truss). *Chelaethiops* showed a similar plot to *Mesobola* and *Rastrineobola* on the Root 1 axis with a slightly more negative plot (Fig. 3.9) indicating differences lying within standard length, the length from the head to pelvic fin origin (B-H truss) and caudal fin origin to anal fin origin (E-G truss) according to the canonical coefficients. The *Chelaethiops* specimens were further plotted in the positive of Root 2 axis in comparison to the other genera (Fig. 3.9) which is representative of differences in the genus lying strongly in morphology of standard length, the depth of the head (B-I truss) and the length of the caudal peduncle (E-G truss).

The *Neobola* specimens appeared to show a similar plot to *Chelaethiops* and *Mesobola* along Root 1, however, they could be seen to plotted slightly lower in the positive axis of Root 2 in comparison to *Chelaethiops* specimens (Fig. 3.9), which is representative of their being slightly more similar to *Mesobola* with small differences in the morphology of standard length, the depth of the head (B-I truss) and the length of the caudal peduncle (E-G truss) as seen in *Chelaethiops*. *Opsaridium* and *Raiamas* showed a similar plot and they appeared further on the negative side of Root 1 (Fig. 3.9) in comparison to *Neobola* and *Rastrineobola,* therefore, showing differences that lie within the standard length, distance from head on dorsal side to origin of pelvic fin (B-H truss) and the length of the caudal peduncle (E-G truss) (Fig. 3.9). On the Root 2 axis *Neobola* and *Rastrineobola* exhibit a placement very similar to *Mesobola* representative of similar characteristics in terms of the coefficients.

The specimens within *Engraulicypris* were plotted very similarly to *Chelaethiops* and *Neobola* on the Root 1 axis showing characteristics with regard to this root that are very similar, however the *Engraulicypris* specimens were plotted on the very high negative regions of the Root 2 axis (Fig. 3.9). The *Engraulicypris* genus, therefore, as a result of its plot (Fig. 3.9), exhibit characteristic differences that lie within the placement of the dorsal fin origin (B-C truss), the length of the head (A-B truss) and the size of the dorsal fin (C-D truss) in comparison to the other tested genera's morphology.

3.4 Discussion

The morphometric analysis showed very little variation among the species or distinctive markers that may be used to identify the genera, or to identify individual species. Some of the populations identified by the phylogenetic study appeared to be morphologically distinct, which did not appear to be in relationship with the other populations. The availability of Boulenger's (1908) holotype and the syntypes of *M. whitei* allowed for the inclusion of the original documented species within the morphometric study.

The genera appeared to be morphologically distinct and the only exception was *Raiamas* where a single specimen was incorrectly assigned to *Opsaridium*. However, this could be accounted for by the close genetic relationship between *Raiamas* and *Opsaridium* that may be mirrored in phenotypic characteristics at certain life history points. *Engraulicypris* and *Rastrineobola,* although identified phylogenetically as being closely related to *M. brevianalis* and therefore likely of the same genus, were all identified correctly within their genera. This may be as a result of these genera containing only a single species each, but may also be because of the habitats in which these specimens, particularly *Engraulicypris* evolved, as *Rastrineobola* is similar to *M. brevianalis* in morphology. The differences within the genera were seen to lie predominantly in the length of the body, head, caudal fin, dorsal fin and depth of the tail, head region and trunk. *Engraulicypris sardella* was distinctive by its long head region and trunk, small body length and shallow body, indicating a long narrow front body region and short tail to account for the short body length. Commonly known as the lake sardine, *E. sardella* is found within Lake Malawi and the upper Shire River (Froese & Pauly 2014), and would be exposed to highly different environmental conditions in comparison to the favoured habitat of *M. brevianali*s. The large number of endemic species that are found within Lake Malawi suggest it may have been isolated for some time (Beadle 1981), which would account for the extreme change in morphometrics when compared to *M. brevianalis.* The large differences found between *M. brevianalis* and *E. sardella* despite their genetic similarity, may be as a result not only of environmental factors but also of population mixing, which limits adaptive diversification between differing environments (Langerhans *et al*. 2003). *Engraulicypris sardella* then would differ morphologically as a result of natural selection due to environmental conditions and low mixing of organisms between alternative environments, which would not occur in systems where flooding often occurred (e.g. Mozambique) (World Meteorological Organization 2012).

The pelagic region (upper 100m) in which *E. sardella* is found in Lake Malawi experiences little to no water current and only experiences wind-induced internal waves, which are slow moving (Beadle 1981) in comparison to some of the river and stream habitats in which *M. brevianalis* is found. This would account for the longer front regions (head and trunk) and narrow, short tail as there is no need for *E. sardella* to have to swim against currents within Lake Malawi. In addition *E. sardella* feeds almost exclusively on zooplankton in the lake (Beadle 1981) which would further decrease the evolutionary drive for a long and powerful tail as they are not dedicated predators. *M. brevianalis,* however, is found in slow to fast flowing currents and feeds on insects, which would account for a stronger tail and head region (Engelbrecht & Mulder 1999). A similar relationship was seen in perch in the pelagic zone in Sweden where they had narrower bodies and heads with shorter fins and fed exclusively on zooplankton, whereas perch in the littoral zone fed on macroinvertebrates and fish (Svanbäck & Eklöv 2002). This was further supported by *Bryconops caudomaculatus* and *Bryconops wavrini* which had differing morphology due to the greatly different conditions between the lagoon and river habitat in which they were respectively found (Langerhans *et al.* 2003). The *Bryconops* species exhibited an anterior shift in the body depth in the river sites in comparison to the lagoons (Langerhans *et al.* 2003). This is in support of the results seen in *E. sardella* with an anterior shift in body depth and a short, narrow tail.

The species level testing showed a similar result with regard to *M. gariepinus,* which is a species recognized for resurrection in the phylogenetic study (Chapter 2). This would support Barnard's (1943) initial conclusion of *M. gariepinus* as a species based on morphometric and meristic characters. This species appeared to show shorter and narrower tails, similarly to *E. sardella,* but also with larger eyes. This may once again be a result of environment, as the lower Orange River has variable currents and conditions (Hoyt *et al*. 1969). The currents in the lower Orange River where *M. gariepinus* is found are slower than other parts of the river at about 330 million m^3a^{-1} (Mare 2007), which would account for the narrower tails, as they do not require them for strong swimming against currents as seen in *Bryconops* species (Langerhans *et al*. 2003). The large eyes which are a feature of *M. gariepinus* could also be linked to functionality as the lower Orange River is one of the most turbid in the world (Hoyt *et al*. 1969; Compton & Maake 2007). Little light and visibility in the water resulting from large amounts of silt, would result in the increase in the size of eye for increased visual ability. This observation supports the value of large eyes as adaptations to low light levels for movement and predation, as was seen in a study on blue marlin (Fritsches *et al*. 2003). These findings show how morphometrics often can pick up functional morphology, in comparison to genetic and distinct morphology, and place species that are genetically distinct together, or place those that are genetically similar apart (Turan 1999).

The *M. brevianalis* Rovuma System population which was also identified as a distinct population within the phylogenetic study (Chapter 2), showed an interesting split in morphology between the two populations. This was apparent in the principal component analysis but was then eradicated during the discriminant function analyses which may show that there is insufficient difference between the populations to be separated morphologically. This may have also been as a result of tapeworms which were found in most of the Lake Chiuta *M. brevianalis* specimens. The internal parasites had grown into the full extent of the body cavity causing the trunk to be widened and deepened in many cases. These internal parasites could therefore have resulted in the body appearing deeper and wider causing inaccurate results in the statistical tests which would account for one population of the Rovuma System specimens being aligned on the axis, depicting a population with broader and deeper trunks.

The species-level statistical tests showed that many of the populations overlapped in their morphology, independent of size. For example, specimens of *R. argentea* that appeared small, possibly because they were juveniles, overlapped extensively with *M. brevianalis* in the ordination plot. These specimens were all correctly identified in the discriminant function analysis, whereas *M. whitei*, which overlapped with *M. brevianalis*, was misidentified as *M. brevianalis* more than half of the time. *M. whitei* is insufficiently distinct to be discriminated from *M. brevianalis*. This would support the phylogenetic finding in Chapter 2, that although *M. whitei* forms a clade, it was not sufficiently well supported to be considered a discrete population.

The *M. brevianalis* Rovuma System population was only slightly misaligned with *M. brevianalis* too, which may be explained by shared habitat preferences and morphological adaptations (Turan 1999). It may also be a result of flooding that occurs down the eastern side of southern Africa that may link these systems, including large river systems (e.g. Rovuma, Nkomati, Limpopo and Save), which all have large flood-plain areas, and could result in these species mixing and passing on genes coding for observed traits (Tauacale 2002), although the molecular phylogeny contradicts this (Chapter 2). This concept is further discussed in Chapter 4.

The regressions of the principal component scores on latitude and on longitude did not show effects of geography on morphological variation. *E. sardella* and *M. gariepinus* were morphologically distinct, but this reflected adaptation to clear lacustrine and turbid riverine conditions, respectively, rather than an effect of space. It has also been found that cyprinids in particular display morphological gradients that correspond to differing levels of adaptation to fast-flowing waters, which would support the markedly different morphological populations with close genetic relationships (Barlow 1961; Langerhans *et al*. 2003).

The truss weighting diagrams (Fig. 3.6) showing where the predominant differences lay between the populations did not indicate efficient identification characteristics to examine in the field. The discriminant function analysis identified body length, body depth and eye size as notable characteristics. However, the principal component analysis indicated that length and depth of the caudal peduncle and the length of the dorsal fin showed more morphological variation.

The results from the morphometric analyses highlight the challenges that morphological methods face as indicators of relationship. Species can be identified through morphology alone, but often characteristics indicate functional adaptation rather than relationship (Turan 1999). Nonetheless, this method may allow identification of populations in the field, where genetic barcoding is unavailable.

3.5 References

ALBERCH, P., GOULD, S.J., OSTER, G.F., WAKE, D.B. 1979. Size and shape in ontogeny and phylogeny. *Paleontological Society*, **5(3)**: 296-317.

BARLOW, G.W. 1961. Causes and significance of morphological variation in fishes. *Systematic Zoology*, **10(3)**: 105-117.

BARNARD, K.H. 1943. Revisions of the indigenous freshwater fishes of the SW Cape Region. *Annals of South African Museum*, **36**: 101-262.

BEADLE, L.C. 1981. *The inland waters of tropical Africa: An introduction to tropical liminology (Second Edition*). Longman Inc. New York. United States of America.

BLANCK, A., LAMOUROUX, N. 2007. Large-scale intraspecific variation in life-history traits of European freshwater fish. *Journal of Biogeography*, **34(5)**: 862-875.

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 of fishes. *The Academy of Natural Sciences of Philadelphia, Special Publication*, **15**: 227.

BOULENGER, G.A. 1908. On the collection of fresh-water fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum*, **1(3)**: 219-235.

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

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

CADRIN, S.X. 2005. *Morphometric landmarks.* In: Cadrin, S.X., Friedland, K.D., Waldmann, J.R (Eds.), Stock Identification Methods: Applications in Fishery Science. Elsevier Academic Press. London. United Kingdom. pp 153-172.

CAVALCANTI, M.J., MONTEIRO, L.R., LOPES, P.R.D. 1999. Landmark-based morphometric analysis in selected species of serranid fishes (Perciformes: Teleostei). *Zoological studies*, **38(3)**: 287-294.

COMPTON, J.S., MAAKE, L. 2007. Source of the suspended load of the upper Orange River, South Africa. *South African Journal of Geology*, **110**: 339-348.

DEHLING, J.M., SINSCH, U. 2013. Diversity of Ptychadena in Rwanda and taxonomic status of *P. chrysogaster* Laurent, 1954 (Amphibia, Anura, Ptychadenidae). *Zookeys,* **356**: 69-102.

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.

FEULNER, P.G.D., KIRSCHBAUM, F., MAMONEKENE, V., KETMAIER, V., TIEDEMANN, R. 2007. Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: Campylomormyrus): a combined molecular and morphological approach. *Journal of Evolutionary Biology*, **20(1)**: 403-414.

FITZGERALD, D.G., NANSON, J.W., TODD, T.N., DAVIS, B.M. 2002. Application of truss analysis for the quantification of changes in fish conditions. *Journal of Aquatic Ecosystem Stress and Recovery,* **9**: 115-125.

FRITCHES, K.A., MARSHALL, N.J., WARRANT, E.J. 2003. Retinal specializations in the blue marlin: eyes designed for sensitivity to low light levels. *Marine and Freshwater Research*, **54**: 333-341.

FROESE, R., PAULY, D. (eds). *FishBase*. World Wide Web electronic publication. (www.fishbase.org). Electronic version accessed 2014.

GOULD, S.J. 1966. Allometry and size in ontogeny and phylogeny. *Biological Review*, **41**: 537-640.

HOYT, J.H., OOSTDAM, B.L., SMITH, D.D. 1969. Offshore sediments and valleys of the Orange River (South and South West Africa). *Marine Geology*, **7**: 69-84.

JUNQUERA, S., PÉREZ-GÁNDARAS, G. 1993. Population diversity in Bay of Biscay anchovy (*Engraulis encrasicolus* L. 1758) as revealed by multivariate analysis of morphometric and meristic characters. *ICES Journal of Marine Science*, **50**: 383-391.

KELLY, C.D., FOLINSBEE, K.E., ADAMS, D.C., JENNIONS, M.D. 2013. Intraspecific sexual size and shape dimorphism in an Australian freshwater fish differs with respect to a biogeographic barrier and latitude. *Evolutionary Biology*, **40(3)**: 408-419.

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.

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.

MAMURIS, Z., APOSTOLIDIS, P., PANAGIOTAKI, P., THEODOROU, A.J., TRIANTAPHYLLIDISDIR, C. 1998. Morphological variation between red mullet populations in Greece. *Journal of Fish Biology*, **52(1)**: 107-117.

MARE, H. 2007. *Orange River Integrated Water Resources Management Plan: Review of Surface Hydrology in the Orange River Catchment ORASECOM 002/2007*. Orange Senqu River Commission (ORASECOM). http://www.orasecom.org/_system/writable/DMSStorage/1941SURFACE_HYDROLOGY.PDF.

MCGEE, H.T. 2013. *The Identification of the Freshwater Fish of Tobago Using DNA Barcoding and Morphological Analysis: An Integrative Approach*. Senior Independent Study Theses. Paper 4931. http://openworks.wooster.edu/independentstudy/4931.

SAJINA, A.M., CHAKRABORTY, S.K., JAISWAR, A.K., PAZHAYAMADAM, D.G., SUDHEESAN, D. 2011. Stock structure analysis of *Megalaspis cordyla* (Linnaeus, 1758) along the Indian coast based on truss network analysis. *Fisheries Research*, **108(1)**: 100-105.

STATSOFT. 1994. *Statistica*, Tulsa, Oklahoma, USA.

STEPIEN, C.A., KOCHER, T.D. 1997. *Molecules and Morphology in Studies of Fish Evolution*. In: Kocher, T.D., Stepien, C.A (eds). Molecular Systematics of Fishes. Academic Press. San Diego. pp 1-11.

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

SVANBÄCK, R., EKLÖV, P. 2002. Effects of habitat and food resources on morphology and ontogenetic growth trajectories in perch. *Oecologia*, **131**: 61-70.

TAUACALE, F.2002. *Water resources of Mozambique "and the situation of the shared rivers"*. Water Resources Paper presented to the River Basin Information Systems Meeting, 13-14 February, Nairobi.

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

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

WORLD METEOROLOGICAL ORGANIZATION. 2012. Limpopo River Basin: a proposal to improve flood forecasting and early warning system. http://www.wmo.int/pages/prog/hwrp/chy/chy14/documents/ms/Limpopo_Report.pdf.

Table 3.1. Sample catalogue numbers and locality information.

Catalogue Number	Full Name	Start Date	River	Latitude1	Longitude1	Country	System	Event Number
27807	*Mesobola brevianalis*	1988/03/18	Mkhuze swamps	-27.5892	32.4114	South Africa	St Lucia System	MK-11
30040	*Engraulicypris whitei*	1933/02/17	Aapies River	-25.5000	28.2333	South Africa	Limpopo System	TM 15024
30041	*Engraulicypris whitei*	1933/02/17	Aapies River	-25.4167	28.2833	South Africa	Limpopo System	TM 16022
38961	*Mesobola brevianalis*	1992-05-31	Olushandja Dam	-17.4333	14.6667	Namibia	Kunene System	KU 1
39012	*Mesobola brevianalis*	1992-05-30	Olushandja Dam	-17.4333	14.6667	Namibia	Kunene System	KU 6
39269	*Mesobola brevianalis*	1992-07-13	Mthubula Beach	-14.7833	35.8333	Malawi	Lake Chiuta System	Site 26
42529	*Mesobola brevianalis*	1983/07/31	Mwambeshi	-12.2500	25.6667	Zambia	Zambezi system	
66145	*Mesobola brevianalis*	2002/07/09	Mlumati River	-25.7567	31.4386	Swaziland	Nkomati System	S7
66270	*Mesobola brevianalis*	2003/03/31	Great Usuthu River	-26.8575	31.9078	Swaziland	Pongolo System	S30
67353	*Opsaridium peringueyi*	2002/12/13	Ngwavuma River	-27.0361	31.3989	Swaziland	Ngwavuma System	S110
73944	*Mesobola brevianalis*	2003/08/19	Rovuma River	-12.5964	36.9356	Mozambique	Rovuma System	N22
74087	*Mesobola brevianalis*	2003/08/25	Lucheringo River	-11.8156	36.2208	Mozambique	Rovuma System	N39
76771	*Mesobola* sp.	2005-09-23	Lwela River	-11.5261	29.1511	Zambia	Luapula System	JPF-05-011
77147	*Mesobola* sp.	2005-10-13	Chambeshi River	-11.4956	31.2797	Zambia	Chambeshi System	JPF-05-048
77701	*Chelaethiops congicus*	2006/02/24	Congo River	5.8580	20.7280	CAR	Congo System	JPF-06-005
78685	*Mesobola brevianalis*	2006/09/20	Luvhuvhu	-22.8964	30.6983	South Africa	Limpopo System	IRB-06-22
78805	*Mesobola brevianalis*	2006/10/18	Orange River	-28.8681	18.6117	Namibia	Orange System	IRB-06-01
78822	*Mesobola brevianalis*	2006/10/20	Orange River	-28.6886	17.5556	Namibia	Orange System	IRB-06-03
78831	*Mesobola brevianalis*	2006/10/21	Orange River	-28.7472	17.6089	Namibia	Orange System	IRB-06-04
82589	*Mesobola brevianalis*	2009/02/05	Albasini Dam	-23.1033	30.1217	South Africa	Limpopo System	83
101196	*Mesobola brevianalis*	2013/04/10	Martin's Drift	-22.9978	27.9408	Botswana	Limpopo System	RB13-Limpopo1
188008	*Rastrineobola argentea*		Lake Victoria	0.0229	32.4902	Uganda	Nile System	RB12-Misc081
190635	*Mesobola brevianalis*	2013/05/29	Mtindzekwa	-26.7486	31.8369	Swaziland	Pongolo System	RB13-Mes18

Catalogue Number	Full Name	Start Date	River	Latitude1	Longitude1	Country	System	Event Number
190643	*Mesobola brevianalis*	2013/05/27	Mkhuze	-27.6069	32.0446	South Africa	St Lucia System	RB13-Mes05
190657	*Mesobola brevianalis*	2013/05/30	White Mbuluzi	-26.1683	31.8790	Swaziland	Mbuluzi System	RB13-Mes19
190665	*Mesobola brevianalis*	2013/05/32	Sabie	-25.0192	31.206	South Africa	Nkomati System	RB13-Mes23
190670	*Mesobola brevianalis*	2013/05/31	Crocodile	-25.5254	31.3295	South Africa	Nkomati System	RB13-Mes22
190676	*Mesobola brevianalis*	2013/05/26	Mfolozi	-28.3856	32.0385	South Africa	St Lucia System	RB13-Mes02
190710	*Mesobola brevianalis*	2013/06/02	Olifants River	-24.1851	30.8233	South Africa	Limpopo System	RB13-Mes26
190719	*Mesobola brevianalis*	2013/05/26	Hluhluwe	-28.1490	32.2824	South Africa	St Lucia System	RB13-Mes04
Uncatalogued	*Engraulicypris sardella*	2013/03/10	Lake Malawi	-14.1201	34.9346	Malawi	Zambezi system	DT13-M066
Uncatalogued	*Neobola bottegoi*	2012/12/11	Wabe River	7.4417	40.1667	Ethiopia		
Uncatalogued	*Raiamas salmolucius*	2010/09/15	Malagarasi	-5.4231	30.2444	Tanzania	Lake Tanganika	MGSI 12

Table 3.2. The 11 landmarks on the body used to form the truss network and measure body length.

Description	Landmark Number
Tip of snout	A
Occipital ridge	B
Anterior insertion point of dorsal fin	C
Posterior insertion point of dorsal fin	D
Dorsal insertion point of caudal fin	E
Ventral insertion point of caudal fin	F
Anterior insertion point of anal fin	G
Anterior insertion point of pelvic fin	H
Anterior insertion point of pectoral fin	I
Hinge of jaw (aligned with first line on gill plate)	J
Mid-lateral origin of caudal fin	K

Table 3.3. Measurements made to form the truss network (1-21) and further measurements made including body length eye length and position and fin lengths.

Measurements	Landmark
1	A-B
2	B-C
3	C-D
4	D-E
5	E-F
6	F-G
7	G-H
8	H-I
9	I-J
10	J-A
11	B-I
12	C-H
13	D-G
14	A-I
15	B-H
16	C-G
17	D-F
18	B-J
19	C-I
20	D-H
21	E-G
22	A-K (Caudal length)
23	Eye length
24	Snout to eye length
25	Dorsal fin length
26	Caudal fin length
27	Pelvic fin length

Table 3.4. Classification matrix for the discriminant function analysis performed for only the specimens within the Bariliine tribe including *E. sardella* and *R. argentea*.

Group	Percent Correct	brevianalis p=0,71047	whitei p=0,06366	Rovuma p=0,11704	Kunene p=0,05544	gariepinus p=0,03285	argentea p=0,01027	sardella p=0,01027
brevianalis	96,2428	333	7	2	4	0	0	0
whitei	41,9355	18	13	0	0	0	0	0
Rovuma	89,4737	4	1	51	1	0	0	0
Kunene	62,9630	10	0	0	17	0	0	0
gariepinus	100,0000	0	0	0	0	16	0	0
argentea	100,0000	0	0	0	0	0	5	0
sardella	100,0000	0	0	0	0	0	0	5
Total	90,3491	365	21	53	22	16	5	5

Table 3.5. Discriminative function classification matrix for the genera based on the 27 morphometric measurements.

Group	Percent Correct	*Mesobola* p=0,94922	*Chelaethiops* p=0,00781	*Engraulicypris* p=0,00977	*Opsaridium* p=0,00977	*Raiamas* p=0,00977	*Rastrineobola* p=0,00977	*Neobola* p=0,00391
Mesobola	99.794	485	0	0	0	0	1	0
Chelaethiops	100.000	0	4	0	0	0	0	0
Engraulicypris	100.000	0	0	5	0	0	0	0
Opsaridium	100.000	0	0	0	5	0	0	0
Raiamas	80.000	0	0	0	1	4	0	0
Rastrineobola	100.000	0	0	0	0	0	5	0
Neobola	100.000	0	0	0	0	0	0	2
Total	99.609	485	4	5	6	4	6	2

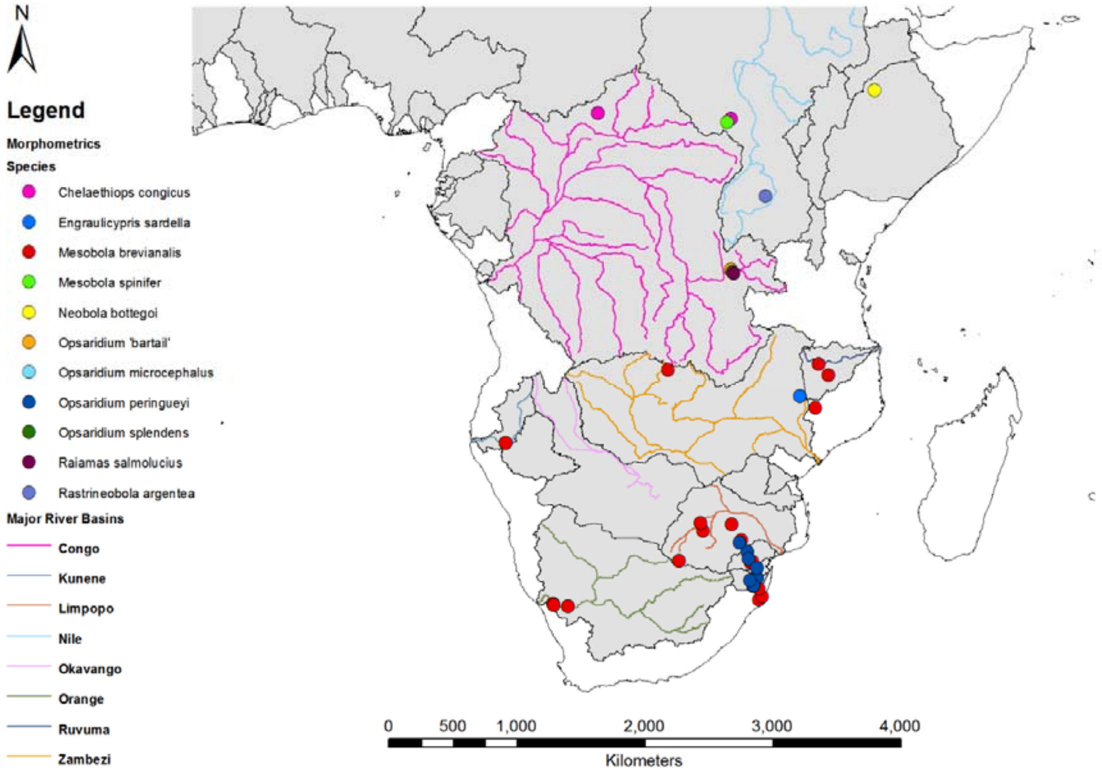


Fig. 3.1. Study site map showing morphometric sample sites within thirteen river systems in southern Africa. (Map generated by H. Retief (Rhodes University, Grahamstown)).

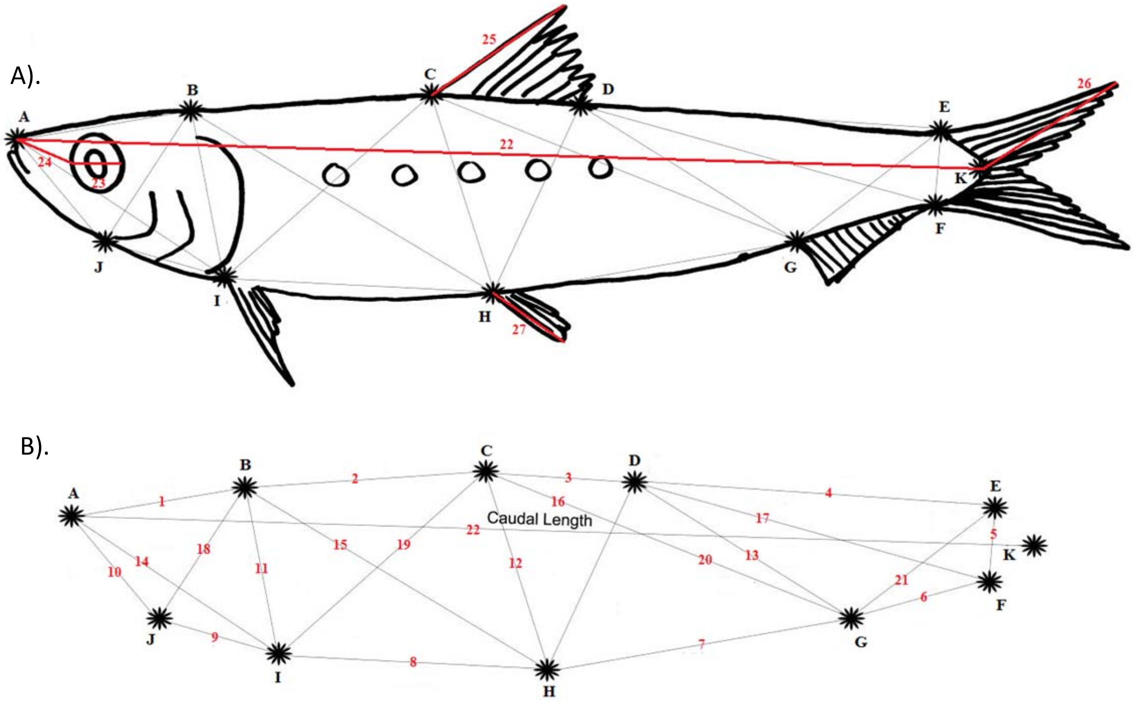


Fig. 3.2. (A) 10 landmarks defined and eye, fin and body length measurements made (red), (B) the Truss network defined using 10 morphometric landmarks on the body .(Image of fish from Wessels(2009)).

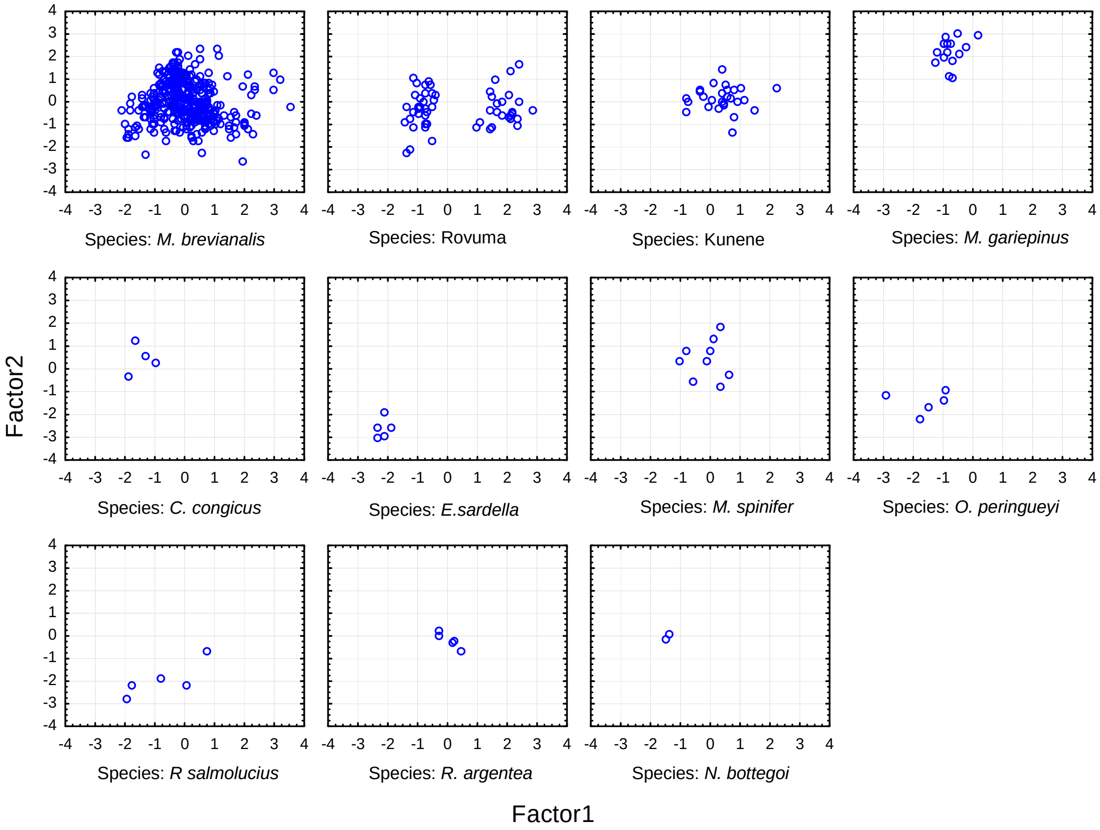


Fig. 3.3. Principal component Factor 1 against Factor 2 at species level plotted individually showing no linear trends indicative of allometric effects.

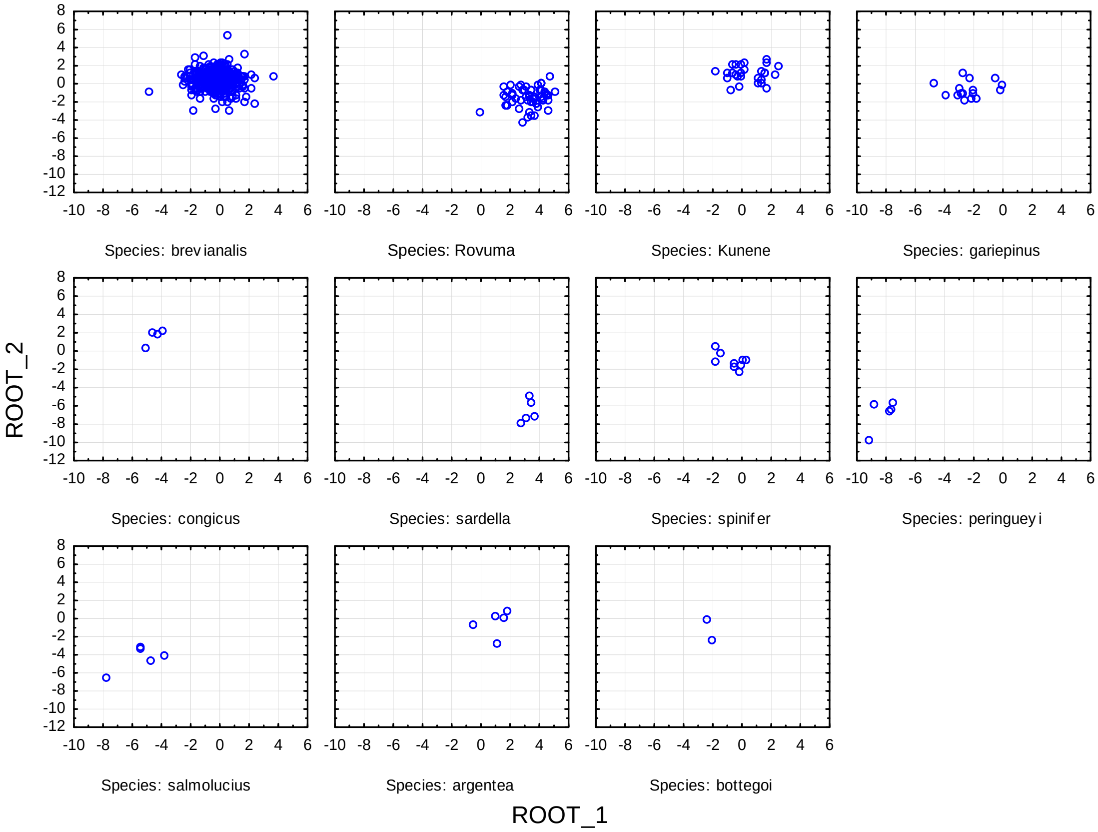


Fig. 3.4. Discriminant function Root 1 against Root 2 at species level plotted individually showing no linear trends indicative of allometric effects.

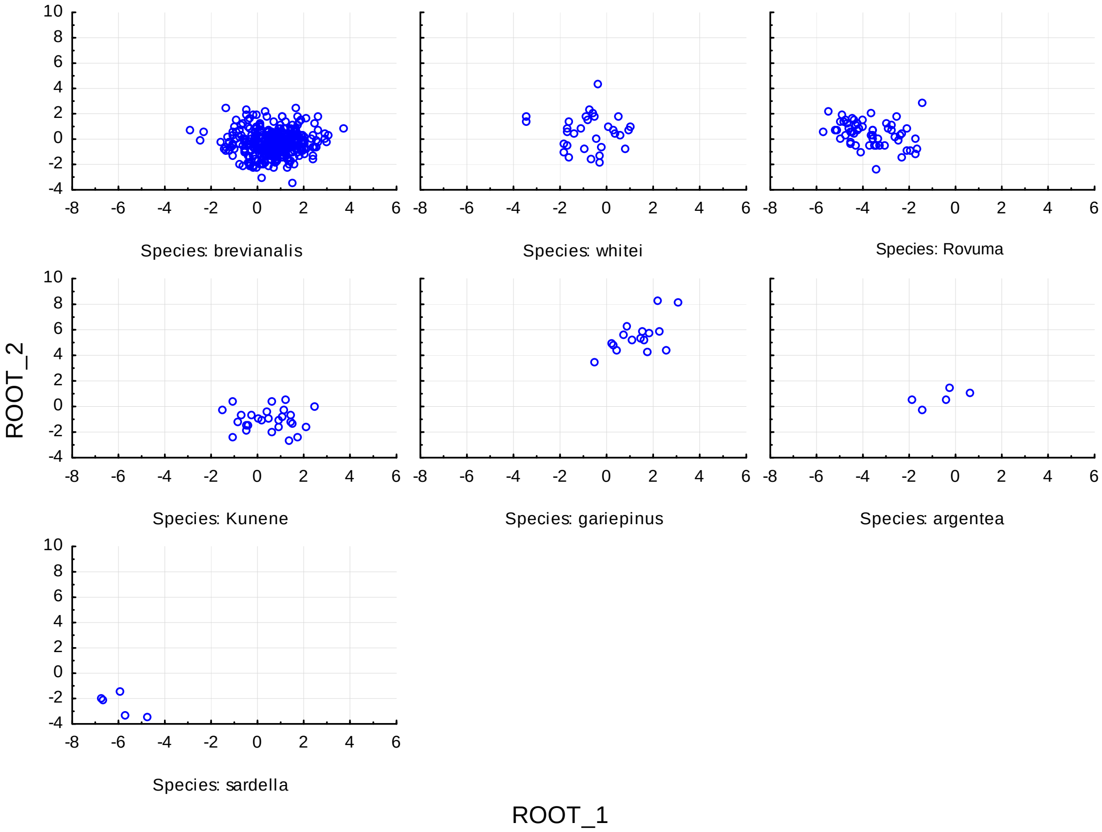


Fig. 3.5. Discriminant function Root 1 against Root 2 at species level plotted individually showing no allometric effects.

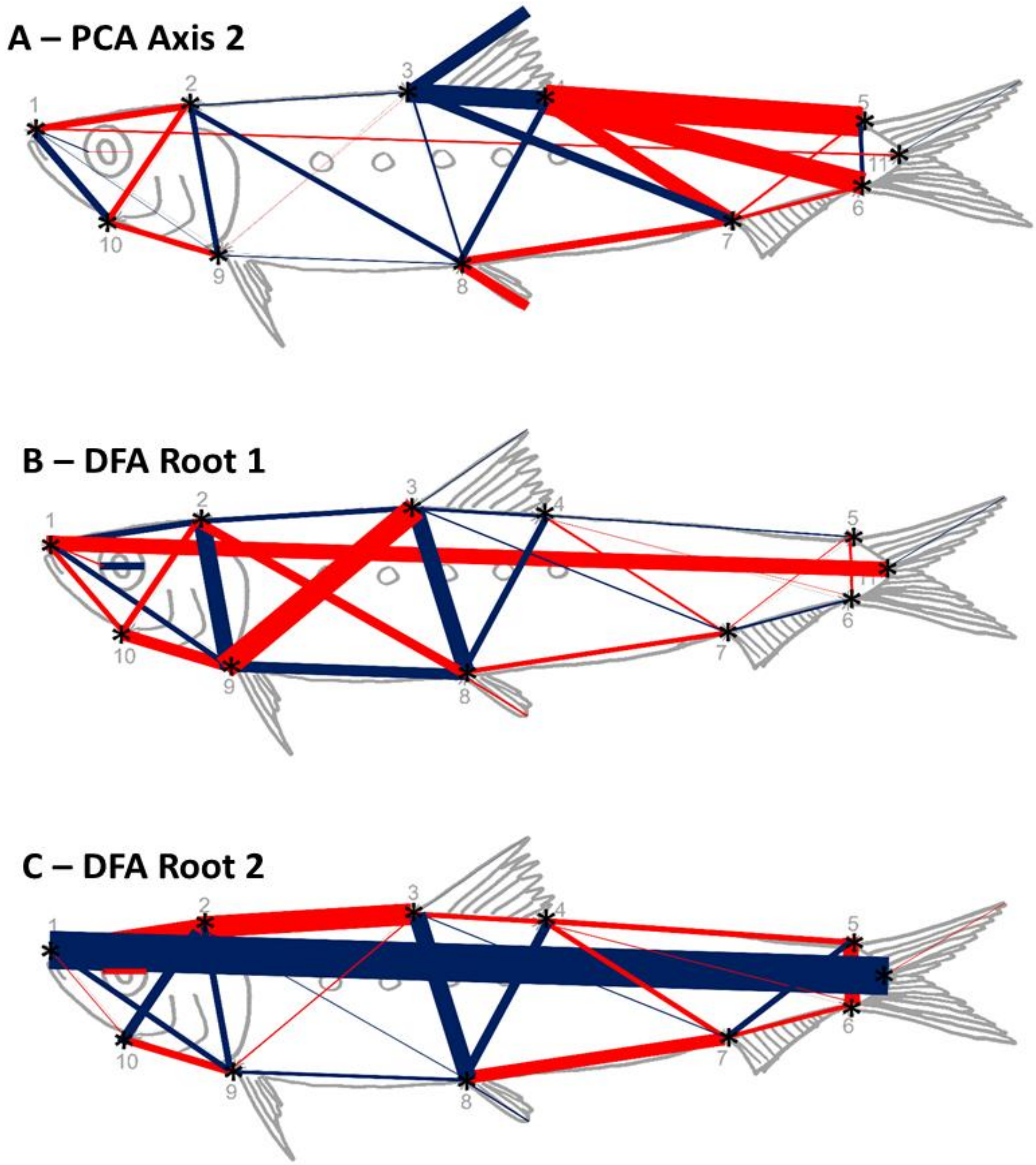


Fig. 3.6. Diagram showing trusses with edges weighted according the Eigenvector coefficients of A – Principal Component 2; B - Discriminant Function Root 1; and C - Discriminant Function Root 2 for the 27 morphometric measurements (blue = positive; red = negative).

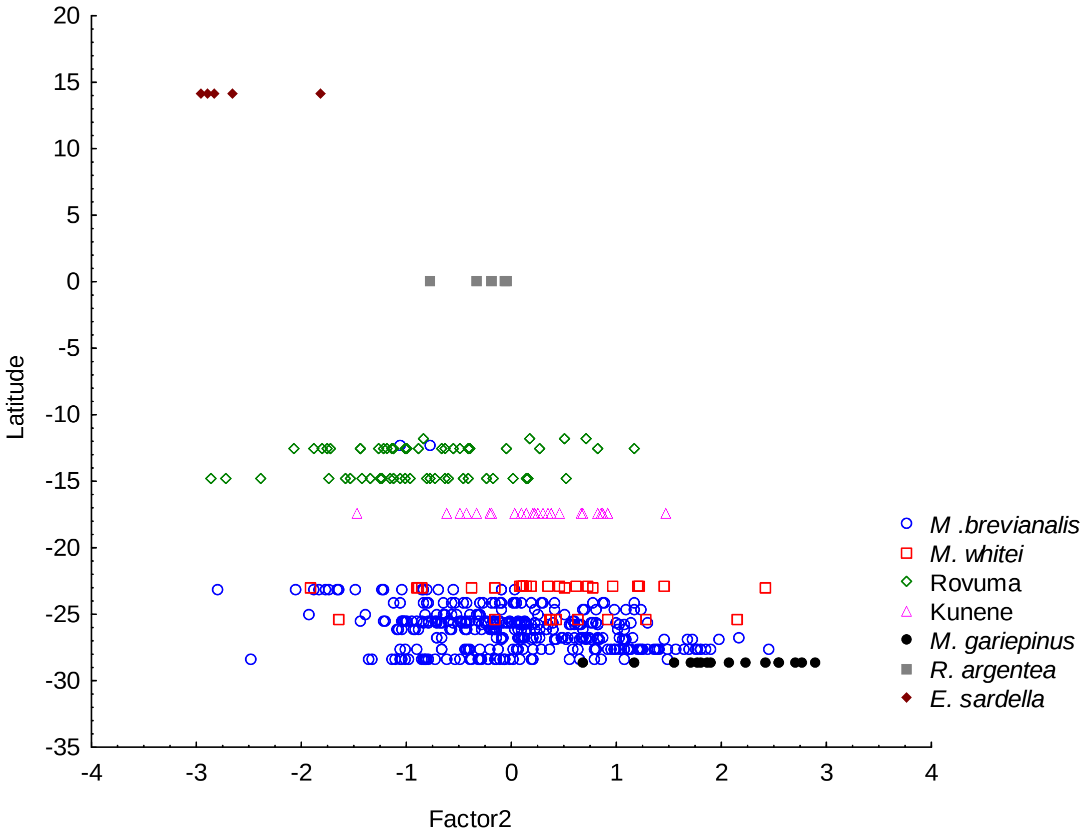


Fig. 3.7. Specimens' scores on Principal Component 2 plotted against the latitude of their collecting locations for each species.

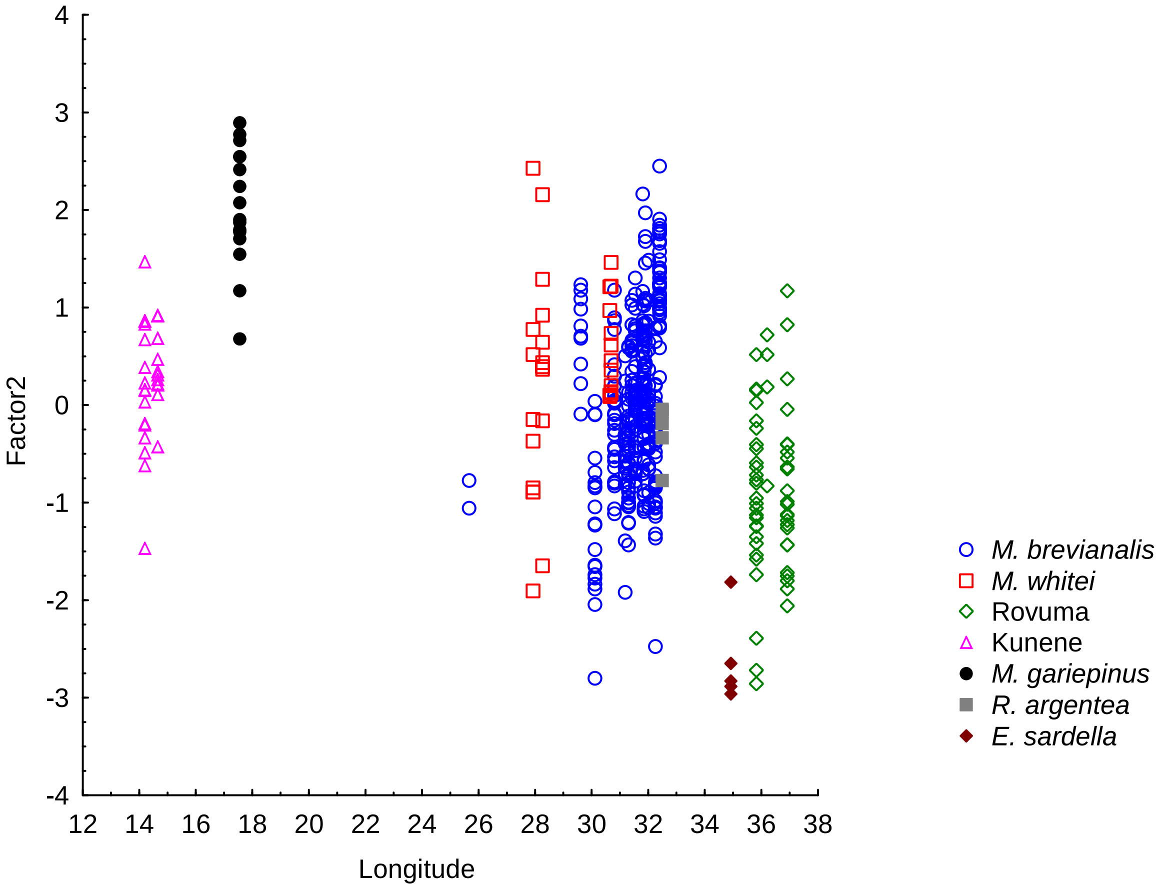


Fig. 3.8. Specimens' scores on Principal Component 2 plotted against the longitude of their collecting locations for each species.

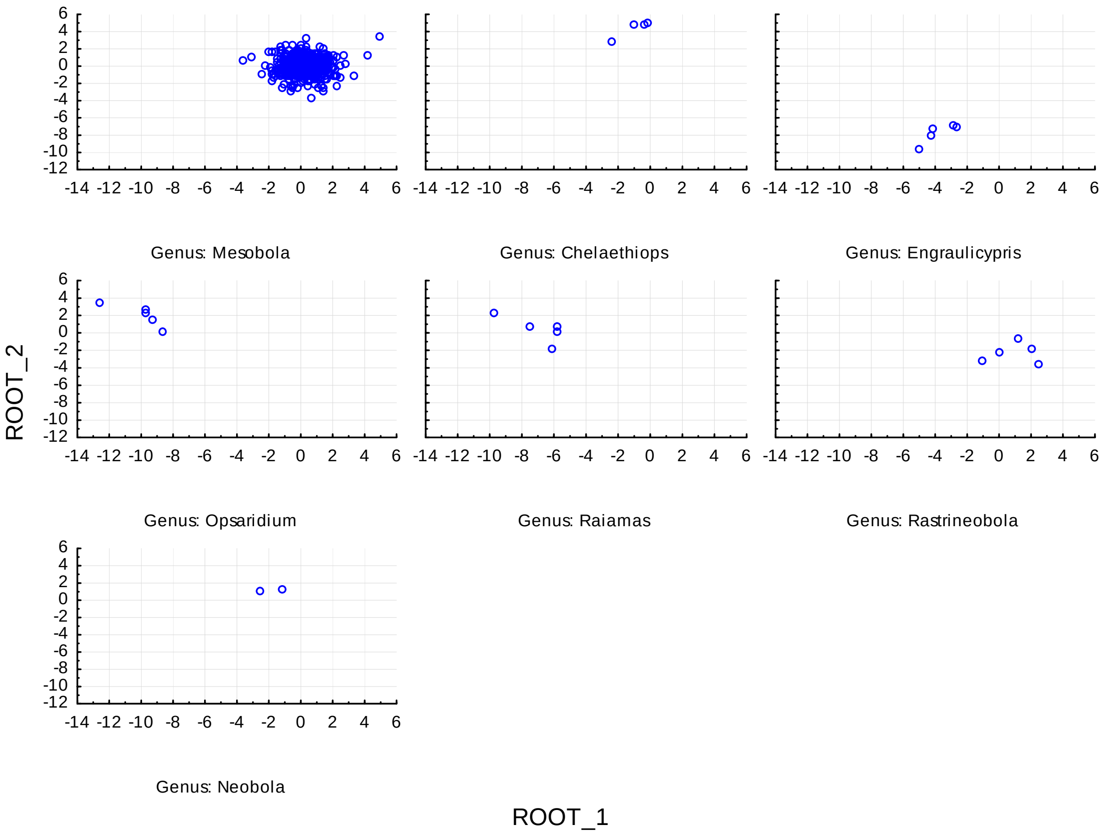


Fig. 3.9. Discriminant function Root 1 against Root 2 at genus level plotted individually, showing no allometric effect.

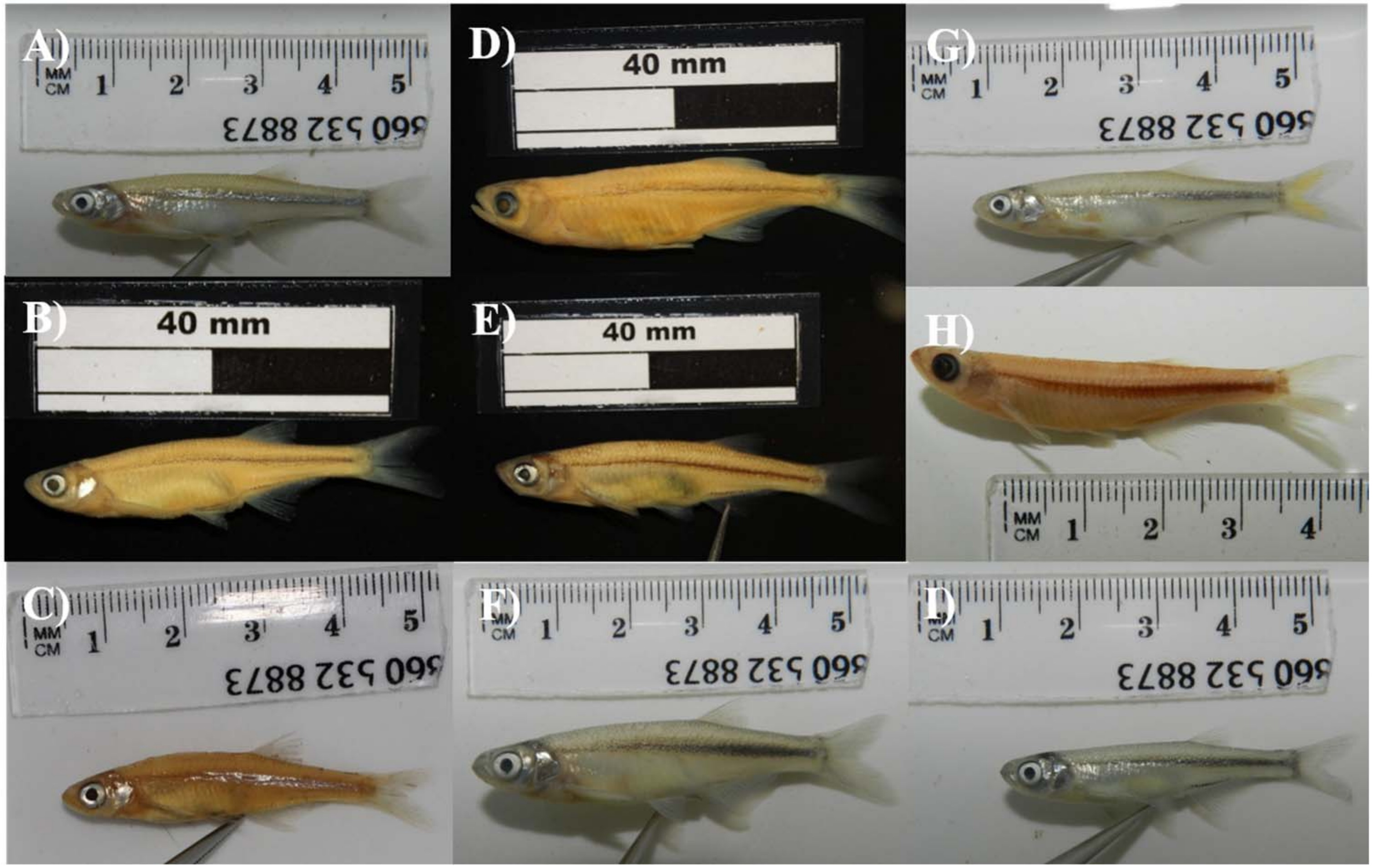
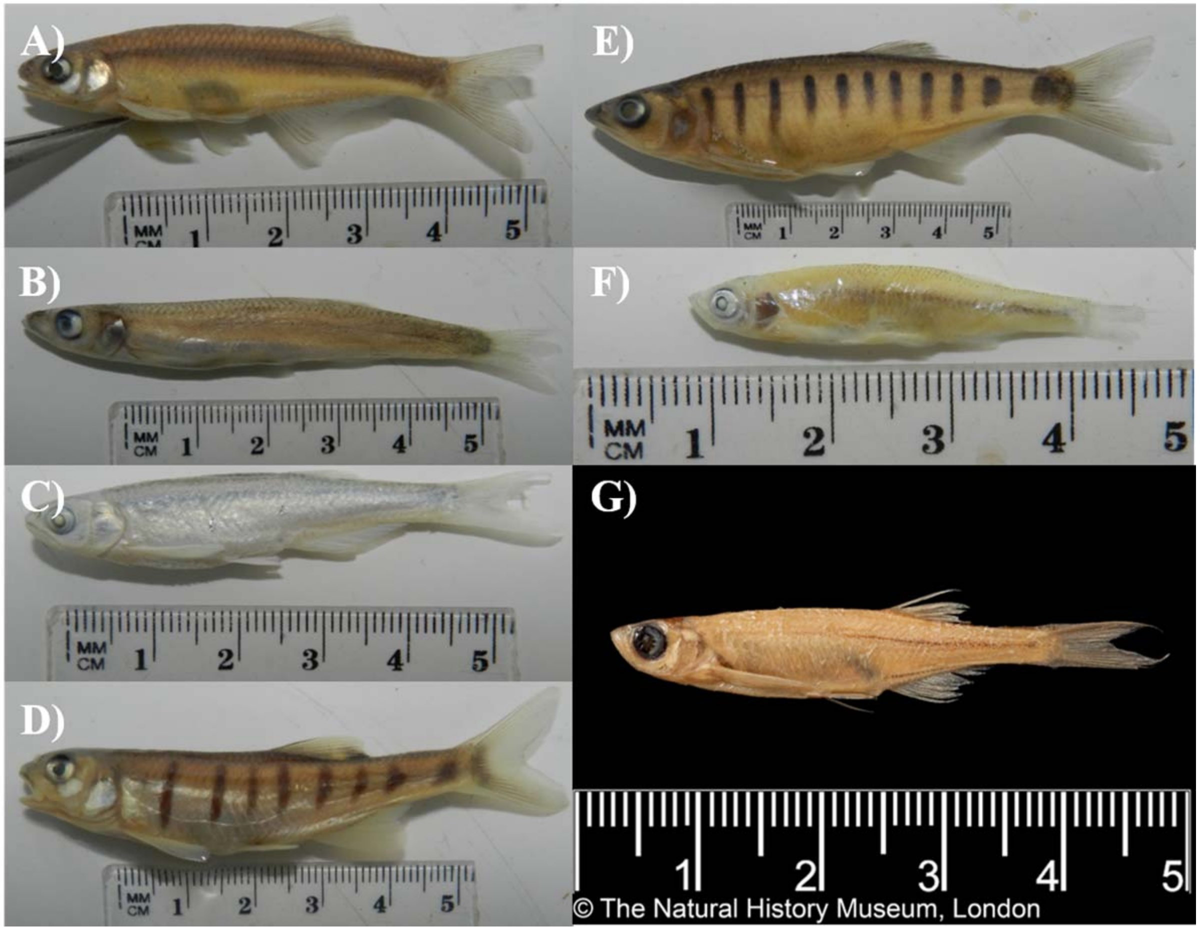


Plate 3.1. Specimens of *M. brevianalis* from the, A - eastern Limpopo System, B - western Limpopo System, C - Kunene System, D - Orange System, E - Rovuma System, F - St Lucia System, G - Pongolo System, H - Zambezi System, I - Nkomati System.



© The Natural History Museum, London

Plate 3.2. A - *Chelaethiop*s, B - *Engraulicypris*, C - *Neobola*, D - *Opsaridium*, E - *Raiamas*, F - *Rastrineobola*, G - *M. spinifer*.

4 The paleobiogeography of *Mesobola brevianalis* and the southern African Chedrini.

4.1 Introduction

The biogeography of freshwater fish species can act as an indicator of landscape evolution (Goodier *et al*. 2011) and has enabled research into evolutionary pathways (Stewart 2001). This is because the evolution of freshwater fish biogeography is determined by ecological responses to habitat change, particularly large paleobiogeographical events that caused alterations to large drainage systems (Goodier *et al*. 2011) as they have limited dispersal capabilities (Day *et al*. 2013). It is presumed, therefore, that biogeographical patterns within phylogenies can act as 'spatial-temporal signatures' for the region's history and through paleobiogeographical evidence of events, concurrence of events and biogeographical populations can be found (Goodier *et al.* 2011). Paleobiogeography, the study of patterns and links in life on Earth, enables the discovery and understanding of dominant mechanisms of the speciation process (Lieberman 2000). Most significantly, it allows for the tracking of the 'co-evolution of Earth and the Earth's biota' (Lieberman 2005).

Genetic variation within extant species and populations of African tigerfish (genus *Hydrocynus*) has served as a 'historical archive' that broadly linked tigerfish evolution to geological events (Goodier *et al*. 2011). Conversely, Quaternary geological events were used to determine mtDNA divergence rates of galaxiid fish within isolated drainage systems in New Zealand (Craw *et al*. 2008). The phylogenetic and morphometric relationships of *Mesobola brevianalis* described in Chapters 2 and 3 show interesting patterns but the evolutionary timing and driving forces behind these patterns are unknown and may be linked to paleobiogeological events.

To understand a region's biodiversity it is essential to know its history (Stankiewicz & de Wit 2006). Several authors have attempted to explain the distribution of African freshwater fishes, starting with Boulenger (1905) with a comprehensive list of species and habitats. Later studies presented reasons for the current distribution of indigenous and endemic African freshwater fish species (Roberts 1975; Skelton 1994). Skelton (1994) described comprehensively the diversity and distribution of freshwater fishes in southern Africa. While Roberts (1975) divided southern Africa into ten ichthyofaunal regions, three of which fell within the area for this study: the East Coast, Zambezian and Cape Ichthyofaunal Provinces. Each of the regions is considered to be unique in terms of its biodiversity and underlying model of biogeography (Roberts 1975).

The birth of Africa began at the break-up of Gondwana 120-180 million years ago (MYA) (Stankiewicz & de Wit 2006). With most of the northern regions of Africa under water and only the southern parts of Sub-Saharan Africa above sea level during the breakup, a large amount of uplift and stream capture over millions of years resulted in Africa's current river system assemblage and the biodiversity we see today (McCarthy & Rubidge 2005; Stankiewicz & de Wit 2006). Since the breakup of Gondwana in the early Cretaceous, the great array of changes in the characteristics of the rivers in Africa appear to have been a starting point for a number of large speciation events (Moore & Larkin 2001; Goudie 2005). Changes in the catchments and courses of large river systems have included river captures and the formation of waterfalls, including the Augrabies Falls on the Orange River and Victoria Falls on the Zambezi River, or cataracts such as Stanley Pool on the Congo (Goudie 2005). The age of the continental fluvial sediments that were deposited in the Mesozoic and Cenozoic reflect the river systems that may have been present before and after the breakup of Gondwana, and this aids in understanding the systems of this time (Moore & Larkin 2001). It is possible to assess the timing and the magnitude of changes that occurred in Africa's large drainage systems from the sedimentary sequences that reflect changes in the supply of terrigenous sediments, particularly along the east coast of southern Africa (Moore & Larkin 2001).

Africa is unique in its drainage formation because it is the only continent characterized by passive rather than active plate margins (Goudie 2005). Africa is also strongly dominated by 'basins, faults, rifts and topographic swells which differs from other continents dominated by compressional mountains' (Goudie 2005). A large event that had a great effect on the river drainage systems in southern Africa was the development of north-south rifting that resulted in uplift in eastern African (Otero *et al.* 2009). This led to the formation of the present East African Rift System, creating a new suite of ecological barriers and river system characteristics (Otero *et al*. 2009).

It is recognized that paleobiogeographical methods have limitations that hinder the comprehensive formulation of scenarios of evolution and speciation; these include incomplete paleontological evidence and extinction constraints (Lieberman 2005). Fortunately, while evolutionary scenarios may be only partially supported, knowledge accumulates gradually to clarify them. Although little is known about *M. brevianalis* and its sister taxa, a study of the relationships among its populations and the possible events that resulted in their current biogeography will add to our knowledge of the region. The recent availability of molecular clock analysis enables biologists to estimate the timing of speciation events and periods of divergence (Drummond & Rambaut 2007), enabling the linking of these fishes' past biogeography to ancient tectonic, climatic and environmental changes.

This chapter aims to use molecular clock analysis of the COI mitochondrial sequences from Chapter 2 to estimate the ages of the speciation events shown by *M. brevianalis* and its associated species and to review literature to link paleobiogeographical events to the current diversity and biogeography of the fishes.

4.2 Materials and methods

4.2.1 Molecular clock analysis

A total of 58 sequences of the COI mitochondrial marker, presented in Chapter 2, were aligned in ClustalX (Larkin *et al*. 2007) for a molecular clock analysis. As there was no mutation rate available for COI for Cyprinidae to calibrate the molecular clock, a rate was estimated. This was done by using the ages of three fossils (Table 4.1) associated with Asian and African cyprinid outgroups represented in a study by Wang *et al*. (2013). To accommodate the fossils, 121 additional sequences of African and Asian cyprinid taxa (Wang *et al*. 2013) were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and added to the alignment in ClustalX. A Nexus file was then generated.

The Nexus file was imported into BEAUti (Drummond *et al*. 2012), and an uncorrelated log-normal relaxed clock (UCLN) with ucld.mean (prior of mean substitution rate) set at CTMC Rate Reference (Ferreira & Suchard 2008; Wang *et al.* 2013) was fitted. A model of best fit was estimated using MrModelTest (Nylander 2004) and found to be TVM+I+G with an AIC weighting of 38661.11. The most closely related model available in BEAST was the TN93 model. The tree prior was set as a Yule speciation process with a random starting tree. Three taxon sets were created in BEAUTi based on the fossils presented in Wang *et al*. (2013). The age of the node subtending each fossil taxon and their sister groups (Table 4.1) was constrained with a soft maximum, a lognormal distribution and the mean, standard deviation and offset as described by Wang *et al*. (2013). Only three of the four fossil calibrations used by Wang *et al*. (2013) (Table 4.1) were used to create the BEAST file as no alignable COI sequence was found for *Pogonichthys*, the apparent sister group of the fossil *Mylocheilus* (Wang *et al*. 2013). Following Wang *et al*. (2013), the root age of the tree was set using a normal distribution with a mean of 170.5 and a standard deviation of 10.6 to serve as another calibration point.

A molecular clock analysis was performed in BEAST (Drummond & Rambaut 2007) using a total of 100 million generations sampled every 1000 generations. The BEAST file was loaded into Tracer (Rambaut & Drummond 2007) to determine whether the effective sampling sizes (ESS) were sufficient. Once sampling size was sufficient, the file was imported into TreeAnnotator (Rambaut & Drummond 2006) and 10-25% of the run length was chosen with a discard value of 10000. The tree generated was viewed and edited using FigTree (Rambaut 2009), and most of the outgroups were collapsed in order view the divergence times of the *M. brevianalis* radiation.

4.3 Results

The molecular dating analysis showed that *M. brevianalis* and outgroups have a particularly old lineage for Afrotropical Cyprinidae, beginning around 125 million years ago (Fig. 4.1). *Opsaridium* was the first distinctive lineage to arise ~75 million years ago with a strong posterior node value of 0.9993, followed by *Chelaethiops* at ~ 65 million years ago with a much weaker posterior node value of 0.57 (Fig. 4.1). It can be seen that there was a divergence of *Opsaridium* followed closely by a *Chelaethiops* divergence ~65 million years ago (Fig. 4.1). The next split only occurred between 30-35 million years ago, resulting in the branching of *M. brevianalis* Kunene into its own distinct clade with an extremely strong posterior node value of 1 (Fig. 4.1). This was closely followed by a spit of the remaining *M. brevianalis* populations into one group with *E. sardella* and Rovuma System *M. brevianalis* specimens and a second group comprised of *R. argentea* and all of the remaining *M.*

brevianalis taxa ~27 million years ago with a strong posterior value of 0.91 (Fig. 4.1). *E. sardella* and Rovuma System *M. brevianalis* populations then split with a posterior node value of 1 occurring later ~20 million years ago (Fig. 4.1). The second group saw an earlier split at ~ 23 million years ago with a lower posterior node value of 0.79, into a group containing *M. brevianalis* Orange River specimens and *R. argentea*, and a group containing all of the eastern South African *M. brevianalis* populations (Fig. 4.1). The *M. brevianalis* Orange River and *R. argentea* specimens rapidly split from one another to form their own distinct clades at ~ 21 million years ago with an unexpectedly low posterior node value of 0.1973. This occurred only a few million years after their split with the eastern South African *M. brevianalis* populations (Fig. 4.1). The eastern South African *M. brevianalis* populations, including the type locality (Mkhuze) showed varied levels of branching for a few million years and only show distinct divergence during last 4 million years (Fig. 4.1). It can be seen that the first isolation was the western Limpopo specimens ~4 million years ago (Fig. 4.1). Other branches can be seen, including the Nkomati system, then the Pongolo system and finally the St Lucia and eastern Limpopo specimens, but this occurred very recently, mostly within the last 3 million years, giving little time for divergence (Fig. 4.1).

The node height bars (95%_HPD) were long for the deeper splits between genera and then shortened with species-level splits but the error margins were still quite large, often covering 50 million years (Fig. 4.1). These larger error bars could be seen to be present in the *Opsaridium – Chelaethiops* split and the *Chelaethiops* – (*M. brevianalis* + *R. argentea* + *E. sardella*) split (Fig. 4.1).

4.4 Discussion

The molecular clock analysis of the COI gene provided a basis for linking paleogeographical events to diversification events in the phylogeny. The uncertainties around the nodes of the chronogram that were sometimes relatively large, may have arisen from using a sister clade to estimate the substitution rate for the focal species as there was none available in the literature for Cyprinidae. Using only three of the four fossil calibrations used by Wang *et al.* (2013) may have further compromised the confidence intervals. However, the use of a first estimate is sufficient to link paleobiogeographical events to genetic variation and biogeography (Crisp *et al.* 2011; Goodier *et al.* 2011).

Despite large advances in the understanding of relationships and the current distribution of the Earth's biota, many paleohydrological and paleogeological models are still speculative (Goudie 2005). However, increasing confidence in these models, due to advances in remote sensing, radiometric dates, offshore cores, sonar and seismic studies, tracer studies and plate tectonics (Goudie 2005) means that it is possible to hypothesize informed paleobiogeographical models of evolution for *M. brevianalis* and its associated species.

As in other freshwater fish lineages, and within these interpretive constraints, the radiation of the southern African chedrins mirrors the prior continental river connections and the drastic geological events that took place in Africa (Stewart 2001).

4.4.1 Origin of the African Chedrini 170-80 MYA

The monophyly of the African Danioninae and their origin from an Asian ancestor is supported by several studies (Gaubert *et al*. 2009; Tang *et al*. 2010; Liao *et al*. 2011). This study further supported an Asian origin by recovering two Asian taxa, *Salmostoma bacaila* and *Barilius bendelisis*, as sister groups to the African Danioninae in the molecular clock analysis (Fig. 4.1). These sister groups are members of an Asian radiation which differentiated earlier than the African clade. This indicates a prior Asian diversification that may have given rise to the African clade. The Asian origin is supported by Tang *et al*. (2010), who traced a single African origin from an Asian Danioninae ancestor and Gaubert *et al*. (2009) who located the origin of African and Mesopotamian *Barbus* in a single Oriental-Afrotropical migration. There appears to be a number of faunal elements that are connected to Asia but it is speculative as to how they moved across the continents and intervening ocean, and whether it was as a result of a shared Gondwana ancestral population or if the movement was post-Gondwana through immigrations (Stewart 2001) possibly using the Middle East as a connection (Gaubert *et al*. 2009).

In the chronogram (Fig. 4.1), the Asian genus *Barilius* is linked to the genus *Raiamas*, which is primarily African, but includes *R. guttatus* and *R. bola* from south eastern Asia. This link is weakly supported (pp = 0.3931), so it offers limited contradiction to the hypothesis that the African chedrins are monophyletic (Tang *et al*. 2010). However, it does suggest that the migration of the chedrins to Africa occurred later in the Cretaceous. The very early divergence of the African chedrins however, is questionable as current opinions place other cyprinids as diverging much later than the results suggested as seen in Laebeines and Barbines that have been thought to diverge from an Asian ancestor in the Early Miocene (Tang *et al*. 2009; Tsigenopolous *et al*. 2010). It is then suggested that due to the large margin of error (node height bar) seen for the African chedrin divergence that the period of this divergence would lie towards a more recent divergence on the furthest right point of the error bar around 75 MYA. This would be in support of work performed by Near *et al*. (2012) on ray-finned fish which found divergence of African Cypriniformes to occur around 75-80 MYA. The strong support (pp = 0.9993) for the sister clade containing *Chelaethiops*, *Opsaridium*, *Neobola*, *Rastrineobola*, *Engraulicypris* and *Mesobola* indicates that it was established by the end of the Cretaceous period (Fig. 4.1), probably in north eastern or Central Africa.

4.4.2 Origin of Opsaridium 95-65 MYA

The origin of *Opsaridium* appears to have occurred 95-65 MYA, during the late Cretaceous period and Early Mesozoic (McCarthy & Rubidge 2005). This period coincides with a large amount of interior continental landscape lowering (Partridge 2008) and western uplift along the passive margin of the rift with South America (Stankiewicz & de Wit 2006). The genus shows a pattern of differentiation that starts in east Africa and moves south along the east coast of the continent (Fig. 4.1), starting about 40-50 MYA.

4.4.3 Origin of Chelaethiops and Mesobola 80-50 MYA

Chelaethiops showed divergence from its sister group 83-50 MYA, possibly due to the same general events that affected *Opsaridium* to the east. The Congo Basin was below sea level for most of the Upper Jurassic and the Lower Cretaceous periods, ~200-129 MYA (Stankiewicz & de Wit 2006). Marine fossil and sediment source evidence supports the theory that the Congo Basin was raised above sea level in the Cenozoic period, and it may have become the home for many northern and western species of fish as well as a drainage point for paleo-river systems (Stankiewicz & de Wit 2006). Climatic changes beginning in the Early Cretaceous may have facilitated the creation and maintenance of a lacustrine environment (Giresse 2005), enabling the existence of a large basin drainage system. Furthermore, Stankiewicz and de Wit (2006) supply evidence that the Congo Basin drained east into the Proto-Indian Ocean (Stankiewicz & de Wit 2006) in comparison to its westerly flow today (Fig. 4.2). Peneplain sources from the Late Cretaceous indicate that during the period when *Mesobola* and *Chelaethiops* arose, the Congo Basin became tectonically stable and geological factors combined to facilitate drainage movement to the east (Stankiewicz & de Wit 2006). When South America detached from Africa ~120-130 MYA, a passive Atlantic Rift formed on the western side of Africa (Fig. 4.2) resulting in the flow of water within the Congo Basin being directed predominantly to the south-east to drain out into the Proto-Indian Ocean (Stankiewicz & de Wit 2006; Roberts *et al*. 2012). This general uplift along the margins of southern Africa compelled most of the present southern African paleo-rivers to flow south-eastwards into the Paleo-Indian Ocean (Haddon & McCarthy 2005). This eastward drainage is confirmed by a number of species that appeared to diverge into the eastern side of Africa, predominantly during the Miocene (Stewart 2001). The Congo Basin continued to drain to the east until the time of the East African highland uplift that occurred in the Oligocene and the Eocene, ~30 to 40 MYA (Fig. 4.5) (Stankiewicz & de Wit 2006). This is supported by paleogeography of *Synodontis* suggested by Day *et al.* (2013), attributing the Congo River as the source of dispersal of populations only, followed by range expansion which led to the eventual divergence and isolation within East and West Africa.

During this period it is possible that the ancestors of *Mesobola* and *Opsaridium* moved south via the Paleo-Luangwa or Paleo-Chambeshi (Stankiewicz & de Wit 2006) (Fig. 4.2), and became isolated to give rise to the south eastern *Opsaridium* species that radiated 35-15 MYA and appeared to follow the beginning of the uplift of the East African highlands (Fig. 4.5). Furthermore, the Paleo-Chambeshi appears to have linked with the Paleo-Kafue, which flowed in a south-westerly direction from the eastern side of Africa, together forming a large south-flowing drainage system (Moore *et al.* 2012).

The diversification of *Chelaethiops,* then, correlates with evidence that the Congo Basin became landlocked for some time following the East African uplift, particularly the Ovambo-Kalahari-Zimbabwe (OKZ) axis uplift (Moore *et al*. 2012). This would have isolated a population within the basin or surrounding rivers until the Congo Basin was captured by a small west-coast river that drained it into the Atlantic (Stankiewicz & de Wit 2006) (Fig. 4.2). Therefore, allowing for speciation to occur within the system to produce the radiation of *Chelaethiops* presently found in the Congo River System.

The evidence then follows that prior to the formation of the OKZ uplift (Moore *et al.* 2009), the Congo drainage system's easterly flow and down a Paleo-Chambeshi and Paleo-Luangwa would have connected it to a number of southern systems (Stankiewicz & de Wit 2006). This is postulated to have linked the Luangwa and Kafue to the upper Zambezi, Cuando and Okavango Rivers all of which would have drained out through a southern system, possibly the Limpopo (Moore & Larkin 2001; Stankiewicz & de Wit 2006). The presence of these systems prior to rift formation could have facilitated the movement of the ancestral fishes southwards from the Congo Basin into the southern systems between the Cretaceous and the Paleocene (Moore & Larkin 2001).

4.4.4 Paleo-Kalahari Lake as origin of Mesobola

Following the divergence of *Chelaethiops* and *Opsaridium* the lineage leading to *Mesobola* appeared to remain unchanged until the divergence of a population that is presently found only in the Kunene (Fig. 4.1), ~ 45-23 MYA in the Eocene and Oligocene. Although the chronogram shows the isolation of the Kunene, Orange and Rovuma Rivers and that the eastern South African populations are separate, the estimated times of the events overlap to a large extent (Fig. 4.1) and there may be evidence that a single event or interrelated events caused these four speciation events from the same ancestral population.

Once an ancestral population was within the Paleo-Chambeshi or Paleo-Luangwa, its southward movement could have been through the upper Zambezi from a Paleo-Chambeshi which was draining into the Indian Ocean through the Limpopo (Moore *et al.* 2012) or through the paleo-Luangwa into the southern drainage systems as mentioned (Moore & Larkin 2001). However, the results of this study raise the possibility that a large paleo-lake existed, centred in Botswana (Fig. 4.3). This is a relatively novel postulation which is not explicitly addressed in the literature but models have been suggested prior to this work (McCarthy 1983; McCarthy 2013). The following discussion considers evidence from tangential sources to explore its viability.

The Kalahari Basin is a large basin filled with deposited sands tens to hundreds metres in depth over the area (de Vries *et al.* 2000). The breakup of Gondwana facilitated uplift of the passive margins of the continental rift and downwarping within the interior of Africa which gave rise to the Kalahari Basin, also known as the Kalahari-Cubango-Congo Basin (Thomas

& Shaw 1991; McCarthy 2013). The true extent of the Paleo-Kalahari Basin is speculative but the present drainage system stretches from the Molopo dry valley in the south, 600 km north-east to the Makgadikgadi Pan and beyond the Okavango delta towards Angola in the north-west, covering an area of 60 000-80 000 km^2 (de Vries *et al*. 2000; Moore & Larkin 2001). Kalahari Group sedimentary rocks cover an area of over 2 200 km across northern South Africa through Botswana, Angola and the Democratic Republic of Congo and are thought to have been deposited mainly during the late Cretaceous (Haddon & McCarthy 2005; Moore *et al*. 2012). McCarthy (1983) further spoke of the shales and marls filling the Kalahari Basin which are suited to have origins of formation in a lacustrine environment. Furthermore, Moore *et al*. (2012) depict Miocene drainage as occurring from the north-east (Paleo-Chambeshi) from one single large river system out into a number of endorheic delta-like streams in the southern regions of Africa. It has been postulated that the warmer climate in the middle Miocene (Cooke 1980; Condamine *et al*. 2012; McCarthy 2013), the large contemporary downwarping that occurred in the region and present-day evidence of large, flat pans within the Kalahari Basin led to the forming of a series of inland lakes (Moore & Larkin 2001; McCarthy 2013).

It is proposed here that instead of the existence of several smaller lakes, the Kalahari Basin was occupied by a large paleo-lake system inhabited by the ancestral population of *Mesobola*. It has been hypothesised that the Sowa and Makgadikgadi Pans form major endorheic paleo-lakes in their own right (Cooke 1980; Moore & Larkin 2001). Under the correct conditions, it is possible that the large Kalahari basin area could have accommodated a large lake that subsumed the Sowa and Makgadikgadi Pans, the Okavango Delta and perhaps even the Etosha Pan as suggested by McCarthy (2013). The Kalahari 'sand sea', described by Thomas and Shaw (1993), covers 2.5 million km^2, but little is known of how the sand accumulated in this region. The Kalahari Basin formed as a result of the breakup of Gondwana with later stages of the breakup causing uplift in the African margin (Thomas & Shaw 1993; Haddon & McCarthy 2005; McCarthy 2013) and Thomas and Shaw 1993 spoke of this uplift resulting in the formation of an 'extensive intracratonic, intercontinental basin which acted as a watershed providing a drainage system from coastally-oriented systems into the African interior', which could account for the sand deposits. For some years, the geological literature has hypothesised the existence of a large southern drainage that has previously been interpreted as a large paleo-river system or network of rivers (Moore *et al*. 2012; McCarthy 2013) but this large system must have drained into a region hypothesised here to be the large

Kalahari 'sand sea'. Therefore, although this theory has received little attention prior to this study, the evidence cited and the species divergence seen within *Mesobola* and associated species provides support for a Paleo-Kalahari Basin containing a large paleo-lake. In this discussion, this will be referred to as the Paleo-Kalahari Lake (Fig. 4.3).

The uplift that forged the Kalahari Basin along the epeirogenic axes is also thought to have resulted in back-tilt of paleo-rivers into the Paleo-Kalahari Basin which resulted in the sedimentation accumulation of the Kalahari Group Sediments (Haddon & McCarthy 2005). These rivers could have fed a large endorheic Paleo-Kalahari Lake.

Furthermore, the existence of such a lake would have coincided with a period of warmer climatic conditions following the Cenozoic period and breakup of Gondwana (Condamine *et al*. 2012). This would have facilitated the formation and persistence of a large interior lake. Oxygen isotope records (Condamine *et al*. 2012) indicate that the climate was as much as 4°C warmer on average than it is now (Condamine *et al*. 2012; McCarthy 2013). This would be associated with significantly higher rainfall and support accumulating evidence that mega-lakes have a positive feedback effect on the local climatic conditions by raising precipitation rates and maintaining a warmer and more moderate climate (Krinner *et al*. 2012). This was observed in the Eocene Green River Lake System, where the paleo-lake could raise the minimum winter temperature (Greenwood & Wing 1995), and in mid-Holocene mega-lake systems in the Sahara that enabled the development of a "Green Sahara" (Krinner *et al*. 2012).

It is postulated here that there was sufficient rainfall to push the shores of the Paleo-Kalahari Lake substantially towards the margins of the Kalahari Basin and, based on speculation, perhaps almost as far as the Kalahari sand sea stretches.

If the Paleo-Kalahari Lake was fed by a Paleo-Chambeshi drainage from the north, an ancestral population of fishes could have moved south to spread across much of the interior drainage of southern Africa (Fig. 4.3). Cichlid phylogenetics has allowed the discovery that the River Kwanza, upper River Zambezi and River Congo aided in the radiation proposed within the Kalahari Paleolakes (Schwarzer *et al.* 2012). There is evidence of the presence of *Mesobola* in the Okavango system during the Holocene: a single collection of *Mesobola* [cf.] *brevianalis* was made in the Okavango in 2000 (SAIAB 20379) but they have not been collected in the system since then.

4.4.5 Isolation of a Kunene population 54-25 MYA

Moore and Larkin (2001) hypothesised a connection between the Paleo-Chambeshi or Paleo-Luangwa with the Okavango and Cuando Rivers (Moore & Larkin 2001; Stankiewicz & de Wit 2006). This connection could have also been facilitated by a Paleo-Kalahari Lake (Fig. 4.4). Ancestors of *Mesobola* may have swum up the north-western rivers which would have had a south-easterly flow due to back-tilting of rivers caused by the Atlantic Rift (Fig. 4.4) (Haddon & McCarthy 2005). Extant fish fauna show strong evidence of the prior connection of the Okavango and Cubango Rivers (Moore & Larkin 2001). Once in the headwaters of these rivers, a population may have crossed into the Kunene River through river capture near the source, flooding in the dunefields around the headwaters, or possibly even by swimming directly up the Paleo-Kunene River, which seems to have drained into the Etosha Pan before it was captured by a small coastal river. Wellington (1955) provides evidence from the present day river pattern of southwards flow of the Paleo-Kunene towards the Etosha pan as it flows southwards for most of its course before it sharply changes direction to flow westwards (Fig. 4.4). There is evidence that the Paleo-Kunene River was captured towards the north-west due to backward erosion of a river from the coast (Fig. 4.4) (Goudie 2005).

Further evidence of capture by a coastal drainage system is seen north of the Kunene, in the Rio Coroca System, which flows southwards before sharply turning to the west (Wellington 1955). The prior connection of more south-eastern rivers with the Kunene is further supported by evidence that the Kunene's Atlantic mouth was originally south of its current location and moved further northwards due to dune encroachment (Fig. 4.4) (Goudie 2005). The crossing of the Kunene population may have been facilitated by a connection with the Cubango which is hypothesised to have formed the head waters of the Colui (a Paleo-Kunene tributary) within the Okavango (Fig. 4.4) (Moore & Larkin 2001). The Paleo-Kunene River may have been connected to the Okavango through any one of these means for a sufficient period to allow a population to move into the system. *Synodontis* showed a similar isolation during the Miocene within the Kunene System and therefore may have been as a result of such system movement (Day *et al*. 2013).

It is hypothesised that the eventual isolation of the Kunene population within this system was due to the onset of cooler climatic conditions that began in the Eocene era, coinciding with divergence seen within the time clock (25-45 MYA), which would have resulted in dryer conditions with lower rainfall (Fig. 4.4) (Condamine *et al*. 2012), that spread across Africa.

Such drying conditions would have had the potential to lower the level of the Paleo-Kalahari Lake, dry up peripheral drainages, particularly in the west, and isolate the Etosha pan if it was connected to the paleo-lake. The Kunene *Mesobola* population could therefore have been isolated by the capture of the Paleo-Kunene or by contraction of the paleo-lake. If the Paleo-Kunene was a tributary of the paleo-lake before its capture, that event would have contributed to the fall of the lake's level further.

4.4.6 Mesobola in the Rovuma River 30-20 MYA

The differentiation of the *M. brevianalis* eastern South African Systems, Kunene System, Rovuma System, Orange System, *E. sardella* and *R. argentea* occurred roughly simultaneously during a period after the Early Oligocene glaciations and around the Late Oligocene warming (Plana 2004; Condamine *et al*. 2012). As was stated previously, the marginal uplift that began as a result of the breakup of Gondwana caused the back-tilt of paleo-rivers into the Paleo-Kalahari Basin (Fig. 4.3) (Haddon & McCarthy 2005). With a population movement facilitated via a Paleo-Chambeshi (Moore *et al*. 2012), or a Paleo-Luangwa River (Moore & Larkin 2001), the population would have been able to move through the systems connected to the hypothetical Paleo-Kalahari Lake (Fig. 4.5). As illustrated by Moore *et al*. (2012) the Okavango, Cuando, Cubango, Limpopo and upper Zambezi Rivers were connected to a large system that drained into the Kalahari Basin. An ancestral population in the Paleo-Luangwa, connected to the lower Paleo-Zambezi, could have swum up the lower Zambezi and given rise to present day *E. sardella* and Rovuma *Mesobola* populations following uplift (Fig. 4.5).

A number of studies have provided support for tectonism's ability to reshape continental drainage (Thomas & Shaw 1988; Malik & Mohanty 2007; Moore *et al*. 2012). Moore and Larkin (2001) proposed that the extension of the East African rift system would also have supplied a mechanism for the isolation of populations, as its southward movement transected their rivers. The rifting uplift of eastern Africa was a large event that resulted in great changes to African drainage patterns (Roberts *et al*. 2012). The present East African rift system is thought to have begun 30 million years ago (Goudie 2005), placing it within the time period of the divergence of *M. brevianalis* and its associated species (Fig. 4.5). The Paleo-Rukwa, Paleo-Congo, Paleo-Zambezi and Paleo-Nile systems all experienced large changes and re-routing that continued throughout the Cenozoic period (Roberts *et al.* 2012). By 25 MYA, the rift system had spread over 4000 km from the Gulf of Suez to the Mozambique Channel (Fig. 4.5) (Goudie 2005). The speciation of the Rovuma river population of *Mesobola* and *E. sardella* could therefore be explained by the spread of the East African rift system into southern Africa, estimated at about 20 MYA by McCarthy & Rubidge (2005) and by others as spreading 45-30 MYA (Stankiewicz & de Wit 2006; Moore & Larkin 2001). This rift (Moore & Larkin 2001; Stankiewicz & de Wit 2006; Moore *et al*. 2009; Moore *et al*. 2012), may have resulted in the severing of the Paleo-Luangwa from the Paleo-Rufigi and the Paleo-lower Zambezi, giving rise to a Rovuma/Malawi ancestral population 10-30 MYA.

4.4.7 Engraulicypris sardella 8-4 MYA

The Zambezi River also experienced the Kalahari-Zimbabwe Axis flexure which was part of the larger Ovambo-Kalahari-Zambezi (OKZ) (Moore *et al*. 2009). The increase in sediment flux created as a result of this uplift resulted in a complete change in the Zambezi and Limpopo marine deltas (Moore *et al*. 2009). This may have led to further isolation of the lower Zambezi and Shire Rivers in addition to the Limpopo River in 30-10 MYA. The population in the Paleo-lower Zambezi via the Paleo-Luangwa, had the ability to move up a Paleo-Shire River and into the Paleo-Rovuma System prior to the development of the East African rift system, leading to possible later isolation of an ancestral population within the Rovuma System. The present day Rovuma System shows its drainage reaches the point of rift margin uplift directly opposite Lake Malawi where it dramatically changes direction due to deflection along the edge of the rift margin (Google Earth 2014). Its Lucheringo River tributary drains from Lake Chiuta in the south (Google Earth 2014). This may represent a previous connection between the Shire and Rovuma Rivers before rifting and uplift occurred, that would have isolated populations in both systems (Fig. 4.4).

The uplift suggests a mechanism of early crustal warping causing drainage disruption that would have isolated a Lake Malawi ancestor and a Rovuma River ancestor. This could have begun prior to rifting of the Nyasa region, estimated by the molecular clock analysis to have occurred 25-10 MYA and not in the Late Miocene period (Moore *et al*. 2009). The Lucheringo River then flowed southwards to connect with Lake Chiuta which it presently drains (Lancaster 1981). There is evidence that Lake Chilwa, presently a closed saline lake, was separated from Lake Chiuta by a small isolated uplift or sand bar (Lancaster 1981) that looks like a dyke on aerial photographs (Google Earth 2014). The fish species found within Lake Chilwa and Lake Chiuta are all present in the Rovuma River, showing a recent link between these two lakes (Lancaster 1981). The sand bar formed in the Holocene period, but it is unknown whether there is presently a population of *Mesobola* in Lake Chilwa, which dries out considerably at times.

Lake Malawi, once known as Lake Nyasa or Niassa (Weyl *et al.* 2010), is relatively young (Delvaux 1995). Evidence supports its formation in the Late Miocene (~8 MYA) with a Plio-Pleistocene (4.5 MYA) age for its first deep lake formation (Delvaux 1995; Danley *et al.* 2012) which would have supported the adaptation of *E. sardella* to a pelagic life beginning ~4 MYA (Fig. 4.6). The Malawi Rift Basin began to form ~8.6 MYA during the propagation of the East African Rift System (Delvaux 1995; Danley *et al.* 2012), which may have isolated a Shire River population within conditions with higher niche availability and food resources within the lake. However, a drying period followed during which the lake subsided drastically, and it is speculated that it achieved sufficient depth to support a lake biota only about 4 MYA (Delvaux 1995; Danley *et al.* 2012). If Lake Malawi, which is now 700 m deep in places (Danley *et al.* 2012), was endorheic during the drying phase, this may explain why *E. sardella* developed lacustrine adaptations after originating in a river. The immediate period following the deepening of Lake Malawi coincides with the genetic differentiation and divergence of the *E. sardella* populations (Fig. 4.1) (4 MYA to present). This mimics diversification seen within Cichlids which appeared to coincide with a period following the reaching of sufficient depth however, the exact timing is still uncertain for these species (Danley *et al.* 2012).

The current distribution of *M. brevianalis* within the Zambezi System is not well known. However, it has never been collected in the lower or middle Zambezi and literature supports its recent introduction into this area (Kadye 2008). In the South African Institute for Aquatic Biodiversity (SAIAB) there is a sample that was collected from the upper Zambezi in 2000, but the species has not been found in the area since then. The samples that were collected are only suitable for morphometric analysis as tissues did not result in successful marker isolation. Without relevant genetic data it is not possible to accurately recognise distributions of species currently treated as *M. brevianalis*. The lack of viable genetic tissue further prevents linking paleogeomorphology to the possible existence of an upper Zambezi population.

There is, however, support for separate evolution of the upper and middle Zambezi, with the upper Zambezi being linked to the Limpopo System or the proposed Paleo-Kalahari Lake. This would mean the facilitation of movement of a population into the system prior to being captured by the middle Zambezi (Goudie 2005), an event indicated by a capture elbow in the upper Zambezi (Wellington 1955). It is hypothesised that a former link of the upper Zambezi to a southern drainage system prior to the Oligocene would have been disrupted by uplift events that landlocked the upper Zambezi in the Pleistocene (Stankiewicz & de Wit 2006). This uplift further altered the lower Zambezi drainage, resulting in eastward erosion until the Late Pleistocene when the lower Zambezi captured the middle Zambezi and then the upper Zambezi via the Gwembe Trough (Stankiewicz & de Wit 2006). The isolation within the upper Zambezi may have endured due to downwarping of the Gwembe trough following capture by the lower Zambezi, resulting in the development of the Victoria Falls (Goudie 2005). Although it has been proved that some fish species can survive movement down the falls (Balon 1974), it has also been suggested that large differences in niche and habitat availability between the upper, lower and middle Zambezi may limit upper Zambezi species from colonising the middle and lower Zambezi (Balon 1974).

This hypothesis would support a separate isolation event in the upper Zambezi through the Paleo-Kalahari Lake prior to the connection to the middle Zambezi and the possible extinction of an isolated populations in the lower Zambezi (Kadye 2008) and lower Shire River, caused by a number of ecological factors (Balon 1974) including the formation of the Shire River rapids, which act as a natural barrier (Tweddle *et al.* 1979). Future studies would benefit by addressing this deficit.

4.4.8 Rastrineobola argentea

Although the molecular clock analysis showed a relationship to the Orange River *M. brevianalis,* the extremely low support value would suggest that this relationship is incorrect. Furthermore, there is no literature to suggest a viable paleobiogeographical connection between the Orange River and Lake Victoria. *Rastrineobola argentea* presently occurs in Lake Victoria, Lake Kyoga and the Victoria Nile (Froese & Pauly 2014).

It is possible that taxa missing from the analyses, such as *Mesobola spinifer* Bailey & Matthes, 1971 and *Mesobola bredoi* Poll, 1945, may resolve the anomaly (Eschmeyer 2014). These species are more northern: *M. bredoi* is endemic to Lake Albert and *M. spinifer* is found within the Malagarasi River, Ruaha River and Lake Rukwa (Froese & Pauly 2014).

Lake Victoria is thought to be a relatively young African lake, possibly only 400 000 years old, created by regional tectonic tilting (Johnson *et al*. 2000). The current Lake Victoria formed following an arid event which occurred about 14 600 years ago, which dried up the lake resulting in the lake's formation in a completely arid environment (Johnson *et al*. 2000). It is thought that this event coincided with and accounts for the formation of Lake Kyoga (Goudie 2005). These events may also explain the limited habitat of *R. argentea* to this lake and neighbouring regions.

Although we cannot accurately assess the paleobiogeography of *R. argentea* without further research, it is hypothesised that its estimated relationship to the Orange River population of *Mesobola* is incorrect, and that *R. argentea* would show close relationships to *M. bredoi* and *M. spinifer*. Furthermore, it is thought that these species diverged from the same ancestral population as *E. sardella* in Lake Malawi and the *Mesobola* population in the Rovuma System. The molecular timing of the *R. argentea* divergence shows a close timing to the divergence of *E. sardella* and *M. brevianalis* (Rovuma), 25-10 MYA, which is consistent with these predictions.

Although this relationship can only be hypothesised at this point, it would be supported by paleogeographical literature as evidence from fluvio-deltaic sands and lacustrine muds from Lake Rukwa, found in Lake Malawi, confirm a prior connection between these two lakes (Delvaux 1995). A domal uplift in the Rungwe-Ngozi area, which possibly occurred due to volcanic activity in the Rungwe volcanic province in the late to middle Pleistocene, resulted in the disconnection and regression of both Lake Malawi and Lake Rukwa from one another (Delvaux 1995). This is evident in the geography of the area today (Google Earth 2014). Lake Rukwa has been found to be sensitive to climatic conditions (Delvaux 1995), that would have aided in its isolation during the drying period (Condamine *et al*. 2012).

There is also evidence of a link between Lake Rukwa and Lake Tanganyika, which still occurs today through the flood regions of the Malagarasi River (Nicholson 1999). During a period that experienced far wetter conditions than the present (Condamine *et al*. 2012), the movement of an ancestral population from Lake Rukwa to Lake Tanganyika would have been viable. Lake Kivu, to the north of Lake Tanganyika, is still connected to Lake Tanganyika (Verheyen *et al*. 2003), facilitating movement into a Paleo-Lake Kivu. The final movement from Lake Kivu into Lake Victoria is suggested by the ability of Lake Kivu and Lake Tanganyika to have acted as deep reservoirs prior to the formation of Lake Victoria (Verheyen *et al*. 2003). It is postulated that the unique and highly intricate biodiversity of Lake Victoria, including hundreds of endemic fish species that speciated within the lake during the last 14 000 years (Johnson *et al.* 2000), arose in a 500 year period of high primary production after Lake Victoria refilled when species could have taken refuge from the arid conditions around the lake (Johnson *et al.* 2000).

Cichlid species were found to exhibit very few mutations between their populations in Lake Victoria and Lake Kivu, implying a recent link between these lakes; possibly after the late Pleistocene drying of Lake Victoria (Johnson *et al*. 1996). It is possible that an ancestral population could have moved northwards through Lake Rukwa, Lake Tanganyika, Lake Kivu and Lake Edward into Lake Victoria to give rise to *R. argentea*. This would have occurred particularly between 25-10 MYA when the climate was warmer and wetter and before the onset of the Plio-Pleistocene glaciation (Condamine *et al*. 2012). This would present a more plausible scenario than the very poorly supported relationship with the Orange River shown in Fig. 4.1, and is supported by phylogenetic patterns in, for example, cichlids, that migrated from the geologically older Lake Kivu and Lake Tanganyika into Lake Victoria (Verheyen *et al*. 2003). Further research is required before this species can be confidently placed within the genus and comment made on its relationship, particularly through the inclusion of *M. bredoi* and *M. spinifer*. These species are hypothesised to be the ‘missing links’ to the *R. argentea* speciation events.

4.4.9 Mesobola in the Limpopo and Orange Rivers 25-15 MYA

The speciation of the southern *Mesobola* populations, including the Orange River and eastern South African populations from the Limpopo River to St Lucia, may have begun when the East African rift system propagated into southern Africa. This may have been due particularly to the flexural uplift ~20 million years ago along the Kalahari-Zimbabwe Axis (McCarthy & Rubidge 2005) and the Ovambo-Kalahari-Zambezi Axis flexural uplift (Fig. 4.5) parting the Limpopo River from the endorheic system in the Kalahari Basin (Moore *et al*. 2012), which is the postulated Paleo-Kalahari Lake. This would have coincided with global paleoclimatic cooling and possible consequent positive feedback effects on local climate brought about by the shrinking of the Paleo-Kalahari Lake and diminished inflow from the north, including the Okavango, Cuando and Luangwa Rivers (Fig. 4.5) (Moore & Larkin 2001; Haddon & McCarthy 2005; Stankiewicz & de Wit 2006). Over time, as a result of the drying trend and the onset of Plio-Pleistocene glaciation, the Paleo-Kalahari Lake would have dried to leave isolated pans independent of rifting (Fig. 4.6) (Condamine *et al*. 2012). The molecular clock analysis shows a period when the Orange River and eastern South African river (Limpopo to St Lucia) populations of *Mesobola* diverged as a single population for a short period of ~4 million years (Myr).

Although it is unknown exactly when and where the Molopo River connected to the Orange River System, the involvement of the Molopo River can be proposed as a model of movement due to its present close proximity to the western Limpopo System and the southern Kalahari Basin and its current drainage into the Orange River (Fig. 4.6) (Dingle & Hendey 1984; Bootsman 1997). During the short period of combined divergence a population could have remained in the Limpopo system and made its way up the western Limpopo in particular (Fig. 4.7). With the close proximity of the current Molopo to the western Limpopo (Bootsman 1997; Goudie 2005), it is possible that a small flooding event, extension in the paleo-Molopo River or movement in one of the paleo-Molopo tributaries to connect with the western Limpopo, facilitated movement of a population into the Molopo system (Fig. 4.6). Movement may have been further facilitated by the low paleo-Kalahari beds increasing the flooding area (Fig. 4.6) (Bootsman 1997). When the Plio-Pleistocene glaciation began the Molopo would have begun to recede due to dryer conditions, severing a connection to the Limpopo population (Fig. 4.6).

Alternatively, it is possible that populations of *Mesobola* may have been trapped in the paleo-Molopo River basin in the south of the Kalahari basin (Bootsman 1997) (or perhaps the drying Paleo-Kalahari Lake) when flexural uplift along the Kalahari-Zimbabwe Axis (McCarthy & Rubidge 2005) occurred about 20 MYA. From there, the fish may have moved into the Orange and Limpopo River systems during flooding, as described above. Dingle & Hendey (1984) spoke of a model involving the changing of the Orange River outlet from a 28°S exit to a 31°S and back from the late Cretaceous to the Neogene periods. During the Paleogene, a period that coincides with divergence of the Orange River population, a combination of factors including melilitie and pseudokimberlite plugs and elevation would have prevented northern flow of the Orange River System and sea level dropping resulting in eastward downcutting of northern systems and change to 31°S outlet (Dingle & Hendey 1984). This would have resulted in downcutting and connection of the Molopo with the Orange River System (Dingle & Hendey 1984). In the following Neogene period, river capture across the Gamoep-Pofadder line would have resulted in northern flow once again and a return of the outlet of the Orange River to the 28°S line (Dingle & Hendey 1984). Therefore, this would have facilitated movement of a population into the Orange River during the Paleogene with rapid isolation in the Neogene resulting in the divergence seen within the results.

As there is no defined literature support for either possibility, it is assumed that movement from the western Limpopo River through the Molopo River and to the Orange River is more probable based on the ~4 million year period when the populations were linked to the present east South African populations (Fig. 4.1).

4.4.10 M. brevianalis in the eastern South African Rivers 8-0 MYA

The *Mesobola* populations in the river systems on the eastern side of South Africa from the Limpopo River to St Lucia all showed a close genetic relationship, with limited divergence in the last ~ 6 million years. This may be evidence of a Limpopo River population colonising southwards within the last five million years (Fig. 4.7) or may be due to the repeated connection of these systems by events such as floods.

This sudden divergence may also be explained by significant erosion events that occurred along the margins of the highveld plateau in South Africa (Fig. 4.7) (Dirks & Berger 2013). The high erosion that occurred is evidence that the present-day high relief may be relatively young (Dirks & Berger 2013), possibly less than four million years old (Partridge 2010), coinciding with the sudden divergence seen within eastern South African *M. brevianalis* (Fig. 4.7). The erosion would account for the facilitation of movement of populations along the eastern systems of South Africa by creating larger flood plain regions, with the creation of the southern African ‘high veld’ (Fig. 4.7) (Dirks & Berger 2013).

In addition, this only recent divergence, may be a result of continual gene flow due to flooding along the eastern coast of South Africa which, although there is no literature particularly within the Mozambique region to support this, SAIAB field collectors have noted that it occurs regularly. The World Meteorological Organisation (2012) has also reported on flood events in the Limpopo River that inundated land over 20 km wide. Such floods could progressively facilitate gene flow along the eastern coast of South Africa, from behind the coastal sand dunes all the way to the St Lucia System limiting divergence.

Furthermore, the glacial changes that occurred during the Pleistocene resulted in a drop in the sea level of up to 120 m (Compton 2011) and the expansion of habitat around southern Africa (Været *et al*. 2012). During ice ages, the sea levels around southern Africa could drop by as much as 120-130 m (Cooper & Pilkey 2002; Compton 2011). Maputo Bay is estimated to be 8 to 20 m deep but usually averages a shallow 10 m (Nhapulo 2000). Glacial periods as postulated by Cooper & Pilkey (2002), would cause the sea level to drop, revealing quite large areas of the Maputo Bay land surface and leaving landlocked water systems, with a lowering of only 10m. With the Limpopo, Nkomati and Pongolo systems presently draining through this bay (Fig. 4.7), a glaciation period resulting in a landlocked non-saline water system may have enabled all of its rivers to become confluent at such times. This could have facilitated gene flow among these systems over the period of two to three million years prior to when the systems showed levels of divergence, coinciding with glacial and interglacial periods beginning in the Quaternary ~2.6 MYA (Helmens 2004; Gibbard *et al*. 2010).

4.4.11 Mesobola whitei in the western Limpopo River

The population of *Mesobola* in the western Limpopo appeared to be slightly more distinct with an earlier divergence, beginning about two-to-three million years ago. Events may have resulted in the divergence of this population, but it is thought that speciation may still be in progress, as mentioned in Chapter 2. Severe drought periods in the western Limpopo (Reason & Hachigonta 2005) may have reduced the Limpopo River to disconnected pools (Engelbrecht *et al*. 1997) and randomly decimated its fish populations, allowing for genetic drift to accumulate. Evidence of *Mesobola brevianalis's* ability to survive in small pools during periods of drought would support the survival of a population (Minshull 2008) leading to low heterozygosity and random genetic differentiation from downstream populations (Engelbrecht *et al*. 1997).

4.5 Conclusion

As in *Hydrocynus* (Goodier *et al*. 2011), it appears that *M. brevianalis* and associated species have been affected by paleobiogeomorphic events since their origins in the Late Cretaceous. This study enabled the placement of dates on certain paleobiogeomorphic events not yet explored in detail in the literature, including the connection of the Kunene to either an Okavango tributary or the Etosha Pan and the Kalahari-Okavango-Zambezi uplift that occurred between the Paleogene and the Oligocene. In addition it has shown support for a number of proposed events such as the easterly flow of the Paleo-Congo River. However, due to a lack of supporting literature and research, many of the paleobiogeographical events cannot be fully supported; therefore the models of biogeographical evolution presented in this chapter are working hypotheses of the paleobiogeographical movement of the ancestral populations giving rise to the present *Mesobola* species.

The work done on *Mesobola brevianalis* and its sister taxa in this study show, once again, how the genetic data from freshwater fish helps to decipher continental and drainage system histories.

4.6 References

BAILEY, R.G., MATTHES, H. 1971. A new species of *Engraulicypris* (Cyprinidae) from Tanzania, East Africa. *Revue de Zoologie et de Botanique Africaines*, **83(1-2)**: 79-83.

BALON, E.K. 1974. Fishes from the edge of Victoria Falls, Africa: Demise of a physical barrier for downstream invasions. *Copeia*, **3**: 643-660.

BOOTSMAN, C.S. 1997. On the evolution of the Upper-Molopo Drainage. *South African Geographical Journal*, **79(2)**: 83-92.

BOULENGER, G.A. 1905. The list of the Freshwater Fishes of Africa. *The Annals and Magazine of Natural History*, **7(16):** 36-60.

COMPTON, J.S. 2011. Pleistocene sea-level fluctuations and human evolution on the southern coastal plain of South Africa. *Quaternary Science Reviews*, **30**: 506-27.

CONDAMINE, F.L., SPERLING, F.A.H., WAHLBERG, N., RASPLUS, J., KERGOAT, G.J. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecology Letters*, **15(3)**: 267-277.

COOKE, H.J. 1980. Landform evolution in the context of climatic change and neo-tectonism in the Middle Kalahari of north-central Botswana. *Transactions of the Institute of British Geographers*, **5(1)**: 80-99.

COOPER, J.A. G., PILKEY, O. H. 2002. The Barrier Islands of Southern Mozambique. *Journal of Coastal Research*, **36,** 164-172.

CRAW, D., BURRIDGE, C., NORRIS, R., WATERS, J. 2008. Genetic ages for quaternary topographic evolution: A new dating tool. *Geology*, **36**: 19-22.

CRISP, M.D., TREWICK, S.A., COOK, L.G. 2011. Hypothesis testing in biogeography. *Trends in Ecology and Evolution*, **26(2)**: 66-72.

DANLEY, P.D., HUSEMANN, M., DING, B., DIPIETRO, L.M., BEVERLY, E.J., PEPPE, D.J. 2012. The impact of the geologic history and paleoclimate on the diversification of East African Cichlids. *International Journal of Evolutionary Biology,* **2012(Article ID 574851)**: 1-20.

DAY, J.J., PEART, C.R., BROWN, K.J., FRIEL, J.P., BILLS, R., MORITZ, T. 2013. Continental diversification of an African Catfish Radiation (Mochokidae: *Synodontis*). *Systematic Biology*, **62**: 351-365.

DE VRIES, J.J., SELAOLO, E.T., BEEKMAN, H.E. 2000. Groundwater recharge in the Kalahari, with reference to paleo-hydrologic conditions. *Journal of Hydrology*, **238**: 110-123.

DELVAUX, D. 1995. Age of Lake Malawi (Nyasa) and water level fluctuations. *Royal Museum for Central Africa Tervuren (Belg.) Dept Geol Min Rapp Ann,* **1995-1996**: 99-108.

DINGLE, R.V., HENDEY, Q.B. 1984. Late Mesozoic and Tertiary sediment supply to the Eastern Cape Basin (SE Atlantic) and palaeo-drainage systems in southernwestern Africa. *Marine Geology*, **56**: 13-26.

DIRKS, P.H.G.M., BERGER, L.R. 2013. Hominin-bearing caves and landscape dynamics in the Cradle of Humankind, South Africa. *Journal of African Earth Sciences*, **78**: 109-131.

DRUMMOND, A.J., RAMBAUT, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**: 214.

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.

ESCHMEYER, W. N. (ed). *Catalog of Fishes*. California Academy of Sciences (http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 2014.

ENGELBRECHT, G.D., VAN DER BANK, F.H., MULDER, P.F.S. 1997. Allozyme variation in Shilbe intermedius Rüppel, 1832 (Pisces: Siluriformes) from two localities in the Limpopo River System, South Africa. *Comp. Biochem. Physiol*, **117B(2)**: 179-184.

FERREIRA, M.A.R., SUCHARD, M.A. 2008. Bayesian analysis of elapsed time in continuous-time Markov chains. *Canadian Journal of Statistics*, **36:** 355-368. doi: 10.1002/cjs.5550360302.

FROESE, R., PAULY, D. (eds). *FishBase*. World Wide Web electronic publication. (www.fishbase.org). Electronic version accessed 2014.

GAUBERT, P., DENYS, G., OBERDORFF, T. 2009. Genus-level supertree of Cyprinidae (Actinopterygii: Cypriniformes,), partitioned qualitative clade support and test of macro-evolutionary scenarios. *Biological Reviews*, **84**: 653-689.

GIBBARD, P.L., HEAD, M. J., WALKER, M.J.C. 2010. Formal ratification of the Quarternary System/ period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quarternary Science*, **25(2)**: 96-102.

GIRESSE, P. 2005. Mesozoic-Cenozoic history of the Congo Basin. *Journal of African Earth Sciences*, **43**: 301-315.

GOODIER, S.A.M., COTTERILL, F.P.D., O'RYAN, C., SKELTON, P.H., DE WIT, M.D. 2011. Cryptic diversity of African tigerfish (Genus *Hydrocynus*) reveals paleogeographic signatures of linked Neogene geotectonic Events. *PLoS ONE* **6(12)**: e28775. doi: 10.1371/journal.pone.0028775.

GOUDIE, A.S. 2005. The Drainage of Africa since the Cretaceous. *Geomorphology*, **67**: 437-456.

GREENWOOD, D.R., WING, S.L. 1995. Eocene continental and latitudinal temperature gradients. *Geology*, **23**: 1044-1048.

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

HELMENS, K.F. 2004. The Quaternary glacial record of the Colombo Andes. *Developments in Quartenary Sciences*, **2**: 115-134.

HOWES, G.J. 1984. A review of the anatomy, taxonomy, phylogeny and biogeography of the African neoboline cyprinid fishes. *Bulletin of the British Museum of Natural History*, **47(3)**: 151-185.

JOHNSON, T.C., KELTS, K., ODADA, E. 2000. The Holocene History of Lake Victoria. *Ambio*, **29(1)**: 2-11.

JOHNSON, T.C., SCHOLZ, C.A., TALBOT, M.R., KELTS, K., RICKETTS, R.D., NGOBI, G., BEUNING, K., SSEMMANDA, I., MCGILL, J.W. 1996. Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science*, **273(5278)**: 1091-1093.

KADYE, W.T. 2008. The application of a Fish Assemblage Integrity Index (FAII) in a Southern African river system. *Water SA*, **34(1)**: 25-32.

KRINNER, G., LEZINE, A.M., BRACONNET, P., SEPULCHRE, P., RAMSTEIN, G., GRENIER, C., GOUTTEVIN, I. 2012. A reassessment of lake and wetland feedbacks on the North African Holocene Climate. *Geophysical Research Letters*, **39(7)**: L07701 doi:10.1029/2012GL050992.

LANCASTER, N. 1981. Formation of the Holocene Lake Chilwa sand bar: Southern Malawi. *Catena*, **8(1)**:369-382.

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.

LIAO, T., KULLANDER, S.O., FANG, F. 2011. Phylogenetic position of rasborin cyprinids and monophyly of major lineages among the Danioninae, based on morphological characters (Cypriniformes: Cyprinidae). *Journal of Zoological Systematics and Evolutionary Research,* **49(3)**: 224-232.

LIEBERMAN, B.S. 2000. What is paleobiogeography? *Paleobiogeography: Topics in Geobiology*, **16**: 1-3.

LIEBERMAN, B.S. 2005. Geobiology and paleobiogeography: tracking the coevolution of the Earth and its biota. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **219**: 23-33.

MALIK, J.N., MOHANTY, C. 2007. Active tectonic influence on the evolution of drainage and landscape: Geomorphic signatures from frontal and hinterland areas along the Northwestern Himalaya, India. *Journal of Asian Earth Sciences*, **29(5-6)**: 604-618.

MCCARTHY, T. 1983. Evidence for the former existence of a major, southerly flowing river in Griqualand West. *Transactions of the Geological Society of South Africa*, **86**: 37-49.

MCCARTHY, T., RUBIDGE, B. 2005. *The Story of Earth & Life: A southern African perspective on a 4.6-billion-year journey*. Struik Publishers. Cape Town. South Africa.

MCCARTHY, T.S. 2013. The Okavango Delta and its place in the geomorphological evolution of southern Africa. *South African Journal of Geology*, **116(1)**: 1-54.

MINSHULL, J.L. 2008. Dry season fish survival in isolated pools and within sand-beds in the Mzingwane River, Zimbabwe. *African Journal of Aquatic Science*, **33(1)**: 95-98.

MOORE, A., BLENKINSOP, T., COTTERILL, F.P.D. 2009. Southern African topography and erosion history: plumes of platetectonics? *Terra Nova,* **21(4)**: 310-315.

MOORE, A.E., COTTERILL, F.P.D., BRODERICK, T., PLOWES, D. 2009. Landscape evolution in Zimbabwe from the Permian to present, with implications for kimberlite prospecting. *South African Journal of Geology*, **112**: 65-88.

MOORE, A.E., COTTERILL, F.P.D., ECKARDT, F.D. 2012. The evolution and ages of Makgadikgadi paleo-lakes: Consilient evidence from Kalahari drainage evolution. *South African Journal of Geology*, **115(3)**: 385-413.

MOORE, A.E., LARKIN, P.A. 2001. Drainage Evolution in south-central Africa since the breakup of Gondwana. *South African Journal of Geology*, **101**: 47-68.

NEAR, T.J., EYTAN, R.I., DORNBURG, A., KUHN, K.L., MOORE, J.A., DAVIS, M.P., WAINWRIGHT, P.C., FRIEDMAN, M., SMITH, W.L. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences*, **109(34)**: 13698-13703.

NHAPULO, C.I. 2000. *Correntes de marése circulação geral na b'ia de Maputo*. Licenciatura thesis. Eduardo Mondlane Univ. Maputo. 42 pp. http://hdl.handle.net/10857/353.

NICHOLSON, S.E. 1999. Historical and modern fluctuations of lakes Tanganyika and Rukwa and their relationship to rainfall variability. *Climatic Change*, **41**: 53-71.

NYLANDER, J.A.A. 2004. *MrModelTest 2.0.* Program distributed by the author. Evolutionary Biology Centre, Uppsala University. (http://www.ebc.uu.se/systzoo/staff/nylander.html). Electronic version accessed 2013.

OTERO, O., PINTON, A., MACKAYE, H.T., LIKIUS, A., VIGNAUD, P., BRUNET, M. 2009. Fishes and palaeogeography of the African drainage basins: Relationships between Chad and neighbouring basins throughout the Mio-Pliocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **274**: 134-139.

PARTRIDGE, T.C. 2008. Of diamonds, dinosaurs and diastrophism: 150 million years of landscape evolution in southern Africa. *South African Journal of Geology*, **101(3)**: 167-184.

PARTRIDGE, T.C. 2010. *Tectonics and geomorphology of Africa during the Phanerozoic.* In: Werdelin, L., Sanders, W.J. Cenozoic mammals of Africa. University of California Press. California. United States of America. ISBN: 87-0520257214. pp. 3-17.

PLANA, V. 2004. Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society of Biological Science*, **359**: 1586-1594.

POLL, M. 1945. Descriptions de cinq espèces nouvelles de Cyprinidae du Congo belge appurtenant aux genres Barbus et *Engraulicypris*. *Revue de Zoologie et de Botanique Africaines*, **38 (3-4)**: 298-311.

RAMBAUT, A. 2009. *FigTree v1.3.1*. Institute of Evolutionary Biology, University of Edinburgh. Edinburgh.

RAMBAUT, A., DRUMMOND, A.J. 2006. *TreeAnnotator 1.4.5*, Available at: http://beast.bio.ed.ac.uk/Treeannotator.

RAMBAUT, A., DRUMMOND, A.J. 2007. *Tracer v1.4*, Available at: http://beast.bio.ed.ac.uk/Tracer.

REASON, C.J.C., HACHIGONTA, S. 2005. Interannual variability in rainy season characteristics over the Limpopo region of southern Africa. *International Journal of Climatology*, **25**: 1835-1853.

ROBERTS, E.M., STEVENS, N.J., O'CONNOR, P.M., DIRKS, P.H.G.M., GOTTFRIED, M.D., CLYDE, W.C., ARMSTRONG, R.A., KEMP, A.I.S., HEMMING, S. 2012. Initiation of the western branch of the East African Rift coeval with the eastern branch. *Nature Geoscience*, **5**: 289-294.

ROBERTS, T.R. 1975. Geographical distribution of African freshwater fishes. *Zoological Journal of the Linnaean Society*, **57(4)**: 249-319.

SCHWARZER, J., SWARTZ, E.R., VREVEN, E., SNOEKS, J., COTTERILL, F.P.D., MISHOF, B., SCHLIEWEN, U.K. 2012. Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution. *Proceedings of the Royal Society of Biological Science*, **279**: 4389-4398.

SKELTON, P.H. 1994. Diversity and distribution of freshwater fishes in East and Southern Africa. *Ann. Mus. R. Afr. Centr., Zool.*, **275**: 95-131.

STANKIEWICZ, J., DE WIT, M.D. 2006. Proposed drainage evolution model for Central Africa-Did the Congo flow east? *Journal of African Earth Sciences*, **44**: 75-84.

STEWART, K.M. 2001. The freshwater fish of Neogene Africa (Miocene-Pleistocene): systematics and biogeography. *Fish and Fisheries*, **2**: 177-230.

TANG, K.L., AGNEW, M.K., CHEN, W.J., HIRT, M.V., RALEY, M.E., SADO, T., SCHNEIDER, L.M., YANG, L., BART, H.L., HE, S., LIU, H., MIYA, M., SAITOH, K., SIMONS, A.M., WOOD, R.M., MAYDEN, R.L. 2011. Phylogeny of the gudgeons (Teleostei: Cyprinidae: Gobioninae). *Molecular Phylogenetics and Evolution*, **61**: 103-124.

TANG, K.L., AGNEW, M.K., HIRT, M.V., SADO, T., SCHNEIDER, L.M., FREYHOF, J., SULAIMAN, Z., SWARTZ, E., VIDTHAYANON, C., MIYA, M., SAITOH, K., SIMONS, A.M., WOOD, R.M., MAYDEN, R.L. 2010. Systematics of the subfamily

Danioninae (Teleostei: Cypriniformes: Cyprinidae). *Molecular Phylogenetics and Evolution*, **57**: 189-214.

THOMAS, D.S.G., SHAW, P.A. 1988. Late Cainozoic evolution in the Zambezi Basin: Geomorphological evidence from the Kalahari rim. *Journal of African Earth Sciences*, **7(4)**: 611-618.

THOMAS, D.S.G., SHAW, P.A. 1991. *The Kalahari Environment. Cambridge University Press*. Cambridge. United Kingdom. pp 25-26.

THOMAS, D.S.G., SHAW, P.A. 1993. The evolution and characteristics of the Kalahari, southern Africa. *Journal of Arid Environment*, **25**: 97-108.

TSIGENOPOULOS, C.S., KASAPIDIS, P., BERREBI, P. 2010. Phylogenetic relationships of hexaploid large-sized barbs (genus *Labeobarbus*, Cyprinidae) based on mtDNA data. *Molecular Phylogenetics and Evolution*, **56(2)**: 851-856.

TWEDDLE, D., LEWIS, D.S.C., WILLOUGHBY, N.G. 1979. The nature of the barrier separating the Lake Malawi and Zambezi fish faunas. *Ichthyology Bulletin of Rhodes University*, **39**: 1-9.

VÆRET, L., LEIJNSE, A., CUAMBA, F., HALDORSEN, S. 2012. Holocene dynamics of the salt-fresh groundwater interface under a sand island, Inhaca, Mozambique. *Quaternary International*, **257**: 74-82.

VERHEYEN, E., SALZBURGER, W., SNOEKS, J., MEYER, A. 2003. Origin of the superflock of Cichlid fishes from Lake Victoria, East Africa. *Science*, **300 (5617)**: 325-329.

WANG, M., YANG, J.X., CHEN, X.Y. 2013. Molecular Phylogeny and Biogeography of *Percocypris* (Cyprinidae, Teleostei). *PLoS ONE*, **8(6)**: p1.

WELLINGTON, J. 1955. *Southern Africa: A geographical study. Physical Geography: Volume 1*. Cambridge University Press. Cambridge. United Kingdom.

WEYL, O.L.F., RIBBINK, A.J., TWEDDLE, D. 2010. Lake Malawi: fishes, fisheries, biodiversity, health and habitat. *Aquatic Ecosystem Health & Management*, **13(3)**: 241-254.

WORLD METEOROLOGICAL ORGANIZATION. 2012. Limpopo River Basin: A Proposal to Improve Flood Forecasting and Early Warning System. http://www.wmo.int/pages/prog/hwrp/chy/chy14/documents/ms/Limpopo_Report.pdf.

Table 4.1: Taxon set species and fossil calibrations showing sister groups and log (mean), log (stdev) and offset (Wang *et al*. 2013).

Fossil species	Sister groups	Log (mean)	Log (stdev)	offset
Barbus barbus **(18 MYA)**	Node subtending *Luciobarbus capito; Capoeta capoeta*	0.35	1.0	18.0
Mylopharyngodon piceus **(12.5 MYA)**	*Ctenopharyngodon idella*	0.50	1.0	12.5
Myxocyprinus asiaticus **(37.2 MYA)**	Node subtending *Ictiobus bubalus; Carpoides carpio; Cycleptus elongatus; Catostomus commersonii; Minytrema melanops; Moxostoma poecilurum; Hypentelium nigricans*	1.1	1.0	37.2

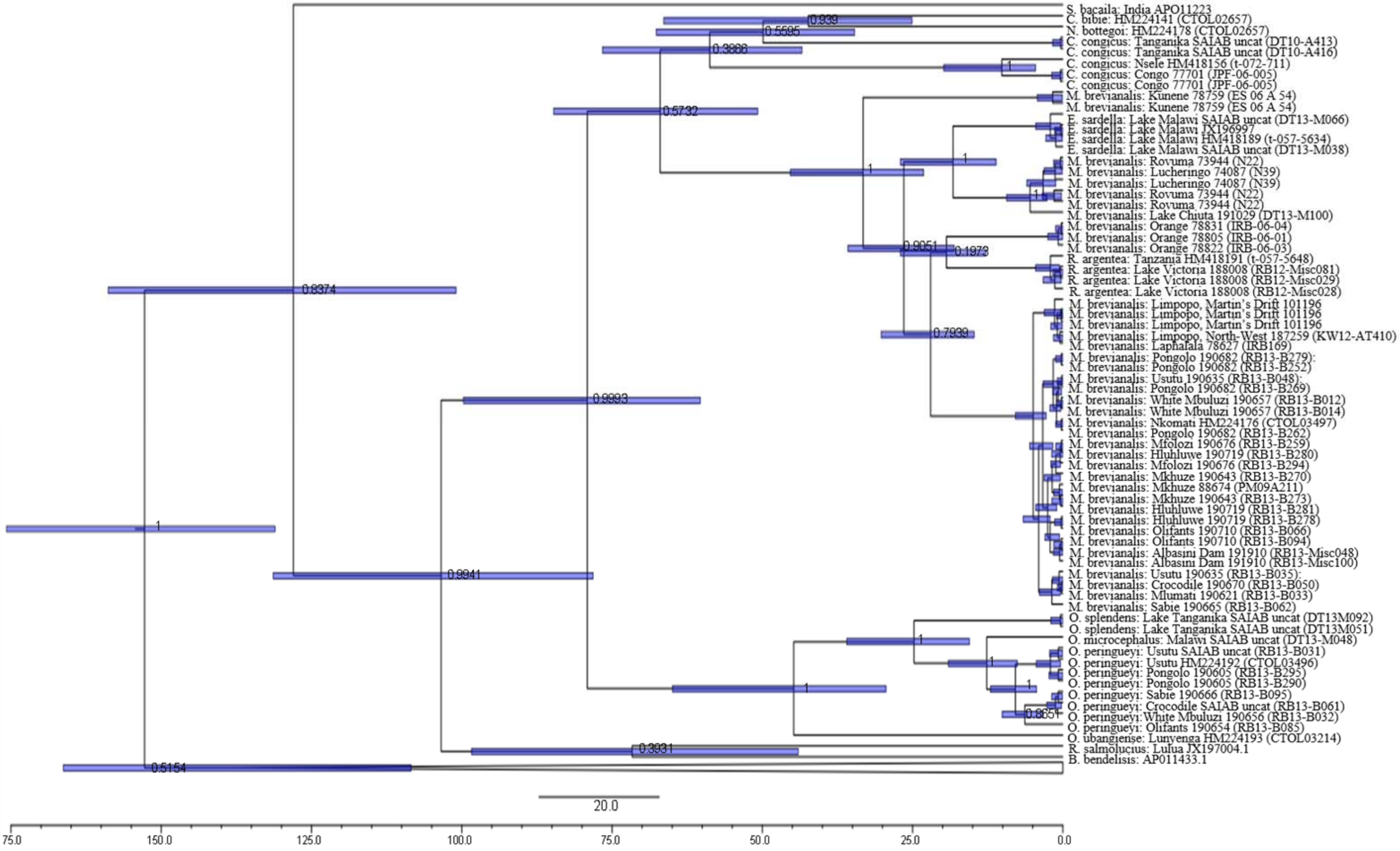


Fig. 4.1. Molecular clock analysis of the COI gene calibrated with three fossils in the outgroup clade using an uncorrelated lognormal clock, Yule prior and single portion with a TN93 model. Numbers after species name represent GenBank accession and SAIAB catalogue numbers.

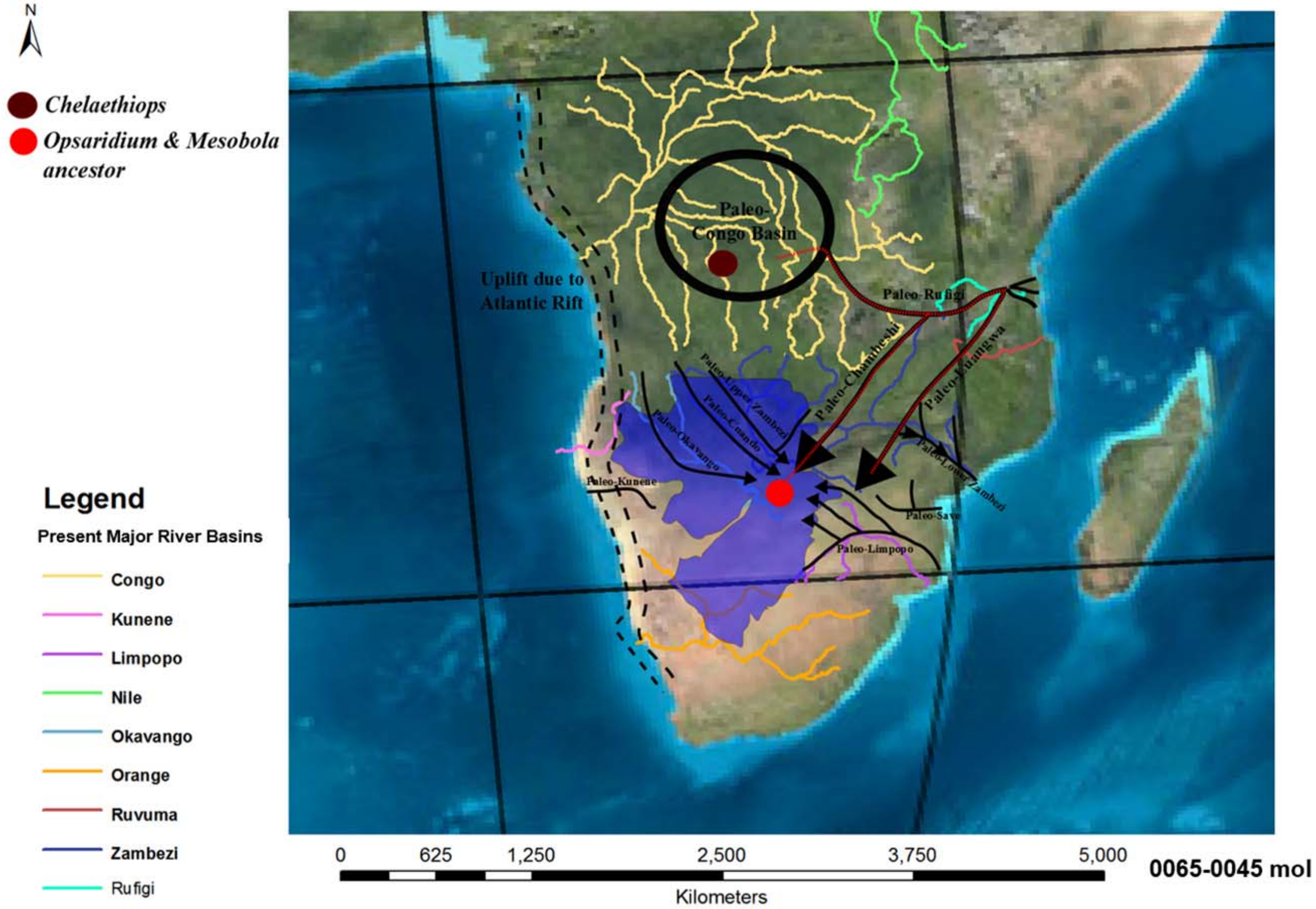


Fig. 4.2. Period from 100-65 MYA showing Congo Basin eastern drainage system facilitating divergence of *Chelaethiops*, *Opsaridium* and *Mesobola* ancestor. Map adapted from: Blakey, R. 2014. Global Paleogeography Mollewide Plate Tectonic Maps: http://www2.nau.edu/rcb7/globaltext2.html. Model adapted from Stankiewicz and de Wit (2006).

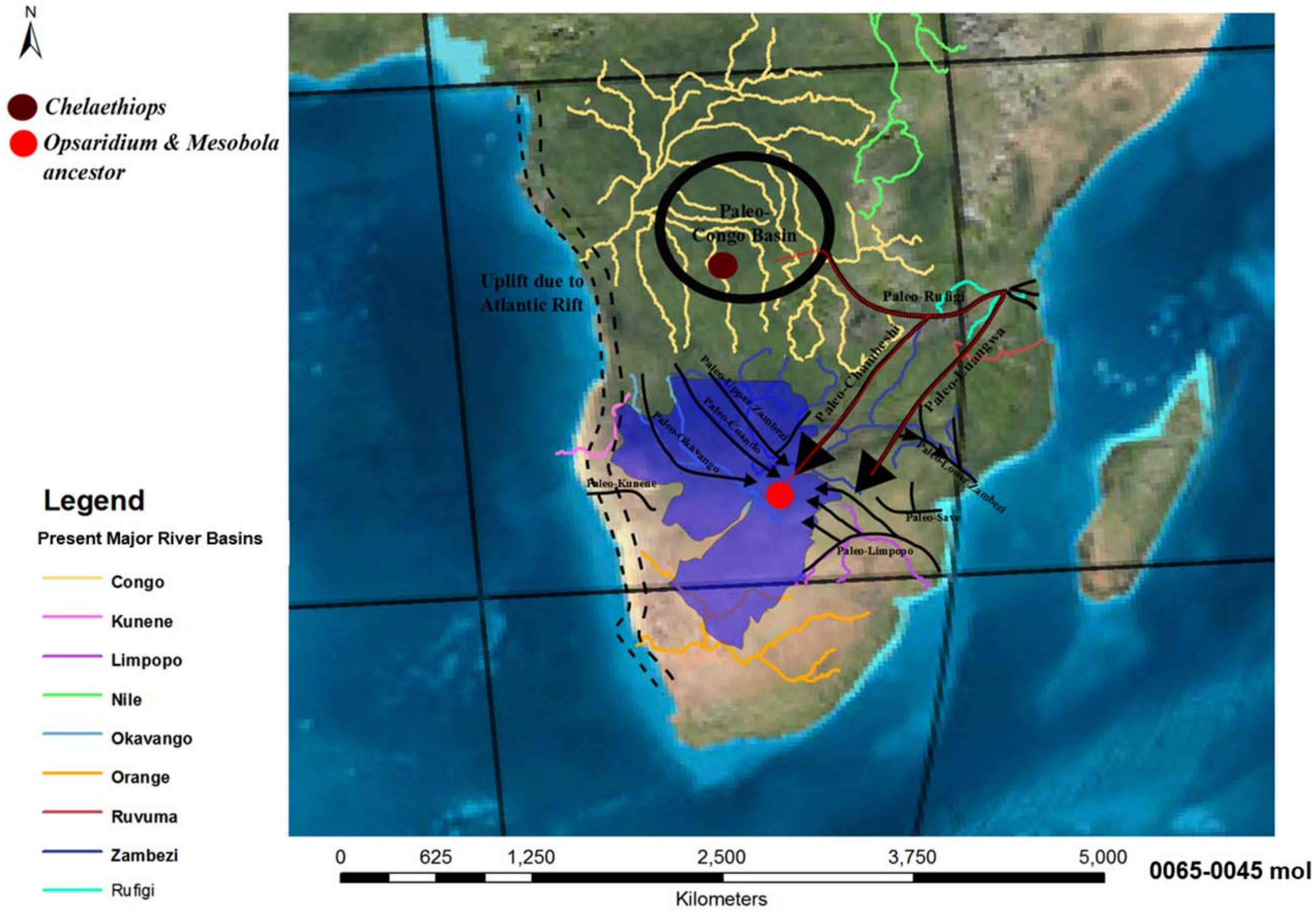


Fig. 4.3. Period from 65-45 MYA showing Paleo-Kalahari Lake and associated rivers. Map adapted from: Blakey, R. 2014. Global Paleogeography Mollewide Plate Tectonic Maps: http://www2.nau.edu/rcb7/globaltext2.html. Model adapted from Stankiewicz and de Wit (2006) and Paleo-Kalahari Lake area adapted from Bootsman (1997).

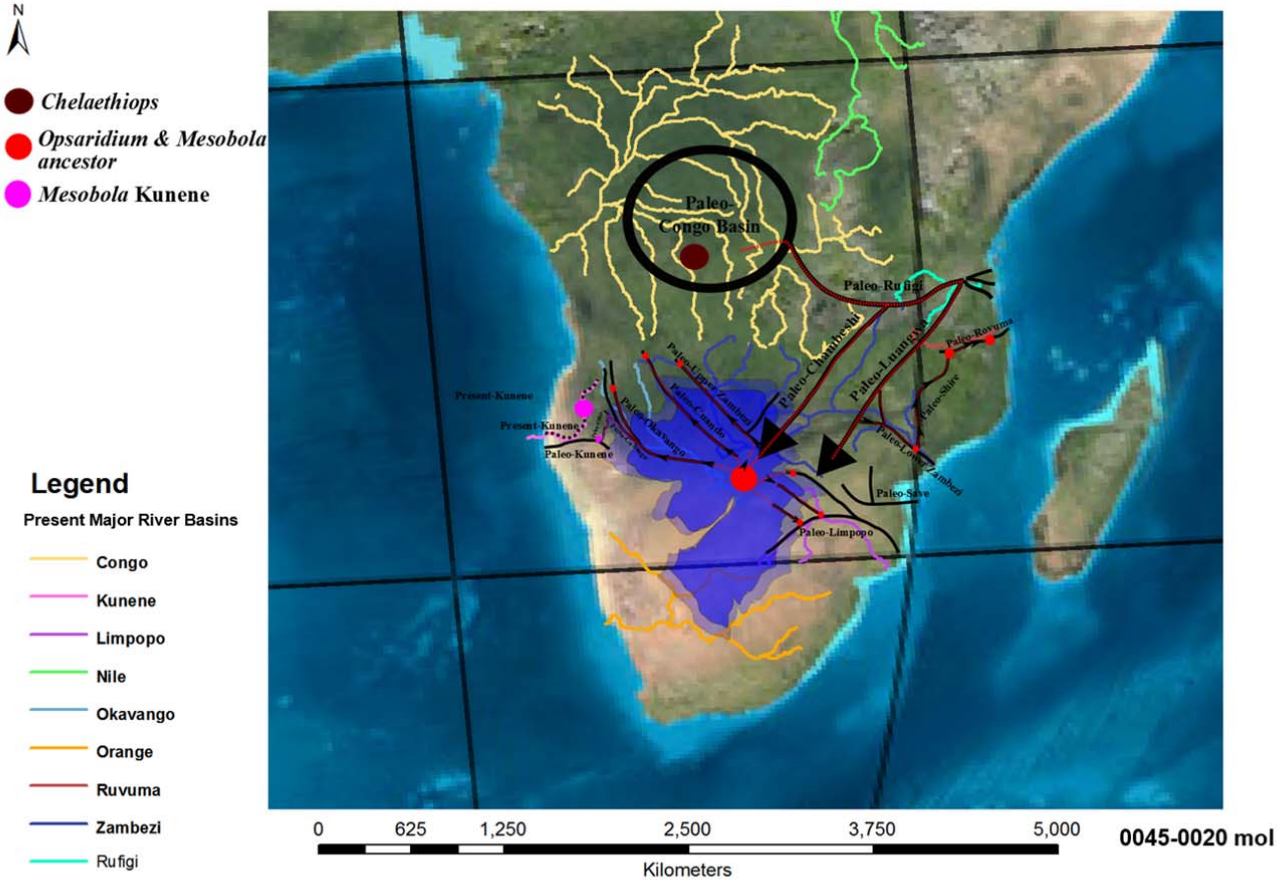


Fig. 4.4. Period from 45-20 MYA showing Oligocene to early Miocene heating and isolation of a *Mesobola* (Kunene) ancestor. Map adapted from: Blakey, R. 2014. Global Paleogeography Mollewide Plate Tectonic Maps: http://www2.nau.edu/rcb7/globaltext2.html. Model adapted from Stankiewicz and de Wit (2006).

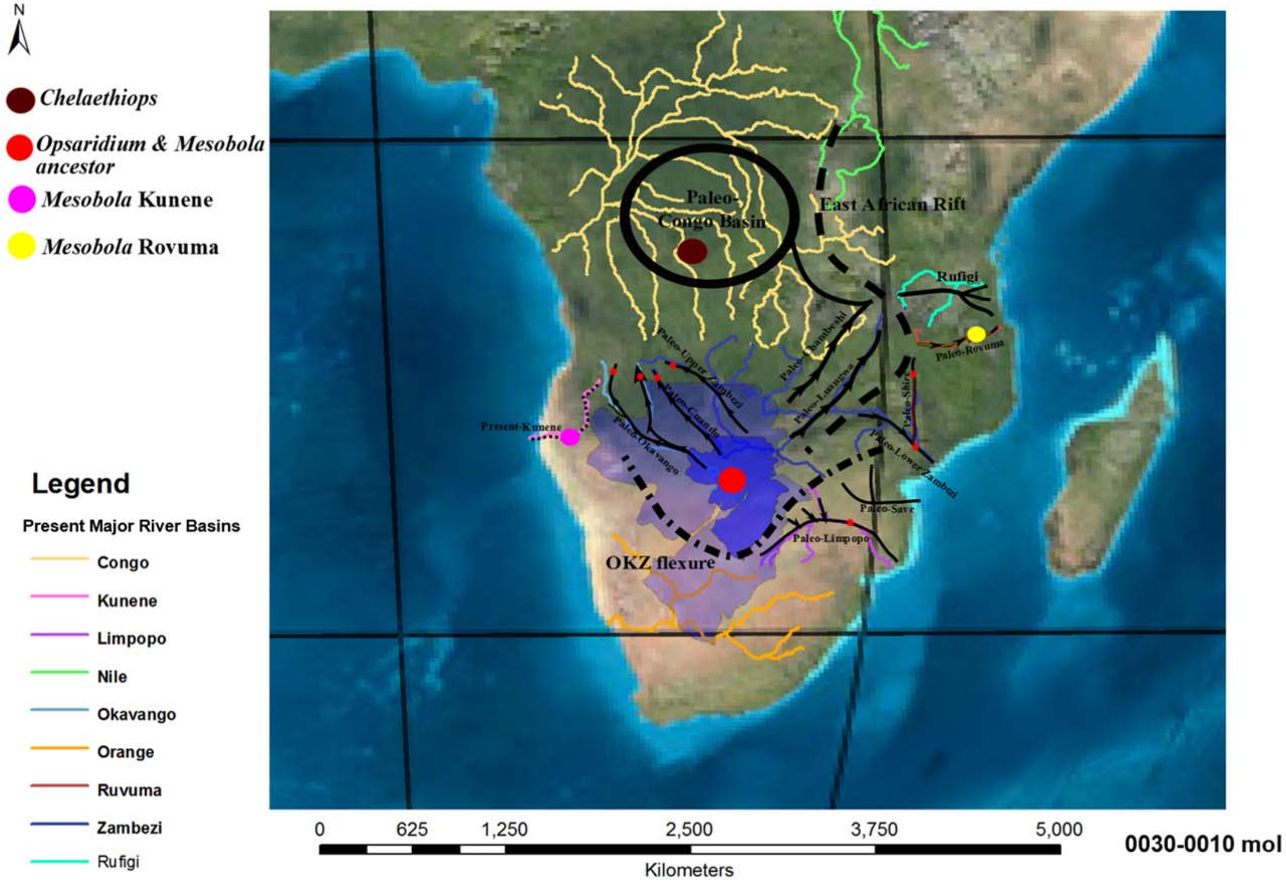


Fig. 4.5. Period from 30-10 MYA showing OKZ flexure facilitating isolation of *M. brevianalis* (Rovuma) and Limpopo *M. brevianalis* and *Opsaridium* ancestor. Map adapted from: Blakey, R. 2014. Global Paleogeography Mollewide Plate Tectonic Maps: http://www2.nau.edu/rcb7/globaltext2.html. Drainage model adapted from Stankiewicz and de Wit (2006).

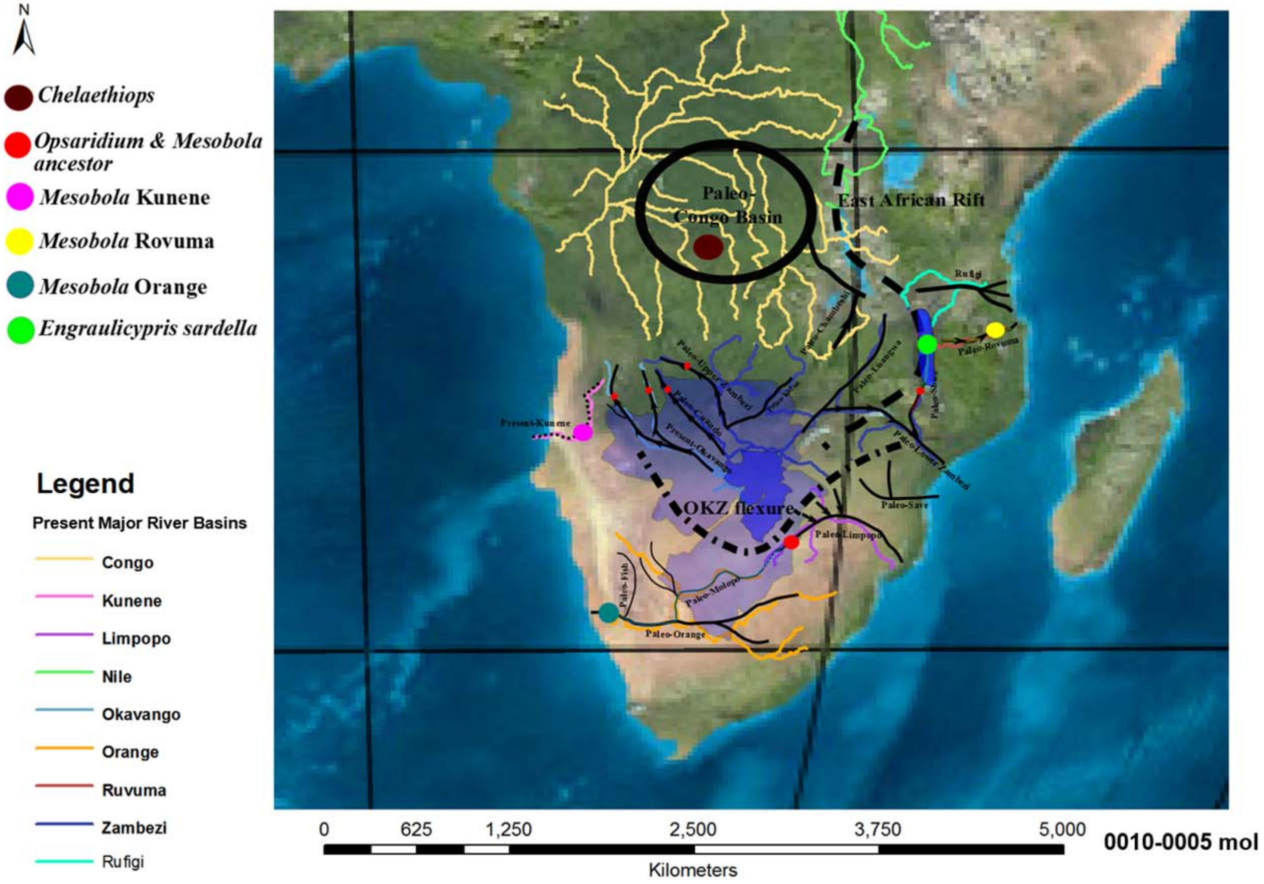


Fig. 4.6. Period from 10-5 MYA. Nyasa eastern drainage system facilitating divergence of *Mesobola* ancestor. Map adapted from Blakey, R. 2014. Global Paleogeography Mollewide Plate Tectonic Maps: http://www2.nau.edu/rcb7/globaltext2.html. Drainage model adapted from Stankiewicz and de Wit (2006).

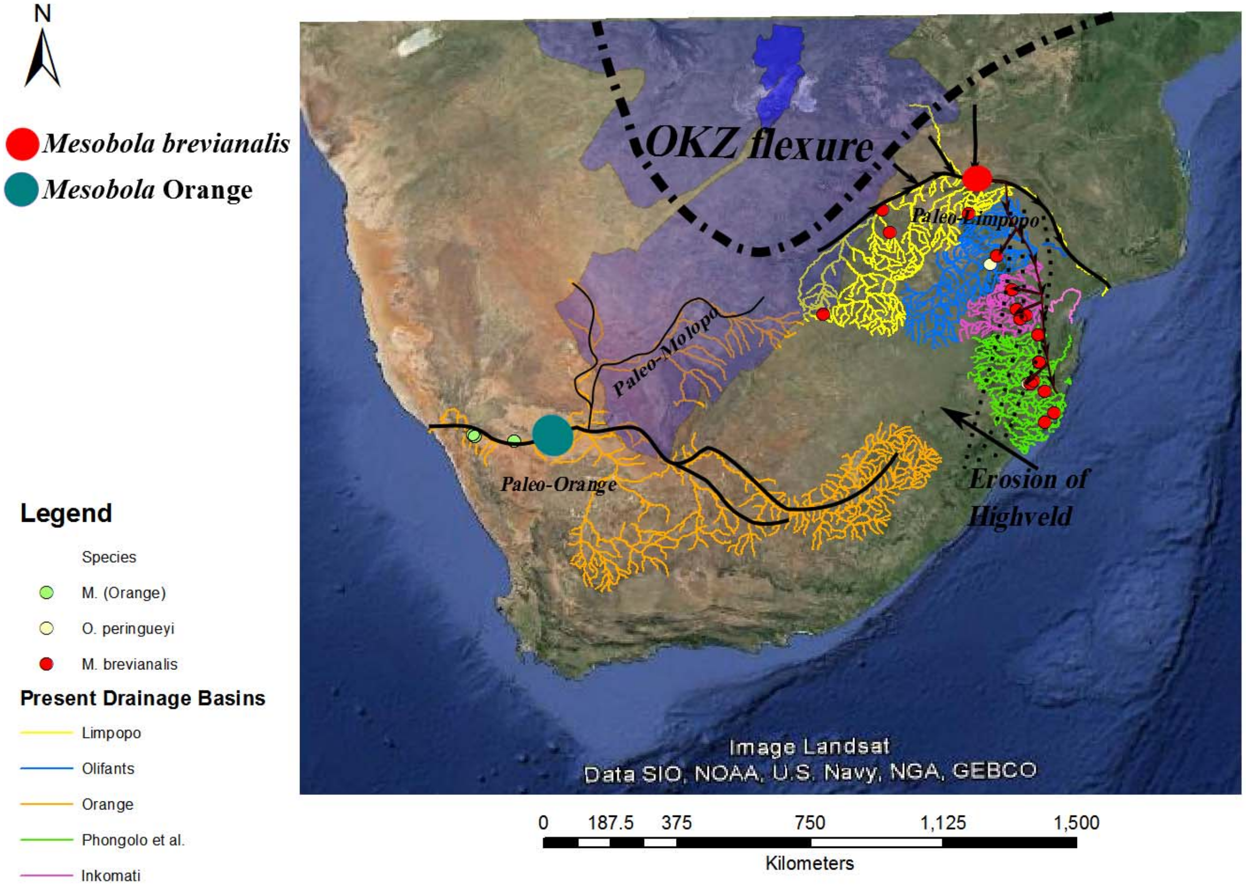


Fig. 4.7. Period from 5 MYA-present in South Africa showing eastern Highveld erosion facilitating divergence of *M. brevianalis* and *O. peringueyi* spread down to the St Lucia System. Map adapted from Google Earth, 2014.

5

Rediagnosis of the genus *Engraulicypris* and description of two new species, *Engraulicypris ngalala* and *Engraulicypris howesi*.

5.1 Introduction

The populations currently treated as *M. brevianalis* and found in the Kunene system, the Rovuma system and the Orange System each exhibited sufficient morphological and genetic differentiation and isolated divergence to be considered as new species in their own right (Chapters 2 and 3). It is essential to describe species that exhibit enough evidence of being unique and independent because without correct identification of species we would not be able to develop effective conservation strategies (Mace 2004). The river sardine exhibits sufficient ecological and economic importance to warrant such prompt attention.

The Orange River system population was previously described by Barnard (1943) based on meristics under the name *Engraulicypris gariepinus* from the lower Orange River and the Fish River, which is a tributary of the Orange River. This decision was made based on branched anal rays and lateral line scale counts (Barnard 1943). This was supported later by Farquharson (1962). The species was then recognized as a junior synonym of ***Engraulicypris brevianalis*** for unstated reasons by Jubb (1963) and supported by Jubb & Farquharson (1965), Jubb (1965) and Fowler (1977) and supported by Lévêque and Daget (1984) as a junior synonym for *Neobola brevianalis*. It was subsequently transferred to *Mesobola brevianalis* (Howes 1984); this is its current taxonomic status (Eschmeyer 2014). Molecular evidence now supports Barnard's (1943) initial description with evidence from a multiple genetic marker approach which is proven to provide an accurate representation of taxonomic relationships and evolutionary histories (Rubinoff 2006). This species, therefore, requires a species name and a description. The Kunene and Rovuma system populations have no prior descriptions.

The *M. brevianalis* population from the western Limpopo showed a short unique divergence period however, they were seen to still be closely nested within type populations. Although they showed this differentiation, these populations are considered to be insufficiently genetically unique to be considered new species. It is suggested that the population should remain under the species name *brevianalis* (Boulenger, 1908), and that *whitei* (van der Horst 1934) described from the Aapies should remain its junior synonym.

The position of *Engraulicypris sardella* (Günther 1868) and *Rastrineobola argentea* (Pellegrin 1904) in the phylogenetic analyses makes the genus *Mesobola* paraphyletic (Chapter 2), a situation that is generally considered to be undesirable (Crisp & Chandler 1996). A solution is to synonymise the genera. This supports Boulenger's (1911) description of an *Engraulicypris* genus with *brevianalis*, *argenteus* and *sardella* with the exception of his further inclusion of *bottegoi*. The inclusion of these species within the clade was strongly supported as they were supported by both Rag1 and COI results that supply multi-marker evidence (Rubinoff 2006). As both *Engraulicypris* Günther, 1894 and *Rastrineobola* Fowler, 1936 are monotypic (Eschmeyer 2014; Froese & Pauly 2014), *Rastrineobola* can be treated as a subjective junior synonym and it is proposed that it should revert to *Engraulicypris* Günther, 1894, the first genus to be described. *Mesobola* Howes, 1984 would also become junior synonyms of *Engraulicypris*. The weaker phylogenetic and biogeographical support for the relationships of *Rastrineobola* (Chapters 2 and 4) suggests that synonymising it may be premature, but the argument for synonymising *Mesobola* and *Engraulicypris* is well supported (Chapter 2).

A morphological key to the southern African species proposed to be within *Engraulicypris* is included following the rediagnosis.

5.2 Materials and methods

Specimens and specimen information were obtained from the South African Institute for Aquatic Biodiversity, Grahamstown (SAIAB), the Albany Museum, Grahamstown (AMGT) and the British Museum of Natural History (BMNH), London. Catalogued SAIAB material was selected for description based on the condition of the specimens (e.g. fin rays and scales intact) and whether specimens had an associated genetic sequence. Photographs of type material in the Natural History Museum, London (BMNH) were supplied by the curator, James Maclaine at the BMNH.

Specimens were photographed with a scale bar. Measurements were made on each specimen with standard unbranded electronic digital calipers. Photographs of the *M. brevianalis* and *E. sardella* holotypes were received from the British Museum of Natural History as the specimens could not be sent and analysed directly. The holotype photographs were measured through the use of imaging analysis software, Analysis Docu (Soft Imaging Systems: http://www.soft-imaging.net/), however measurements that involved the width of the specimen including body width or inter-orbit length could not be measured or included in the description.

Meristic measurements, including fin ray counts, where made on each specimen using a Leica Zoom 2000 microscope. Scale counts were only made on a maximum of two specimens as the scales were very difficult to see and could only be seen by dyeing them with Alizarin Red for an average of five-to-ten minutes and then placing them directly into Alcian Blue dye for a further five-to-ten minutes. Because the dye did not wash out well, the holotype scale counts could not be made.

Vertebra counts were made on X-rays of some specimens including all holotypes except for the *E. sardella* holotype for which no X-ray was available.

A single specimen from each population was cleared and stained using standard methods (Taylor & van Dyke 1985), preserved in 70% glycerol, and dissected to count the gill rakers on both the ceratobranchial and epibranchial arches.

Photographs of recently preserved specimens were taken in the field by SAIAB staff and photographs of preserved specimens and Holotypes from each population were taken in the SAIAB lab.

The data were used to build a character database in the DELTA software package (Dallwitz 1980; Dallwitz *et al*. 1993), which was used to generate the species descriptions and key.

5.3 Results

Descriptions

ENGRAULICYPRIS Günther, 1894

Engraulicypris was described by Günther (1894) from Lake Malawi/Nyasa based on a single species *Engraulicypris pinguis* Günther, 1894. He described a genus possessing an elongated body and head, scales deciduous and of moderate size, dorsal fin short and directly over origin of anal fin with less than nine branched rays and the anal fin with ~14 rays. Günther (1894) further described it as bearing a 'compressed snout with a lateral mouth of moderate width, slender gill rakers, double series pointed pharyngeal teeth, suborbitals covering almost entire cheek and upper jaw completely hidden below large præorbital' (Günther 1894). This genera and species placement was supported by the following studies: Lèvêque and Daget (1984), Poll and Gosse (1995), Tang *et al*. (2010) (Eschmeyer 2014).

However, with the finding of *Mesobola* within the genus, the rediagnosis must include features of *Engraulicypris gariepinus*, and the new Rovuma and Kunene species to be described below. The genus is then identified by the lack of pelvic or pectoral fin lobes, mouth reaching to below anterior border of orbit, dorsal fin origin more or less above origin of anal fin, pectoral fin not reaching origin of anal fin, and body colouration lacking vertical bars or bands (Skelton *et al*. 1991). The description is summarised as follows:

***Engraulicypris* (Günther, 1894**)

Diagnosis. Mouth reaching anterior border of orbit. Dorsal fin more or less above origin of anal fin. Pectoral fin lacking lobe at base. Pelvic fin lacking a basal lobe.

Live colouration. Body colouration without vertical bars or bands.

Rastrineobola shows similar morphology but is not synonymised here due to uncertainty about its relationships within *Engraulicypris* and the lack of paleobiogeographical support (Chapters 2 and 4). The proposed species of *Engraulicypris* are described below.

Engraulicypris sardella **(Günther, 1868)**

Barilius sardella. Günther, 1868. Catalogue of the fishes in the British Museum, vol. 7: 292. Holotype, BMNH 1863.11.12.23, unsexed; "Shire River, eastern Africa"; [BMNH].

= *Engraulicypris pinguis* Günther, 1894. *Proceedings of the Zoological Society of London* 1893(4): 626, fig. (Lévêque & Daget 1984: 301). Syntypes: 2 unsexed, Lake Nyasa [= Lake Malawi], south eastern Africa [BMNH 1893.1.17.15–16].

Etymology. '*Sardella*' refers to the sardine-like form and habits of this species. '*Pinguis*' is Latin that alludes to plumpness or oilyness.

Morphology. (Fig. 5.1.A, 5.2.A & 5.3.A ; Table 5.1 & 5.6.A). Maximum SL 13 cm (Eccles, 1992). Body elongated, fusiform and cylindrical. Maximum body depth in the middle of pelvic and pectoral fin origin. Pre-dorsal profile straight or slightly convex behind head. Head length 22 % of SL, without tubercles (Tweddle, pers. com.). Snout pointed; elongated (slightly compressed); 35 % of head length. Mouth terminal; crescent-shaped; reaching anterior border of orbit. Nostrils small; level with dorsal margin of eye; separated from orbit by less than one orbit radius. Tubular anterior naris absent. Eye lateral; slightly visible above and from below; 31 % of head length. 1st Ceratobranchial gill arch with 27 gill rakers. Gill rakers very long; pointed; closely-spaced. 1st Epibranchial with 9 gill rakers. Pharyngeal teeth in four rows. Pharyngeal teeth 5,3,2,1; slender and long; falcate.

Fins moderate to small. Dorsal fin closer to caudal fin than tip of snout; more or less above origin of anal fin; length 13 % of SL; posterior margin slightly concave; rays soft; anterior-most branched fin ray longest. Dorsal and anal fin project at 45 ° angle. Pectoral fins largest; reaching 1/2 of distance to base of pelvic fins; fin lacking lobe at base. Pelvic fins reaching 1/2 distance to base of anal fin; small; rounded; fin lacking a basal lobe. Anal fin short; extending 1/3 length of caudal peduncle. Ano-genital opening at anterior of base of anal fin. Caudal peduncle long, 18 % of standard length; narrow, depth 40 % of caudal length. Caudal fin forked; lobes pointed; lower lobe shorter.

Scales small to medium relative to body size; in regular rows; cycloid; radially striate; rounded, slightly elongate. Base of anal fin lacking sheath of enlarged, elongate scales. Lateral line present; complete; dipping gradually down towards ventral at tip of pectoral fin; does not join midline; scale count 51–55.

Live colouration. Body and head metallic silver with distinct countershading. Body midline silver (broad and distinct from ventral); colouration without vertical bars or bands. Dorsal surface dark silver grey with ubiquitous melanophores. Snout silver grey. Operculum entirely shiny. Iris dark to light grey. Dorsal fin membrane clear; rays pale grey; melanophores fading towards tips. Caudal fin membrane clear; rays blue-grey; rays lightening towards edge; melanophores abundant and fading towards tips. Anal fin rays clear; membrane clear; melanophores minimum to absent. Pectoral and pelvic fin membranes clear; rays clear.

Preserved colouration. Body and head white and silver with dark spotting sparse to densely packed towards caudal fin on midline and on dorsal surface. Scales on dorsal surface lightly pigmented. Ventral scale pigmentation less intense than dorsal. Melanophores small, dark; along midline, increasing in intensity to caudal fin; intermittent on dorsum to origin of caudal fin; large and clustered from origin of anal fin to caudal fin. Operculum with silver sheen. Side of body with silver sheen extending from pectoral fin to anal fin origin. Membranes between fin rays clear. Pelvic fin clear rays. Dorsal, caudal and pectoral fin rays with melanophores small, dark, widely-spaced; rays pale grey.

Geographical distribution. MALAWI, MOZAMBIQUE, TANZANIA: Lake Malawi, upper Shire River (Eccles, 1992).

Material examined. *E. sardella*: **Holotype,** BMNH No 1863.11.12.23, SL: 56.22 mm. "Shire River, eastern Africa". [BMNH]. **Other material.** SAIAB 191026, 5 unsexed: SL 56.76–64.12 mm. Malawi, Lake Malawi, Chirombo Bay, 14°07'12"S 34°56'4"E, 9 March 2013, D. Tweddle [SAIAB].

Biology. Demersal and pelagic species and only known species in Lake Malawi to have pelagic larvae which remain in the upper 100 m of water column (~2 mm when hatched) (Thompson & Irvine 1997). Tends to form large schools (Lewis & Tweddle, 1990). Occurs over sandy substrates in the morning and moves offshore during the day (Konings, 1990). Caught at night using plunge nets, chirimilas or beach seines (lamparas) with light (Eccles, 1992). Experiences great fluctuations in populations within and between years due to food availability (Thompson & Irvine 1997). Diet experiences ontogenetic shifts and predominantly feed on zooplankton including copepodite copepods and claderocerans

(Thompson & Irvine 1997). *E. sardella* rarely live longer than a single year (Thompson & Irvine 1997).

Engraulicypris brevianalis (Boulenger, 1908)

Neobola brevianalis. Boulenger, 1908. *Annals of the Natal Government Museum* **1**(3): 281. Holotype, BMNH 1907.4.17.90, unsexed; "Mkuzi River, Zululand, Transvaal"; [BMNH].

= *Engraulicypris whitei* van der Horst, 1934. *Annals of the Transvaal Museum* **15**(3): 281, fig (Lévêque & Daget 1984: 326). Syntypes: 5 unsexed, Petronella [SAIAB (formerly RUSI) 30040 ex TMP 15024]; 4 unsexed, Hammanskraal [SAIAB 30041 ex TMP 16022].

Etymology. '*Brevianalis*' alludes to the relatively short anal fin. '*Whitei*' honours Mr A. G. White, who collected the type specimens in the Aapies River near Petronella and again near Hammanskraal.

Morphology. (Fig. 5.1.B, 5.2.B. 5.3.B; Table 5.2, 5.6.B). Maximum SL 7.5 cm (Skelton, 1993). Body elongated; somewhat fusiform; laterally compressed. Maximum body depth at middle point between pelvic and pectoral fin origin. Pre-dorsal profile straight or slightly convex behind head. Head length 20 % of SL; with tubercles along jaw and head in breeding males (Barnard 1943). Snout rounded; short; 30 % of head length. Mouth terminal; slightly crescent-shaped with long anterior side; reaching anterior border of orbit. Nostrils large; level with dorsal margin of eye; separated from orbit by less than one orbit radius. Tubular anterior naris short; adjacent to open posterior naris. Eye lateral; visible from above and below (more prominent); 35 % of head length. 1st Ceratobranchial gill arch with eight gill rakers. Gill rakers long; pointed; widely-spaced. 1st Epibranchial with three gill rakers. Pharyngeal teeth in three rows. Pharyngeal teeth 4,3,2; robust and long; falcate.

Modal fin formulae in Table 5.2. Fins large in relation to body size. Dorsal fin closer to caudal fin than tip of snout; more or less above origin of anal fin; length 17 % of SL; posterior margin straight; rays soft; anterior-most branched fin ray longest. Dorsal and anal fin point parallel. Pectoral fins largest; reaching $^1/_2$ to $^3/_4$ distance to base of pelvic fin; fin lacking lobe at base. Pelvic fins reaching $^2/_3$ distance to base of anal fin; relatively small; pointed; fin lacking a basal lobe. Anal fin moderately long; extending $^2/_3$ length of caudal peduncle; last unbranched ray longest. Ano-genital opening at anterior of base of anal fin. Caudal peduncle moderately long, 16 % of standard length; deep, depth 65 % of caudal length. Caudal fin forked; lobes with slightly concave interior and extending into point; upper lobe shorter.

Scales small to medium relative to body size; in regular rows; cycloid; radially striate; rounded, slightly elongate. Base of anal fin lacking sheath scales. Lateral line present; complete; dipping sharply down towards ventral at tip of pectoral fin; joins midline at posterior of caudal peduncle; scale count 53–57.

Live colouration. Body and head silver with pale brown dorsal surface. Body midline silver; colouration without vertical bars or bands. Dorsal surface with small dark brown melanophores. Snout darker dorsally. Operculum entirely shiny. Iris dark to light grey. Dorsal fin membrane clear; rays clear with olive melanophores; fading towards tips. Caudal fin membrane clear towards vivid yellow at fork; rays light olive; rays lightening towards edge; melanophores small, dark, fading towards rear. Anal fin rays clear; membrane clear; dark spotting above origin; melanophores dark olive fading towards tips. Pectoral fin membranes clear; rays clear; first ray with few dark melanophores. Pelvic fin rays clear; membrane clear.

Preserved colouration. Body and head white and silver with dark spotting sparse to densely packed towards caudal fin on midline and on dorsal surface. Scales on dorsal surface lightly pigmented. Ventral scale pigmentation as intense as dorsal surface. Dorsal surface of head lightly pigmented. Melanophores small, dark; clustered on rear of head, below orbit and on lips and snout; along midline, increasing in intensity to caudal fin; browner on dorsal surface above midline; forming a small, dark line above anal fin. Operculum and posterior edge of orbit with silver sheen. Membranes between fin rays clear. Pelvic fin clear membranes and rays. Dorsal, caudal and pectoral fin membranes clear; rays with melanophores small, dark, widely-spaced; rays pale grey.

Geographical distribution. SOUTH AFRICA, BOTSWANA, ZIMBABWE, MOZAMBIQUE: Mkhuze River, Limpopo River, St Lucia System, Nkomati System, Pongolo System.

Material examined. *E. brevianalis*: **Holotype**, BMNH No 1907.4.17: 90, SL 67.39 mm. "Mkuzi River, Zululand, Transvaal". [BMNH]**.**

E. whitei: **Syntypes**, SAIAB 30040 (ex TM No 15.024) (5) and SAIAB 30041 (ex 16.022) (4), "Aapies River (Limpopo system) near Petronella and near Hammanskraal (Transvaal)" [SAIAB]**.**

Other material, SAIAB 190643, 5 unsexed: SL 41.13–43.78 mm. South Africa, Mkhuze River, 27°36'24.9"S 32°03'40.6"E, 27 May 2013, R.Bills & M. Riddin [SAIAB].

Biology. Pelagic species preferring close proximity to substrate and seeking out slacker areas such as backwater, eddies and pools below rifles (Bills, pers.com). Occurs in shoals and prefers well-aerated, open water in flowing rivers (Skelton 2001), favouring the upper stratum (Engelbrecht & Mulder 1999). Feeds from water column on planktonic crustaceans and insects (e.g. midges and ants) (Skelton 2001). Caught at night with light (Bills, pers.com). Breeding occurs in early summer (Skelton 2001). Caught in dams where appears to propagate successfully with little predation and moves around in rivers according to seasonal flows (Bills, pers. com). Appears to migrate up streams in spring to breed where it is found in tributaries (Bills, pers. com).

Engraulicypris gariepinus (Barnard, 1943) stat. rev.

Engraulicypris gariepinus. Barnard, 1943. *Annals of the South African Museum* **36**(2): 220. Syntypes, SAM 18722–23, 2 unsexed (unlocatable). Neotype, SAIAB 78831, unsexed; "Orange River and Fish River"; [SAIAB].

Etymology. '*Gariepinus*' refers to the Gariep, a San name for the Orange River, that means "Great water".

Taxonomic notes. The two syntypes for *E. gariepinus* Barnard, 1943 were originally stored in the South African Museum, but were moved to the Albany Museum, Grahamstown, South Africa (106, 1009) (Eschmeyer 2014). The specimens cannot be found there now. For this reason, a neotype is nominated.

Morphology. (Fig. 5.1.C, 5.2.C. 5.3.C; Table 5.3, 5.6.C). Maximum SL 4.6 cm. Body elongated; somewhat fusiform; laterally compressed. Maximum body depth before pelvic fin. Pre-dorsal profile straight or slightly convex behind head. Head length 21 % of SL; with tubercles on jaw and head in breeding males (Barnard 1943). Snout rounded; short; 32 % of head length. Mouth terminal; slightly crescent-shaped with long anterior side; reaching anterior border of orbit. Nostrils large; level with dorsal margin of eye; separated from orbit by less than one orbit radius. Tubular anterior naris short; adjacent to open posterior naris. Eye lateral; visible from above and below (more prominent); 32 % of head length. 1st Ceratobranchial gill arch with seven gill rakers. Gill rakers long; pointed; widely-spaced. 1st Epibranchial with three gill rakers. Pharyngeal teeth in three rows. Pharyngeal teeth 4,3,2; robust and long; falcate.

Modal fin formulae in Table 5.3. Fins large in relation to body size. Dorsal fin closer to caudal fin than tip of snout; more or less above origin of anal fin; length 17 % of SL; posterior margin straight; rays soft; anterior-most branched fin ray longest. Dorsal and anal fin point parallel. Pectoral fins largest; reaching $^1/_2$ to $^3/_4$ distance to base of pelvic fin; fin lacking lobe at base. Pelvic fins reaching $^2/_3$ distance to base of anal fin; relatively small; pointed; fin lacking a basal lobe. Anal fin moderately long; extending over $^3/_4$ length of caudal peduncle; last unbranched ray longest. Ano-genital opening at anterior of base of anal fin. Caudal peduncle short, 14 % of standard length; and deep, depth 60 % of caudal length. Caudal fin forked; lobes pointed; upper lobe shorter.

Scales small to medium relative to body size; in regular rows; cycloid; radially striate; rounded, slightly elongate. Base of anal fin lacking sheath of enlarged, elongate scales. Lateral line present; complete; dipping sharply down towards ventral at tip of pectoral fin; joins midline at posterior of caudal peduncle; scale count 49–51.

Live colouration. Body and head silver with transparent pale brown dorsal surface. Body midline silver; colouration without vertical bars or bands. Dorsal surface with melanophores concentrated around dorsal fin. Snout with dense dark spotting on tip. Operculum entirely shiny. Iris dark to light grey. Dorsal fin membrane clear; rays clear; melanophores fading towards tips. Caudal fin membrane clear to pale orange towards midline; rays dark grey; rays lightening towards edge; melanophores small, dark, fading towards rear. Anal fin rays clear; membrane clear; pale orange spotting above origin; melanophores minimum to absent. Pectoral fin membranes clear; rays clear; first ray few dark melanophores. Pelvic fin rays clear; membrane clear.

Preserved colouration. Body and head orange with small dark brown spotting along dorsal surface, midline and above anal fin. Scales on dorsal surface lightly pigmented. Ventral scale pigmentation less intense than dorsal. Dorsal surface of head lightly pigmented. Melanophores small, dark; grouped on rear of head, below orbit, and on lips and snout; along midline, increasing in intensity to caudal fin; brownish on dorsal surface, darkening between origin of pectoral and dorsal fin; forming small dark line above anal fin. Membranes between fin rays clear. Pelvic fin clear membranes and rays.

Geographical distribution. SOUTH AFRICA, NAMIBIA: Lower Orange River system, Fish River (Barnard, 1943).

Material examined. **Neotype**, SAIAB 78831, unsexed: SL 42.48 mm. Namibia, Orange River, Noordower, 28°44'50"S 17°36'32"E [SAIAB]**. Other material**, SAIAB 193617, 4 unsexed: SL 43.45–43.75 mm. Collected with neotype [SAIAB]; SAIAB 78822, 5 unsexed: SL 38.60–41.28 mm. Namibia, Orange River, Felix Unite Camp, 28°41'19"S 17°33'20"E, 20 October 2006, R. Bills [SAIAB]; 2 unsexed, SAIAB 78805, 42.15–46.61 mm. Namibia, Orange River, Houms River Camp Site, 28°52'5"S 18°36'42"E, 18 October 2006, R. Bills [SAIAB]; SAIAB 74232, 8 unsexed: SL 28.76–40.78 mm. South Africa, Orange River, Pella Drift lower site, 28°57'47"S 19°6'36"E, 28 January 2004, R. Bills & N. Jones [SAIAB].

Biology. Shoaling fish favouring open, shallow water, normally in slacker pools and particularly below rifles. Populations found in lower Orange and Fish River limited by Augrabies and Fish River Falls. They are thought to feed on small invertebrates (planktonic crustaceans or insects), autocthonous species mainly. They are caught in large populations and therefore abundant where they occur. Appear to favour turbid waters due to small shoaling fish with no defence (Bills, pers. com).

Engraulicypris howesi **Riddin, Bills & Villet n. sp.**

Holotype, SAIAB 39012, unsexed; [SAIAB].

Etymology. This species is named in honour of Gordon Jon Howes (1938-2013), whose studies of the anatomy of the Danioninae laid the foundations of their modern classification (Howes 1980, 1984).

Morphology. (Fig. 5.2.D. 5.3.D; Table 5.4, 5.6.D). Maximum SL 4.3 cm. Body elongated; somewhat fusiform; laterally compressed. Maximum body depth at middle pelvic and pectoral fin origin. Pre-dorsal profile straight or slightly convex behind head. Head length 23 % of SL; lack of breeding males limited knowledge on presence of tubercles. Snout rounded; short; 29 % of head length. Mouth terminal; slightly crescent-shaped with long anterior side; reaching anterior border of orbit. Nostrils large; level with dorsal margin of eye; separated from orbit by less than one orbit radius. Tubular anterior naris short; adjacent to open posterior naris. Eye lateral; visible from above and below (more prominent); 41 % of head length. 1st Ceratobranchial gill arch with 8 gill rakers. Gill rakers long; pointed; widely-spaced. 1st Epibranchial with 3 gill rakers. Pharyngeal teeth in four rows. Pharyngeal teeth 5,3,2,1; slender and long; falcate.

Modal fin formulae in Table 5.4. Fins large in relation to body size. Dorsal fin closer to caudal fin than tip of snout; more or less above origin of anal fin; length 14 % of SL; posterior margin straight; rays soft; anterior-most branched fin ray longest. Dorsal and anal fin point parallel. Pectoral fins largest; reaching $^1/_2$ to $^3/_4$ distance to base of pelvic fin; fin lacking lobe at base. Pelvic fins reaching $^2/_3$ distance to base of anal fin; relatively small; pointed; fin lacking a basal lobe. Anal fin moderately long; extending $^2/_3$ length of caudal peduncle; last unbranched ray longest. Ano-genital opening at anterior of base of anal fin. Caudal peduncle moderately long, 16 % of standard length; deep, depth 50 % of caudal length. Caudal fin forked; lobes pointed; upper lobe shorter.

Scales small to medium relative to body size; in regular rows; cycloid; radially striate; rounded, slightly elongate. Base of anal fin lacking sheath of scales. Lateral line present; complete; dipping sharply down towards ventral at tip of pectoral fin; joins midline at posterior of caudal peduncle; scale count 51–52.

Live colouration. Body and head unknown. Body colouration without vertical bars or bands.

Preserved colouration. Body and head orange with small dark brown spotting along dorsal surface, midline and above anal fin. Scales on dorsal surface lightly pigmented. Ventral scale pigmentation less intense than dorsal. Dorsal surface of head lightly pigmented. Melanophores small, dark; grouped on rear of head, below orbit, and on lips and snout; along midline, increasing in intensity to caudal fin; browner on dorsal surface, darkening between origin of pectoral and dorsal fin; forming small dark line above anal fin. Operculum with silver sheen. Side of body with silver sheen extending from pectoral fin to anal fin origin.

Membranes between fin rays white to clear towards end. Pelvic fin clear membranes and rays. Dorsal, caudal and pectoral fin membranes white to clear; rays with small, widely-spaced, melanophores fading towards end; rays pale brown to clear.

Geographical distribution. NAMIBIA, ANGOLA: Cunene River system.

Material examined. **Holotype**, SAIAB 39012, unsexed: SL 42.84 mm. Namibia, Kunene River, Olushandja, 17°43'S 14°67'E [SAIAB]. **Paratypes**, SAIAB 193063, 9 unsexed: SL 21.90–42.11 mm. Collected with holotype [SAIAB]; SAIAB 78759, 6 unsexed, SL 26.82–36.44 mm, Namibia, River Kunene at Hippo Pool below Ruacana Waterfall, 17°24'24"S 14°13'1"E, 21 August 2006, E. Swartz & Kramer [SAIAB]; SAIAB 38961, 4 unsexed, SL 28.97–34.59 mm, Namibia, Kunene River, 17°38'33"S 14°21'67"E, 27 May 1992, C. Hay, N. James & P. Skelton [SAIAB].

Biology. Very little is known of their biology. Appear to favour turbid, rocky, river regions where they can gather in pockets of recirculating currents. Predominantly found in turbid slack regions of the river. Feed on drifting invertebrates, larvae (diptera etc) and adult insects. (Bills, pers. com.).

Engraulicypris ngalala Riddin, Villet & Bills n. sp.

Holotype, SAIAB 74087, unsexed; [SAIAB].

Etymology. In the Cyao language spoken in the Niassa region of northern Mozambique, the name '*ngalala*' denotes all, small, compressed, silvery fishes, including *Mesobola*, but also *Brycinus* and *Hemigrammopetersius* species.

Morphology. (Fig. 5.1.D, 5.2.E. 5.3.E; Table 5.5, 5.6.E). Maximum SL 50.46 mm. Body elongated; somewhat fusiform; laterally compressed. Maximum body depth at middle pelvic and pectoral fin origin. Pre-dorsal profile straight or slightly convex behind head. Head length 18 % of SL; lack of breeding males limited knowledge on presence of tubercles. Snout rounded; short; 33 % of head length. Mouth terminal; slightly crescent-shaped with long anterior side; reaching anterior border of orbit. Nostrils large; level with dorsal margin of eye; separated from orbit by less than one orbit radius. Tubular anterior naris short; adjacent to open posterior naris. Eye lateral; visible from above and below (more prominent); 43 % of head length. 1st Ceratobranchial gill arch with 13 gill rakers. Gill rakers long; pointed; widely-spaced. 1st Epibranchial with three gill rakers. Pharyngeal teeth in four rows. Pharyngeal teeth 5,3,2,1; slender and long; falcate.

Modal fin formulae in Table 5.5. Fins large in relation to body size. Dorsal fin closer to caudal fin than tip of snout; more or less above origin of anal fin; length 14 % of SL; posterior margin straight; rays soft; anterior-most branched fin ray longest. Dorsal and anal fin point parallel. Pectoral fins largest; reaching $^1/_2$ to $^3/_4$ distance to base of pelvic fin; fin lacking lobe at base. Pelvic fins reaching $^2/_3$ distance to base of anal fin; relatively small; pointed; fin lacking a basal lobe. Anal fin moderately long; extending $^2/_3$ length of caudal peduncle; last unbranched ray longest. Ano-genital opening at anterior of base of anal fin. Caudal peduncle moderately long, 16 % standard length; deep, depth 50 % of caudal length. Caudal fin forked; lobes slightly concave interior lobe into point; upper lobe shorter.

Scales small to medium relative to body size; in regular rows; cycloid; radially striate; rounded, slightly elongate. Base of anal fin lacking sheath of enlarged, elongate scales. Lateral line present; complete; sharply down towards ventral at tip of pectoral fin; joins midline at posterior of caudal peduncle; scale count 51–52.

Live colouration. Body and head white ventrally with pale brown dorsal surface. Body midline black; colouration without vertical bars or bands. Dorsal surface with ubiquitous melanophores. Snout with dense dark spotting on tip. Operculum shiny only on ventral posterior edge and small area at posterior edge of orbit. Iris white to light grey with a few melanophores. Dorsal fin membrane clear; rays clear with dark melanophores. Caudal fin membrane clear to pale orange towards midline; rays dark brown to black; rays lightening towards edge; melanophores abundant and fading towards tips. Anal fin rays clear; membrane clear; pale orange spotting above origin; melanophores dark brown fading towards tips. Pectoral fin membranes clear; rays clear; first ray with abundant dark melanophores. Pelvic fin rays clear; membrane clear.

Preserved colouration. Body and head pale yellow with dark brown spotting on dorsal surface and midline. Scales on dorsal surface lightly pigmented. Ventral scale pigmentation less intense than dorsal. Dorsal surface of head lightly pigmented. Melanophores small, dark; grouped on rear of head, below orbit, and on lips and snout; along midline, increasing in intensity to caudal fin; browner on dorsal surface, darkening between origin of pectoral and dorsal fin; forming small dark line above anal fin. Operculum and posterior base on orbit with silver sheen. Membranes between fin rays white to clear towards end. Pelvic fin clear membranes with melanophores on first ray. Dorsal, caudal and pectoral fin rays with melanophores small, widely-spaced, fading towards end; pale brown to clear.

Geographical distribution. MOZAMBIQUE, MALAWI, TANZANIA: Rovuma River system and Lake Chiuta.

Material examined. **Holotype**, SAIAB 74087, unsexed: SL 40.03 mm. Lucheringo River, Singa Hunting Camp, 11°48'56"S 36°13'15"E, 25 August 2003, I.R. Bills [SAIAB]. **Paratypes**, SAIAB 193064, 1 unsexed: SL 42.25 mm. Collected with holotype. [SAIAB]; SAIAB 73944, 6 unsexed, SL 40.18–50.46 mm. Malawi, Rovuma River below Chamba, 12°35'47"S 36°56'8"E, 19 August 2003, I.R. Bills [SAIAB]; SAIAB 39269, 11 unsexed, SL 42.42–52.53 mm. Lake Chiuta at Mthubula Beach, 14°78'33"S 35°83'33"E, 13 July 1992, P. Skelton & D. Tweddle [SAIAB].

Biology. Very little is known about this new species. Found in conditions very similar to those of *gariepinus* as favour big rivers, gathering in the slack, turbid and shallow regions. In Lake Chiuta they are caught along the edges in reed beds. Biology in Lake Chiuta and Rovuma Rivers may differ as morphometry suggests slight differences in adaptation. Lake Chiuta would supply planktonic community associated with river beds (copepods etc.) for food that is not found in the rivers within the Rovuma. Within the rivers no available planktonic community and therefore, survive predominantly on invertebrate drift. Similar to other species favour turbid, sandy, rocky or muddy substrates where there are abundant areas of slack flow (Bills, pers.com.).

Key to African chedrin genera and the genus *Engraulicypris*.

The pigmentation of *E. howesi* is unknown, and it lacks diagnostic external characters, so it is omitted from the following key.

1. Pelvic fin lacking a basal lobe 2
Pelvic fin with a basal lobe 6

2(1). Mouth not reaching anterior border of orbit .... ***Rastrineobola***
Mouth reaching anterior border of orbit 3

3(2). Operculum shiny only on ventral posterior edge and small area at posterior edge of orbit; body midline black; iris white to light grey with a few melanophores; pelvic fin melanophores dark and widely dispersed .... ***Engraulicypris ngalala* Riddin, Villet & Bills n. sp.**
Operculum entirely shiny; body midline silver; iris dark to light grey; pelvic fin melanophores absent. 4

4(3). Body laterally compressed; head with tubercles in breeding males; snout rounded 5
Body fusiform; body cylindrical; head never with tubercles; snout pointed .... ***Engraulicypris sardella* (Günther, 1868)**

5(4). Caudal fin membrane clear towards vivid yellow at fork; anal fin extending 2/3 length of caudal peduncle; caudal peduncle moderately long; snout darker dorsally .... ***Engraulicypris brevianalis* (Boulenger, 1908)**
Caudal fin membrane clear to pale orange towards midline; anal fin extending over $^3/_4$ length of caudal peduncle; caudal peduncle short; snout with dense dark spotting on tip .... ***Engraulicypris gariepinus* Barnard, 1943 stat. rev.**

6(1). Dorsal fin distinctly forward of origin of anal fin..7
Dorsal fin more or less above origin of anal fin .... ***Neobola***
Dorsal fin slightly behind origin of anal fin.. ***Chelaethiops***

7(6). Body colouration without vertical bars or bands .... ***Leptocypris***
Body colouration with numerous vertical bars 8

8(7). Mouth reaching anterior border of orbit .... ***Opsaridium***

Mouth reaching middle of orbit .... ***Raiamas***

5.4 References

BARNARD, K.H. 1943. Revision of the indigenous freshwater fishes of the s.w. Cape region. *Annals of the South African Museum*, **36(2)**: 101-262.

BOULENGER, G.A. 1908. On a collection of freshwater fishes, batrachians and reptiles from Natal and Zululand, with descriptions of new species. *Annals of the Natal Government Museum*, **1(3)**: 219-235.

BOULENGER, G.D. 1911. *Catalogue of the fresh-water fishes of Africa in the British Museum (Natural History)*. London. v. 2: 529.

CRISP, M.D., CHANDLER, G.T. 1996. Paraphyletic species. *Telopea*, **6(4)**: 813-844.

DALLWITZ, M.J. 1980. A general system for coding taxonomic descriptions. *Taxon*, **29**: 41-46.

DALLWITZ, M.J., PAINE, T.A., ZURCHER, E.J. 1993. User's guide to the DELTA System: a general system for processing taxonomic descriptions. 4th edition. http://delta-intkey.com.

ECCLES, D.H., 1992. *FAO species identification sheets for fishery purposes. Field guide to the freshwater fishes of Tanzania*. Prepared and published with the support of the United Nations Development Programme (project URT/87/016). FAO, Rome. 145 pp

ENGELBRECHT, G.D., MULDER, P.F.S. 1999. Allozyme variation in the river sardine, *Mesobola brevianalis* (Pisces, Cyprinidae). *Water SA*, **25(3)**: 293–296.

ESCHMEYER, W. N. (eds). *Catalog of Fishes*. California Academy of Sciences (http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 2014.

FOWLER, H.W. 1936. Zoological results of the George Vanderbilt African Expedition of 1934: Part III-The fresh water fishes. *Proceeding of the Academy of Natural Sciences of Philadelphia*, **88**: 243-335.

FOWLER, H.W. 1977. A catalog of world fishes. *Quarterly Journal of the Taiwan Museum*, **30**: 88.

FROESE, R., PAULY, D. (eds). *FishBase*. World Wide Web electronic publication. (www.fishbase.org). Electronic version accessed 2014.

GÜNTHER, A. 1868. *Catalogue of the fishes in the British Museum. Catalogue of the Physostomi, Containing the Families Heteropygii, Cyprinidae, Gonorhynchidae, Hyodontidae, Osteoglossidae, Clupeidae, Chirocentridae, Alepocephalidae,* *Notopteridae, Halosauridae, in the Collection of the British Museum, vol 7.* Taylor and Francis. London. pp. 292.

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*, **4**: 616-628.

HOWES, G.J. 1980. The anatomy, phylogeny and classification of bariliine cyprinid fishes. *Bulletin of the British Natural History Museum (Zoology Series)*, **37 (3)**: 129-198.

HOWES, G.J. 1984. A review of the anatomy, taxonomy, phylogeny and biogeography of the African neoboline cyprinid fishes. *Bulletin of the British Museum (Natural History) Zoology*, **47(3)**: 151-185.

JUBB, R.A. 1963. A revised list of the freshwater fishes of southern Africa. *Annals of Cape Provincial Museums*, **3**: 5-39.

JUBB, R.A. 1965. Freshwater fishes of the Cape Province. *Annals of Cape Provincial Museums*, **4**: 1-72.

JUBB, R.A., FARQUHARSON, F.L. 1965. The freshwater fishes of the Orange River drainage basin. *South African Journal of Science*, **61(3)**: 118-125.

KONINGS, A., 1990. Ad Konings's book of cichlids and all of the other fishes of Lake Malawi. T.F.H. Publications, Inc. 495 pp Lewis, D.S.C. and D. Tweddle, 1990. The yield of Usipa (*Engraulicypris sardella*) from the Nankumba Peninsula, Lake Malawi (1985–1986). *Collected Reports on Fisheries Research in Malawi. Occas. Pap.* **1**: 57–66.

LÉVÊQUE, C. DAGET, J. 1984. *Cyprinidae* (pp. 217-342). In: Daget, J., GOSSE, J.P., TEUGELS, G.G., THYS VAN DEN AUDENAERDE, D.F.E. 1991. Check list of the freshwater fishes of Africa. Musée Royal de l' Afrique Centrale. Paris. France. ISBN 907324203-7.

LEWIS, D., TWEDDLE, D. 1991. The yield of Usipa (*Engraulicypris sardella*) from the Nankumba Peninsula, Lake Malawi (1985-86). *Collected Reports of Fisheries Research in Malawi, Occas. Papers*, **1**: 57-66.

MACE, G.M. 2004. The role of taxonomy in species conservation. Philosophical *Transactions of the Royal Society of Biological Sciences*, **359**: 711-719.

PELLEGRIN, J. 1904. Diagnoses préliminaires de Poissons nouveaux du Lac Victoria, recueillis par M. Alluaud. *Bulletin de la Société zoologique de France*. **29 (8)**: 184-186.

POLL, M., GOSSE, J.P. 1995. Genera des poissons de l'eau douce de l'Afrique. Mémoire de la Classe des Sciences. *Académie royale de Belgique*, **9**: 1-324.

RUBINOFF, D. 2006. Utility of Mitochondrial DNA Barcodes in Species Conservation. *Conservation biology*, **20(4):** 1026-1033.

SKELTON, P.H., 1993. *A complete guide to the freshwater fishes of southern Africa*. Southern Book Publishers. pp 388.

SKELTON, P.H., 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Book Publishers. Cape Town South Africa. pp 119-120.

SKELTON, P.H., TWEDDLE, D., JACKSON, P.B.N. 1991. *Cyprinids of Africa*. In: Winfield, I.J., Nelson, J.S. (eds). Cyprinid Fishes- Systematics, Biology and Exploitation. Chapman & Hall. London, pp 211–239.

TANG, K.L., AGNEW, M.K., HIRT, M.V., SADO, T., SCHNEIDER, L.M., FREYHOF, J., SULAIMAN, Z., SWARTZ, E., VIDTHAYANON, C., MIYA, M., SAITOH, K., SIMONS, A.M., WOOD, R.M., MAYDEN, R.L. 2010. Systematics of the subfamily Danioninae (Teleostei: Cypriniformes: Cyprinidae). *Molecular Phylogenetics and Evolution*, **57**: 189-214.

TAYLOR, W.R., VAN DYKE, G.C. 1985. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. *Cybium*, **9(2)**: 107-119.

THOMPSON, A.B., IRVINE, K. 1997. Diet-shifts and food-dependant survival in *Engraulicypris sardella* (Cyprinidae) larvae from Lake Malawi, Africa. *Journal of Plankton Research*, **19(3)**: 287–301.

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

Table 5.1. Morphometric and meristic measurements and counts for *Engraulicypris sardella* type series.

Measurement	N	Holotype	Range			
			Max	Min	Mean	SD
SL	6	56.22	64.12	56.22	59.71	3.36
Head length (%SL)	6	13.62	24.22	20.90	22.48	1.28
Head depth (%HL)	6	7.26	61.78	53.35	56.23	3.36
Snout length (%HL)	6	4.34	40.60	31.58	35.32	3.34
Orbit diameter (%HL)	6	3.34	34.47	24.55	31.21	3.57
Postorbit length (%HL)	6	5.53	40.60	26.38	31.21	5.01
Inter-orbit length (%HL)	5	Unknown	28.77	23.89	25.72	1.88
Predorsal length (%SL)	6	36.24	67.53	64.46	66.30	1.13
Prepelvic length (%SL)	6	29.84	53.08	49.43	51.14	1.28
Dorsal fin Length (%SL)	6	6.73	13.21	11.70	12.56	0.60
Pectoral fin length (%SL)	6	7.94	17.70	14.12	15.32	1.40
Pelvic fin length (%SL)	6	5.73	10.85	10.19	10.44	0.26
Anal fin length (%SL)	6	3.87	14.68	6.89	11.62	2.59
Body depth (%SL)	6	9.02	16.51	14.73	15.63	0.71
Body width (%SL)	5	Unknown	11.06	8.47	9.57	0.96
Caudal peduncle length (%SL)	6	9.01	21.41	16.02	18.91	1.83
Caudal peduncle depth (%SL)	6	4.57	8.31	7.05	7.71	0.47

Meristics	N	Holotype	Range
Dorsal-fin rays	n/a	Unknown	>9 (Günther 1894)
Anal-fin rays	n/a	Unknown	~14 (Günther 1894)
Pectoral-fin rays	n/a	Unknown	Unknown
Pelvic-fin rays	n/a	Unknown	Unknown
Lateral line scales	2	Unknown	55 (N=2)
Caudal peduncle scale	2	Unknown	12 (N=1), 14 (N=2)
Scale rows lat.line-dorsal	2	Unknown	7 (N=2)
Scale rows lat.line-pelvic	2	Unknown	2 (N=2)
Scale rows lat.line-anal	2	Unknown	2 (N=2)
Total vertebrae	5	Unknown	43 (N=2), 44 (N=3)
Abdominal vertebrae	5	Unknown	23 (N=5)
Caudal vertebrae	5	Unknown	21 (N=2), 22 (N=4)
Rib pairs	5	Unknown	15 (N=1), 16 (N=1), 17 (N=3)

Table 5.2. Morphometric and meristic measurements and counts for *Engraulicypris brevianalis* type series.

Measurement	**N**	**Holotype**	**Range**			
			Max	**Min**	**Mean**	**SD**
SL	6	67.39	67.39	41.13	46.46	10.30
Head length (%SL)	6	16.56	24.57	19.74	21.20	1.86
Head depth (%HL)	6	11.59	86.24	69.80	74.61	6.09
Snout length (%HL)	6	3.53	37.96	21.29	32.95	5.96
Orbit diameter (%HL)	6	5.65	45.85	34.14	39.69	4.68
Postorbit length (%HL)	6	6.34	38.27	25.62	30.68	4.74
Inter-orbit length (%HL)	5	Unknown	47.19	36.40	44.52	4.59
Predorsal length (%SL)	6	44.35	65.80	62.36	64.00	1.33
Prepelvic length (%SL)	6	33.46	50.78	46.48	48.23	1.80
Dorsal fin Length (%SL)	6	12.32	18.64	16.01	17.55	0.97
Pectoral fin length (%SL)	6	13.40	22.32	19.88	21.66	0.99
Pelvic fin length (%SL)	6	8.98	13.66	11.37	12.31	0.97
Anal fin length (%SL)	6	12.08	17.92	14.33	16.07	1.33
Body depth (%SL)	6	14.27	23.54	20.53	21.96	1.10
Body width (%SL)	5	Unknown	13.46	10.69	12.08	1.05
Caudal peduncle length (%SL)	6	10.04	18.10	14.10	15.74	1.57
Caudal peduncle depth (%SL)	6	6.78	11.01	9.52	10.24	0.58

Meristics	**N**	**Holotype**	**Range**
Dorsal-fin rays	5	Unknown	ii + 8 (N=5)
Anal-fin rays	5	Unknown	iii + 13 (N=1), iii + 14 (N=3), iii + 15 (N=1)
Pectoral-fin rays	5	Unknown	i + 10 (N=4), i + 11 (N=1)
Pelvic-fin rays	5	Unknown	i + 7 (N=5)
Lateral line scales	2	Unknown	53 (N=1), 57 (N=1)
Caudal peduncle scale	2	Unknown	18 (N=2)
Scale rows lat.line-dorsal	2	Unknown	9 (N=1), 11 (N=1)
Scale rows lat.line-pelvic	2	Unknown	2 (N=2)
Scale rows lat.line-anal	2	Unknown	2 (N=2)
Total vertebrae	5	37	37 (N=1), 38 (N=4)
Abdominal vertebrae	5	19	18 (N=2), 19 (N=3)
Caudal vertebrae	5	18	19 (N=4), 20 (N=1)
Rib pairs	5	14	13 (N=1), 14 (N=3), 15 (N=1)

Table 5.3. Morphometric and meristic measurements and counts for *Engraulicypris gariepinus* type series.

Measurement	N	Neotype	Range			
			Max	Min	Mean	SD
SL	20	42.48	46.61	28.76	38.36	5.41
Head length (%SL)	20	21.96	23.78	18.25	21.18	1.79
Head depth (%HL)	20	76.31	93.35	59.71	77.75	7.75
Snout length (%HL)	20	27.87	40.37	23.10	31.77	4.69
Orbit diameter (%HL)	20	34.51	50.35	36.89	40.27	3.32
Postorbit length (%HL)	20	39.34	38.75	27.56	33.06	3.03
Inter-orbit length (%HL)	20	36.33	48.23	27.79	38.10	5.68
Predorsal length (%SL)	20	64.24	68.09	62.17	64.83	1.71
Prepelvic length (%SL)	20	50.71	55.17	45.07	49.03	2.32
Dorsal fin Length (%SL)	20	16.83	20.39	15.19	17.00	1.37
Pectoral fin length (%SL)	20	22.03	25.04	21.60	22.90	1.00
Pelvic fin length (%SL)	20	12.88	14.25	11.80	12.79	0.68
Anal fin length (%SL)	20	15.58	18.45	14.38	16.41	1.12
Body depth (%SL)	20	22.76	25.64	18.39	21.81	1.91
Body width (%SL)	20	11.46	13.27	9.39	11.11	1.19
Caudal peduncle length (%SL)	20	17.51	16.64	11.13	13.95	1.52
Caudal peduncle depth (%SL)	20	8.38	9.84	7.13	8.18	0.79

Meristics	N	Neotype	Range
Dorsal-fin rays	20	ii + 8	ii + 7 (N=3), ii + 8 (N=17)
Anal-fin rays	20	iii + 15	iii + 14 (N=4), iii + 15 (N=9), iii + 16 (N=7)
Pectoral-fin rays	20	i + 10	i + 9 (N=10), i + 10 (N=8), i + 11 (N=2)
Pelvic-fin rays	20	i + 7	i + 7 (N=19), i + 8 (N=1)
Lateral line scales	2	Unknown	49 (N=1), 51 (N=1)
Caudal peduncle scale	2	Unknown	14 (N=1), 16 (N=1)
Scale rows lat.line-dorsal	2	Unknown	11 (N=2)
Scale rows lat.line-pelvic	2	Unknown	2 (N=2)
Scale rows lat.line-anal	2	Unknown	2 (N=2)
Total vertebrae	12	39	36 (N=1), 38 (N=3)
Abdominal vertebrae	12	19	17 (N=1), 18 (N=1), 19 (N=10)
Caudal vertebrae	12	20	19 (N=3), 20 (N=8), 21 (N=1)
Rib pairs	12	14	13 (N=3), 14 (N=7), 15 (N=2)

Table 5.4. Morphometric and meristic measurements and counts for *Engraulicypris howesi* type series.

Measurement	N	Holotype	Range			
			Max	**Min**	**Mean**	**SD**
SL	20	42.84	42.84	21.90	33.35	5.03
Head length (%SL)	20	21.62	25.04	21.43	22.76	1.11
Head depth (%HL)	20	64.25	74.52	56.81	65.62	5.61
Snout length (%HL)	20	25.27	34.30	21.22	28.53	3.95
Orbit diameter (%HL)	20	37.80	46.29	32.49	40.62	3.42
Postorbit length (%HL)	20	37.37	43.86	29.95	36.34	4.01
Inter-orbit length (%HL)	20	32.07	43.25	5.79	33.35	8.38
Predorsal length (%SL)	20	62.61	67.19	62.61	64.78	1.33
Prepelvic length (%SL)	20	46.27	51.11	40.37	47.48	2.57
Dorsal fin Length (%SL)	20	16.15	17.63	9.52	14.30	2.43
Pectoral fin length (%SL)	20	20.12	24.16	15.50	18.86	2.17
Pelvic fin length (%SL)	20	11.83	14.19	9.05	12.06	1.19
Anal fin length (%SL)	20	15.90	16.80	9.22	13.69	2.16
Body depth (%SL)	20	18.49	20.42	14.22	17.83	1.59
Body width (%SL)	20	10.04	10.96	5.82	9.36	1.46
Caudal peduncle length (%SL)	20	18.98	18.98	13.15	15.91	1.52
Caudal peduncle depth (%SL)	20	9.45	9.86	7.01	8.51	0.92

Meristics	N	Holotype	Range
Dorsal-fin rays	20	ii + 8	ii + 6 (N=2), ii + 7 (N=6), ii + 7 (N=12)
Anal-fin rays	20	iii + 13	iii + 13 (N=9), iii + 14 (N=6), iii + 15 (N=5)
Pectoral-fin rays	20	i + 10	i + 8 (N=2), i + 9 (N=17), i + 10 (N=1)
Pelvic-fin rays	20	i + 7	i + 6 (N=1), i + 7 (N=17), i + 8 (N=2)
Lateral line scales	2	Unknown	51 (N=1), 52 (N=1)
Caudal peduncle scale	2	Unknown	14 (N=2)
Scale rows lat.line-dorsal	2	Unknown	9 (N=2)
Scale rows lat.line-pelvic	2	Unknown	2 (N=2)
Scale rows lat.line-anal	2	Unknown	2 (N=2)
Total vertebrae	11	38	38 (N=3), 39 (N=7), 40 (N=1)
Abdominal vertebrae	11	19	19 (N=10), 20 (N=1)
Caudal vertebrae	11	19	19 (N=4), 20 (N=7)
Rib pairs	11	14	13 (N=5), 14 (N=6)

Table 5.5. Morphometric and meristic measurements and counts for *Engraulicypris ngalala* type series.

Measurement	N	Holotype	Range			
			Max	Min	Mean	SD
SL	20	40.03	50.46	19.37	43.95	2.90
Head length (%SL)	20	7.59	22.79	16.90	18.44	1.26
Head depth (%HL)	20	5.67	77.08	56.85	73.59	2.75
Snout length (%HL)	20	2.54	40.91	15.34	33.10	3.28
Orbit diameter (%HL)	20	3.47	46.01	30.50	43.01	2.67
Postorbit length (%HL)	20	2.54	38.22	22.50	33.18	3.24
Inter-orbit length (%HL)	20	2.68	43.41	21.26	39.13	2.46
Predorsal length (%SL)	20	25.35	68.23	62.27	64.10	1.82
Prepelvic length (%SL)	20	19.27	51.15	42.62	47.72	2.34
Dorsal fin Length (%SL)	20	5.43	19.27	11.89	14.04	1.81
Pectoral fin length (%SL)	20	8.88	23.15	19.29	21.84	0.98
Pelvic fin length (%SL)	20	5.01	17.66	10.57	13.20	1.53
Anal fin length (%SL)	20	6.31	17.52	12.66	14.57	0.97
Body depth (%SL)	20	7.46	20.68	13.89	18.37	0.99
Body width (%SL)	20	3.43	9.26	3.05	8.07	0.69
Caudal peduncle length (%SL)	20	5.98	18.31	12.82	15.89	1.32
Caudal peduncle depth (%SL)	20	3.24	10.13	5.33	8.46	0.64

Meristics	N	Holotype	Range
Dorsal-fin rays	20	ii + 7	ii + 7 (N=15), ii + 8 (N=5)
Anal-fin rays	20	iii + 14	iii + 13 (N=3), iii + 14 (N=6), 3 + 15 (N=9), iii + 16 (N=2)
Pectoral-fin rays	20	i + 10	i+ 8 (N=3), i + 9 (N=11), i + 10 (N=6)
Pelvic-fin rays	20	i + 7	i + 6 (N=2), i + 7 (N=17), i + 8 (N=1)
Lateral line scales	2	Unknown	51 (N=1), 52 (N=1)
Caudal peduncle scale	2	Unknown	14 (N=1), 16 (N=2)
Scale rows lat.line-dorsal	2	Unknown	9 (N=2)
Scale rows lat.line-pelvic	2	Unknown	2 (N=2)
Scale rows lat.line-anal	2	Unknown	1 (N=1), 2 (N=1)
Total vertebrae	14	38	38 (N=2), 39 (N=1)
Abdominal vertebrae	14	19	19 (N=12), 20 (N=2)
Caudal vertebrae	14	19	19 (N=4), 20 (N=10)
Rib pairs	14	14	14 (N=1), 15 (N=13)

Table 5.6. COI sequences extracted from holotype or other material populations.

A) *Engraulicypris sardella* (Lake Malawi: cataloguing in process) populations COI-

CCCTTTATCTTGTATTTGGTGCTTGAGCTGGTATAGTAGGAACCGCCCTTAGTCTCTTAATCCGCGCTGAGCTG
AGCCAGCCAGGATCACTTCTGGGCGATGATCAAATCTACAATGTTATCGTTACTGCTCATGCTTTCGTAATAAT
TTTCTTTATAGTGATGCCGATCCTCATTGGGGGGTTTGGGAACTGATTAGTACCACTAATGATTGGGGCCCCAG
ACATGGCGTTCCCTCGAATAAACAATATAAGCTTCTGACTCCTACCCCCATCATTCCTCTTGCTTCTAGCCTCTT
CTGGCGTAGAAGCTGGTGCCGGAACTGGGTGAACAGTGTACCCCCCATTGTCAGGGAACCTAGCCCACGCGG
GAGCATCAGTAGATTTAACAATCTTTTCTCTCCATCTGGCGGGTGTGTCATCTATCTTAGGGGCTATTAACTTT
ATTACTACAACTATTAACATGAAACCTCCAGCCATTTCTCAGTACCAAACACCCCTGTTTGTCTGAGCCGTCTT
GGTAACAGCTGTTCTTCTTCTACTATCGCTTCCCGTGTTAGCTGCTGGCATTACAATACTTTTAACAGACCGTA
ACCTTAACACCTCATTTTTTGATCCTGCAGGAGGGGGAGATCCTATTCTCTACCAACACCTGTTCTGATTTTTT

B) *Engraulicypris brevianalis* 190643 population COI-

CCCTTTATCTTGTATTTGGTGCTTGAGCTGGTATAGTAGGAACCGCCCTTAGTCTCTTAATCCGCGCTGAGCTA
AGCCAGCCGGGATCACTTCTGGGCGATGATCAAATCTACAATGTTATCGTTACTGCTCATGCTTTCGTAATAAT
TTTCTTTATAGTGATGCCCATCCTCATTGGGGGGTTTGGGAACTGATTAGTACCACTAATGATTGGGGCCCCAG
ACATGGCGTTCCCTCGTATAAACAATATAAGCTTCTGACTCCTGCCCCCGTCATTCCTCTTGCTTCTAGCCTCTT
CTGGCGTAGAGGCTGGCGCCGGGACTGGATGAACAGTATACCCCCCATTAGCAGGGAACCTAGCCCACGCGG
GAGCATCAGTAGATTTAACAATCTTTTCTCTTCATCTGGCAGGTGTATCATCTATCTTAGGGGCTATTAACTTT
ATTACTACAACTATTAATATGAAACCTCCAGCCATTTCTCAGTACCAAACGCCCCTGTTTGTCTGAGCCGTCTT
AGTAACAGCTGTTCTTCTACTACTTTCACTTCCCGTGCTAGCTGCTGGCATTACAATACTTTTAACAGACCGTA
ATCTTAACACCTCGTTTTTTGATCCTGCGGGAGGGGGAGATCCTATTCTTTACCAACACCTATTTTGATTTTTT

C) *Engraulicypris gariepinus* 78831 Holotype population COI-

AAAGATATTGGCACCCTTTATCTTGTATTCGGTGCTTGAGCTGGTATAGTAGGAACCGCCCTTAGTCTCTTAAT
CCGCGCTGAGCTGAGCCAGCCGGGATCACTTCTGGGCGATGATCAAATCTACAATGTTATCGTTACTGCTCAT
GCTTTCGTAATAATTTTCTTTATAGTGATGCCCATCCTCATTGGGGGGTTTGGGAACTGATTGGTACCATTAAT
GATTGGGGCCCCGGACATGGCATTCCCTCGAATAAACAACATAAGCTTCTGACTTCTGCCCCCGTCATTCCTCC
TGCTTTTAGCCTCTTCTGGTGTAGAAGCTGGCGCCGGGACTGGGTGAACAGTATACCCCCCATTGTCGGGAAA
CCTGGCCCACGCGGGAGCATCAGTAGATTTAACAATCTTTTCTCTTCATCTGGCGGGTGTATCATCTATCTTAG
GGGCTATTAACTTTATTACTACAACTATTAATATGAAACCTCCAGCCATTTCTCAGTACCAGACACCCCTGTTT
GTCTGAGCCGTCCTAGTAACAGCTGTTCTTCTCCTACTATCGCTTCCAGTGCTAGCTGCTGGCATTACAATGCT
TCTAACAGACCGTAATCTTAACACCTCATTTTTTGATCCTGCAGGAGGAGGAGATCCTATTCTTTACCAACACC
TATTCTGATTTTTT

D) *Engraulicypris howesi* 39012 Holotype population COI-

AAAGATATTGGCACCCTATATCTCGTATTTGGTGCTTGAGCTGGTATAGTAGGAACCGCCCTTAGTCTCTTAAT
CCGCGCTGAGCTGAGCCAGCCGGGATCACTTCTGGGCGATGATCAAATCTACAATGTTATCGTTACTGCTCAT
GCTTTCGTAATAATTTTCTTTATAGTGATGCCCATCCTCATTGGGGGATTTGGGAACTGACTGGTACCCCTAAT
GATTGGGGCCCCGGACATAGCATTCCCTCGGATAAACAATATAAGCTTCTGACTCCTACCCCCATCATTCCTCT
TGCTTCTAGCCTCTTCTGGTGTAGAAGCTGGCGCTGGGACCGGATGAACAGTATACCCCCCATTGGCAGGGAA
CCTGGCCCACGCGGGAGCATCAGTAGATTTAACAATCTTTTCTCTTCATCTAGCGGGTGTGTCATCTATCTTAG
GGGCTATTAACTTTATTACTACAACCATTAACATGAAACCTCCAGCCATTTCTCAGTACCAAACACCCCTGTTT
GTTTGGGCCGTTTTAGTAACAGCTGTTCTTCTGCTACTATCGCTTCCCGTACTAGCTGCTGGCATTACAATGCTT
TTAACAGACCGTAATCTTAACACCTCATTTTTTGACCCTGCAGGAGGAGGAGATCCTATTCTTTATCAACACCT
GTTCTGATTTTTT--

E) *Engraulicypris ngalala* 74087 Holotype population COI-

AAAGATATTGGCACCCTTTATCTTGTATTTGGTGCTTGAGCTGGTATAGTAGGAACCGCCCTTAGTCTCTTAAT
CCGCGCTGAGCTTAGCCAGCCAGGATCACTTCTGGGCGATGATCAAATTTACAATGTTATCGTTACTGCTCATG
CTTTCGTAATAATTTTCTTTATAGTGATGCCGATCCTTATTGGGGGGTTCGGGAACTGATTAGTACCATTAATG
ATTGGGGCCCCAGACATGGCATTCCCTCGAATAAACAATATAAGCTTCTGACTCCTGCCCCCATCATTCCTCTT
GCTTCTAGCCTCTTCTGGTGTAGAAGCTGGCGCCGGAACTGGATGAACAGTGTACCCCCCATTGTCAGGGAAC
CTGGCTCACGCGGGAGCATCAGTAGATTTAACAATCTTTTCTCTCCATCTGGCAGGTGTGTCATCTATCTTAGG
GGCTATTAACTTTATTACTACGACTATTAACATGAAGCCTCCAGCCATTTCTCAGTACCAGACACCCCTGTTTG
TCTGAGCCGTTCTAGTAACAGCTGTTCTTCTCCTATTATCGCTTCCCGTGCTAGCTGCTGGCATTACAATACTTT
TAACAGACCGTAATCTTAACACCTCGTTTTTTGACCCTGCAGGAGGGGGAGATCCTATTCTTTACCAGCACCTG
TTCTGATTTTTT

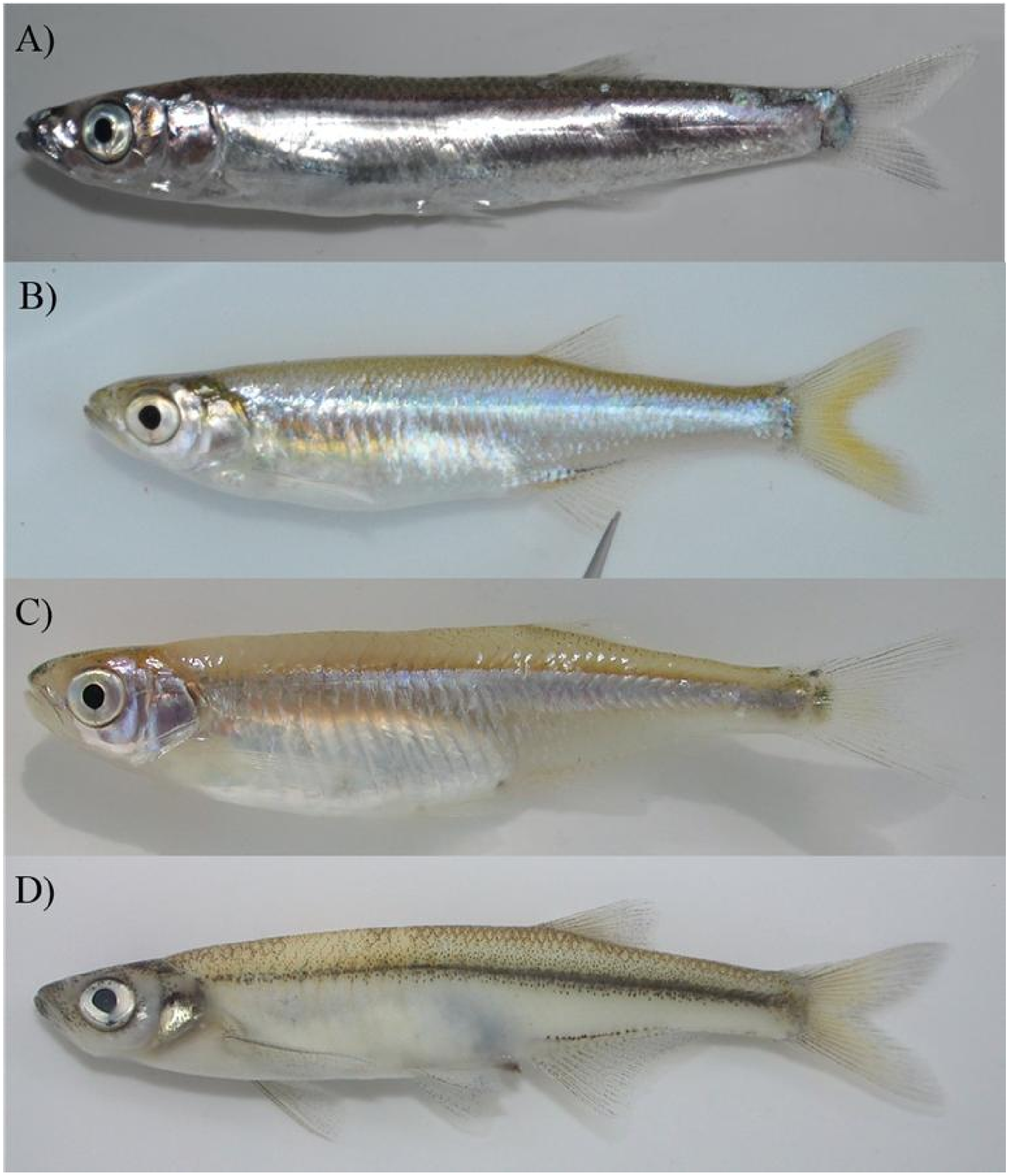


Fig. 5.1. Photographs of live specimens from each species: A - *Engraulicypris sardella*, B - *Engraulicypris brevianalis*, C - *Engraulicypris gariepinus*, D - *Engraulicypris ngalala.*

Fig. 5.2. Photographs of preserved holotypes or neotypes for each species: A - *Engraulicypris sardella*, B - *Engraulicypris brevianalis*, C - *Engraulicypris gariepinus* (neotype), D - *Engraulicypris howesi*, E - *Engraulicypris ngalala*.

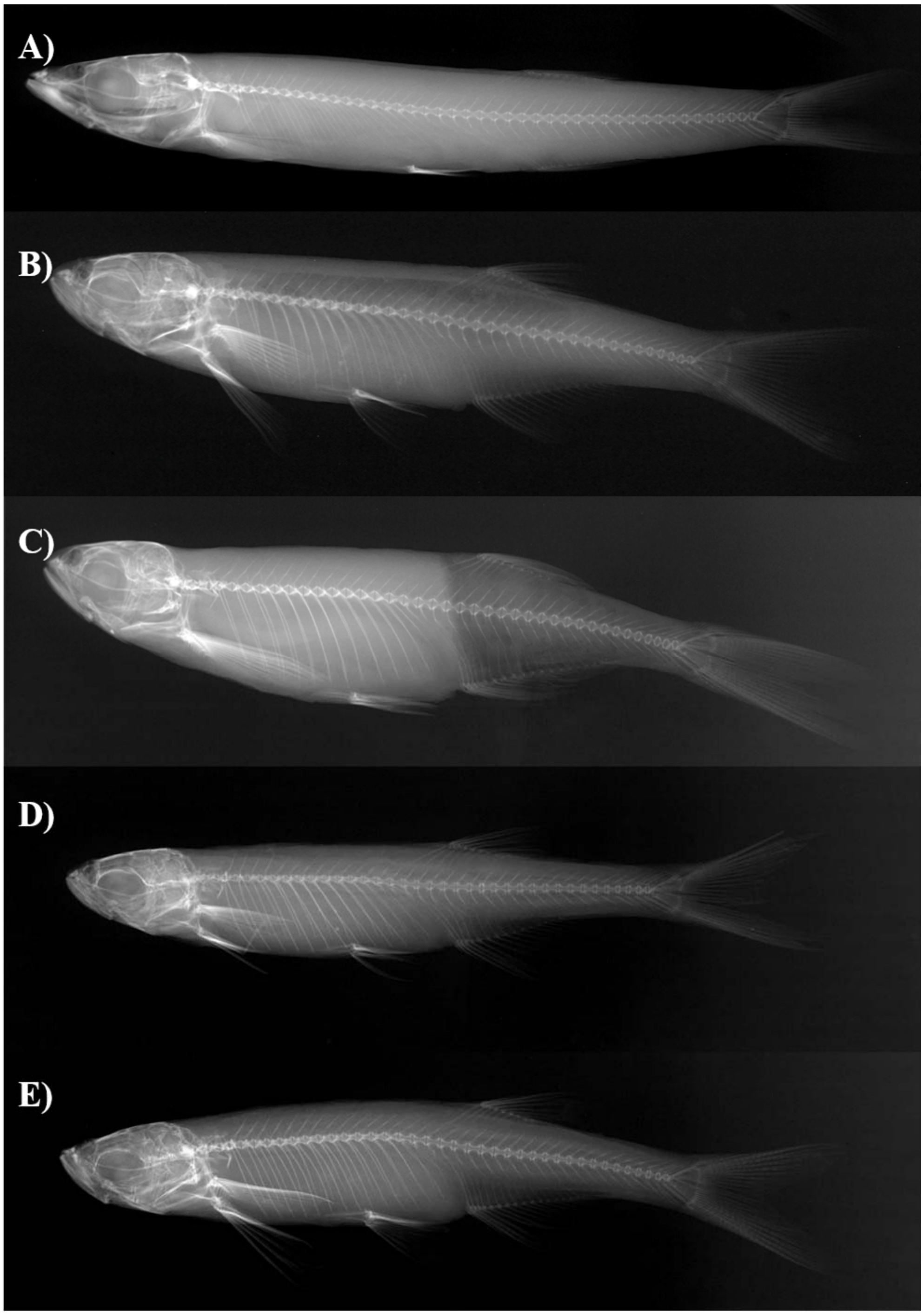


Fig. 5.3. X-rayed specimen from each species showing vertebrae structure: A – *Engraulicypris sardella*, B - *Engraulicypris brevianalis,* C - *Engraulicypris gariepinus*, D - *Engraulicypris howesi*, E - *Engraulicypris ngalala*.