Genetic diversity, evolutionary relationships and conservation of southern African *Labeo* fishes in relation to water management

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

DOCTOR OF PHILOSOPHY

of

RHODES UNIVERSITY

by

MPHO RAMOEJANE

October 2016

ABSTRACT

Labeo spp. are large, herbivorous fishes that are important components of aquatic ecosystems and are a high conservation priority in South Africa. This thesis contributes to determination of conservation priorities for *Labeo umbratus* (Smith 1841) by resolving the taxonomic status of this species in the evolutionary context of southern African *Labeo* spp., assessing the presence of unique lineages in historically isolated river basins, and assessing the threat of intra- and interspecific hybridisation associated with introductions. Phylogenetic analyses of five DNA sequence data sets (cytochrome *c* oxidase subunit I gene [COI], cytochrome *b* gene [Cyt *b*], Recombination activating gene 1 [Rag1], COI+Rag1 and COI+Cyt *b*+Rag1) showed that the *Labeo umbratus* group (*sensu* Reid, 1985), which comprises the species *Labeo umbratus*, *Labeo capensis* (Smith 1841), *Labeo seeberi* Gilchrist and Thompson 1911 and *Labeo rubromaculatus* Gilchrist and Thompson 1913, is monophyletic, morphologically distinct and geographically disjunct from other African *Labeo* spp. groups except in the Tugela River system were *L. rubromaculatus* co-occurs with *Labeo molybdinus* Du Plessis 1963. Phylogeographic analysis of mitochondrial DNA (Cyt *b*) sequence data demonstrated that the populations of the *L. umbratus* from the Orange and the southward-flowing river systems are reciprocally monophyletic and were identified as evolutionary significant units. The populations in the southward-flowing river systems were further divided into southwestern (Gourits and Gamtoos) and southeastern (Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon) polyphyletic sublineages. Four management units (Gourits Basin; Gamtoos Basin; Sundays+Bushmans+Great Fish River Basins; and Keiskamma+Buffalo+Nahoon River Basin) were not reciprocally monophyletic but were proposed on the basis of

containing unique haplotype frequencies for conservation purposes. To evaluate the threat of hybridisation to the genetic integrity of *L. umbratus*, the occurrence of *Labeo umbratus* × *L. capensis* hybrids was investigated using mtDNA Cyt *b* and nDNA S7 intron sequence data and morphological data. Genetic evidence for interspecific hybridisation was detected for populations in two impoundments, Hardap Dam (Orange River Basin) and Darlington Dam (Sundays River Basin, Eastern Cape, South Africa). Some putative hybrids were identifiable morphologically on account of intermediacy between the parental species in meristic and morphometric characters. Translocation via direct stocking (Hardap Dam) or via an inter-basin water transfer scheme (Darlington Dam) was identified as a driver for hybridisation. Introductions associated with an inter-basin water transfer scheme has resulted in introgression between the previously isolated Orange River and southern lineages of *L. umbratus*. Further translocation of fish from these affected areas to non-contaminated river systems and impoundments such as Kat River (Great Fish River) and Slagboom (Sundays River) should be avoided.

Table of Contents

DECLARATION

I, the undersigned, declare that to the best of my knowledge, the thesis hereby submitted for the Doctor of Philosophy in Ichthyology in the Department of Ichthyology and Fisheries Science, Rhodes University, is my original work. This thesis has not been previously submitted in any form to another university. I have not included ideas, phrases, passages or illustrations from another person's work without acknowledging their authorship.

Name: Mpho Ramoejane

Student Number: g08R6192

Signature..... Date.....

DEDICATION

This thesis is dedicated to the three very important women in my life. Firstly, my grandmother ("Mme" Mary Nomasonto Ramoejane) who recently passed on (03-08-2015). Mme, I am the man I am today because of your guidance and love. May your soul rest in peace. Secondly, my mother (Dibuseng Gladys Ramoejane), for standing by my decision to do a PhD, and for her support and love thereafter during what was a very difficult journey. Finally, to Tebogo Malatsi: our relationship is as old as this PhD and you were my strong shoulder on which to lean.

ACKNOWLEDGEMENTS

My supervisor, Dr Olaf Weyl, for his patience and understanding during what seemed like the end of days in my PhD time. I could not adhere to my deadlines, but you still stood by me and pushed me hard to complete. I know this was difficult for you too, as this PhD was based mostly on a field of study that is not your speciality; you still got me to finish.

My co-supervisor, Dr Ernst Swartz, for your help whenever you could, even though you were far away. Your leaving SAIAB made things harder for me, but you managed to help where you could.

My other co-supervisor, Dr Brian Sidlauskas; I'm better at analysing morphological characteristics because of you. Thank you also for your comments on the chapters sent to you to review.

Dr Gavin Gouws for informally supervising a very large part of the final PhD process. You also helping me with the use of genetics programmes and the interpretation of some of my data analysis and how to properly discuss them. You helped me despite your busy schedules. I do not think I would have completed my degree if you had not intervened. You were patient, compassionate and critisised my chapters throughly. Thank you very much you were a blessing indeed.

Dr Tshoanelo Miya and Dr Pholoshi Maake, who were a year ahead of me, for their help with some of the genetics programmes and valuable conversations that we had about our

work, and for your friendship. It helped me a lot because I understood that you had been through similar difficulties to those I was going through, and you made it. You got your PhDs and you always encouraged me never to give up.

Mr Mandla Magoro, for being a friend in need. We became close friends in the last year of my write-up and you had to listen to me every day, moaning about how difficult the write-up was. Thank you for listening, and for your words of encouragement.

My three best friends at home (Thabo Moloi, Tshepo Radebe, Sebetseli Makhele). As far away you were, you still made me believe in myself.

Sis Naniswa Nyoka (Mamzo), for your advice and words of encouragement throughout my student life here at SAIAB. You played a mother's role here in Grahamstown, together with Sis Noxi and Sis Mavis. Thank you very much.

Dr Albert Chakona for helping me with the use of some genetics programmes and the interpretation of some of my data analysis. Despite your busy schedule, you still found time to help me when I came to you.

My fellow students and postdocs at SAIAB, for making me feel that I am still at home here at the Institute.

Rhodes psychologist, Mrs Nomangwane Mrwetyana, for helping me resolve the personal problems that were interfering with my PhD. Noma, I needed someone closer to talk to and you were there. Thank you very much!

Mr Garth van Heerden (Human Capital Development Manager). Thank you for seeing throughout that my financial circumstances were tolerable. You also listened to me complain about the difficulties I was in, and advised where possible.

Dr Robert Mckanzie (Proof reader). Thank you for the time and effort you took to proof reading my chapters even though at times I was not giving you enough time, you still came through for me.

Lastly, I would like to thank Dr Angus Paterson (Managing Director of SAIAB) for his words of encouragement. It is true that taking things slowly (one step at a time, like a tortoise) will enable one to reach ones goals, rather than trying to rush through it like a rabbit. Thank you very much; you are the best managing director I know.

This research was supported and funded by SAIAB, Rhodes University (Henderson Scholarship) and the National Research Foundation of South Africa (FA2005021600012). Any opinions, findings, conclusions or recommendations expressed in this material are those of the author and therefore the NRF does not accept any liability in regard thereto.

CHAPTER ONE: GENERAL INTRODUCTION

Overexploitation, water pollution, flow modification, destruction or degradation of habitats, and invasion by non-native species, acting separately and interactively (Dudgeon *et al.*, 2006, Gene, 2007, Leprieur *et al.*, 2009), have resulted in freshwater fishes being among the most imperilled organisms on the planet (Carrizo *et al.*, 2013). South Africa is no exception and the most recent assessment of threats to southern African aquatic ecosystems (Darwall *et al.*, 2009), listed invasive species, pollution, water abstraction and modification of water courses for human use as major threats.

The International Union for Conservation of Nature and Natural Resources (IUCN) Red List of Threatened Species™ provides a peer-reviewed assessment on the conservation status of evaluated species using information on the distribution, habitat preference, taxonomy, conservation priorities and threats in relation to the IUCN Red List Categories and Criteria (IUCN, 2012) in order to assign a threat status (e.g., Extinct, Extinct in the Wild, Critically Endangered, Endangered, Vulnerable, Near Threatened, Least Concern or Data Deficient) to individual species (Darwall *et al.*, 2008). In 2008 the conservation status of a total of 355 southern African freshwater fish species was assessed using the IUCN Red List criteria (Darwall *et al.*, 2008). Of the total number, 12 species were assessed to be Critically Endangered, 19 were Endangered, 9 were Vulnerable, 9 were Near Threatened, 235 were Least Concern and 71 were Data Deficient. Of the 12 species that were evaluated as Critically Endangered, one was a *Labeo* Cuvier 1816 species, *Labeo seeberi* Gilchrist and Thompson 1911 (IUCN, 2016). The *L. seeberi* evaluation was based on severe declines in population sizes resulting

from predation by non-native fishes, as well as from the deterioration of habitat quality (Paxton *et al.*, 2002).

The conservation status of 11 other southern African species of the cyprinid genus *Labeo* were assessed using Red-List criteria in 2008 (Table 1.1). Although, all remaining species are currently evaluated as Least Concern, evaluators listed several threats including the impact of non-native fishes, habitat alterations, man-made barriers to migration and hybridisation with closely related, introduced species (Table 1.1).

Hybridisation as a threat to genetic integrity was listed as a threat to one species, *Labeo umbratus* (Smith 1841). The threat is a consequence of inter-basin water transfer schemes (IBWTs), which have facilitated the introduction of species across a geographic divide between the Orange River system and two southern-flowing river systems, namely the Great Fish River and Sundays River systems (Swartz & Impson, 2007). Given that hybridisation between *L. umbratus* and *Labeo capensis* (Smith 1841) has been observed previously in impoundments (Gaigher & Bloemhof, 1975), this study is intended to contribute to the knowledge base required to better understand this threat to *L. umbratus* populations in the Eastern Cape Province of South Africa, where this *Labeo* species is the largest native primary freshwater fish species. This will be achieved by determining the phylogenetic relationships and evolutionary history of southern African *Labeo* spp.; assessing the phylogeography of *L. umbratus* to assess for regions of conservation importance and evaluating the threat of hybridisation resulting from introductions to develop recommendations for the better conservation of *L. umbratus* genetic diversity.

TABLE 1.1. Southern African *Labeo* spp. and their conservation status. Data taken from the IUCN Red List of Threatened Species™ (IUCN, 2012).

Species	Threats	Conservation status	Reference
L. altivelis Peters 1852	Heavily fished in much of its distribution range.	Least Concern	Bills *et al.* (2010b)
L. ansorgii Boulenger 1907	A dam in the Cunene River separates the populations in the upper and lower reaches of the river.	Least Concern	Da Costa (2007)
L. capensis (Smith 1841)	Industrial pollution on some sections of the Vaal River.	Least Concern	Swartz & Impson (2007)
L. congoro Peters 1852	Habitat and river flow modification, pollution, the use of toxic plants for fishing and overfishing are threats. East African populations are also threatened by water turbidity and sedimentation of spawning beds.	Least Concern	Bayona *et al.* (2010)
L. cylindricus Peters 1852	Habitat degradation by sedimentation as a result of agricultural practices.	Least Concern	Bills *et al.* (2010a)
L. lunatus Jubb 1963	Heavy fishing pressure may affect abundance.	Least Concern	Marshall & Tweddle (2007)
L. molybdinus Du Plessis 1963	Not known.	Least Concern	Bills & Cambray (2007)
L. rosae Steindachner 1894	Susceptible to weir construction, net fishing and sedimentation.	Least Concern	Bills *et al.* (2007)
L. rubromaculatus Gilchrist and Thompson 1913	Inter-basin water transfers may result in invasion of non-native *Labeo* species and possible hybridisation.	Least Concern	Cambray (2007)
L. ruddi Boulenger 1907	Sedimentation and loss of pools in the Kruger National Park, South Africa.	Least Concern	Engelbrecht *et al.* (2007)
L. seeberi Gilchrist and Thompson 1911	Predation by non-native fish centrarchids (*Micropterus dolomieui* Lacepède, 1802, *Micropterus salmoides* and *Lepomis macrochirus* Rafinesque 1819) and possibly *Clarias gariepinus*, fragmented populations (little to no recruitment), decline of habitat quality and number of mature individuals.	Critically Endangered	Lubbe *et al.* (2015)
L. umbratus (Smith 1841)	Industrial pollution on some sections of the Vaal River, invasion by non-native fish and habitat deterioration (weirs) in the Bushmans, Gourits Gamtoos Rivers and hybridisation with *L. capensis* in the Great Fish and Sundays Rivers in South Africa.	Least Concern	Swartz & Impson (2007)

FAMILY CYPRINIDAE

The genus *Labeo* belongs to the family Cyprinidae, which is the largest and freshwater fish family with a wide geographic distribution that includes North America, Africa, Europe, Asia and most of islands of South and Southeast Asia (Howes, 1991; Nelson, 2006). The family comprises about 2991 described species in 11 subfamilies and at least 220 genera (Howes, 1991; Eschmeyer & Fong, 2015). About 326 species have been described since 2006 (Eschmeyer & Fong, 2015). Included among these recently described species are *Labeo fulakariensis* Tshibwabwa, Stiassny & Schelly 2006 from the lower Congo River (Tshibwabwa *et al.*, 2006), five species of *Garra* Hamilton 1822 from Ethiopia (Stiassny & Getahun, 2007), and the redfin species *Pseudobarbus skeltoni* Chakona & Swartz 2013 from the Cape Floristic Region, South Africa (Chakona & Swartz, 2013). In addition, two ongoing international research initiatives are focused on elucidating the diversity and evolution of Cypriniformes fishes: Cypriniformes Tree of Life (CTol) (CTol, 2016) and All Cypriniformes Species inventory (ACSI-2) (ACSI, 2016).

The infrafamilial classification of the Cyprinidae remains highly controversial. The Cyprininae is the largest cyprinid subfamily generally recognised. The Cyprininae is classified into about 11 tribes, which include the Labeonini (Fig. 1.1(a)) (Yang *et al.*, 2015). Rainboth (1991, 1996) formed the tribe Labeonini by combining the 'subfamily' Labeoninae of Chen *et al.* (1984) and the 'labeine cyprinids' group of Reid (1982, 1985). However, the taxonomy of the tribe is undergoing continued refinement (Zhang & Chen, 2004; Stiassny & Getahun, 2007; Yang & Mayden, 2010; Yang *et al.*, 2012; Zheng *et al.*, 2012). The species placed in the Labeonini are distributed in the freshwaters of tropical Africa and Asia (Yang & Mayden, 2010). Most of the species

are adapted to rapid, fast-flowing waters (rheophilic) and have oral modifications (e.g., the terminal or inferior positioning of the mouth, presence or absence of the upper lip and tongue-like or suctorial disc-shaped lower lip of various sizes) and a streamlined body structure suited for such habitats (Yang *et al.*, 2012). Species in this tribe also have a vomero-palatine organ (Reid, 1982), which is a double row of fleshy transverse folds situated on the buccopharyngeal membrane (the roof of the mouth). The organ is used during feeding to co-mix the ingested particles with mucous secretions, but it is in a regressed state in rheophilic aufwuchs scrapers such as *Labeo* and *Garra* (Reid, 1982). The diverse morphology of the mouth and presence of the vomero-palatine organ is a synapomorphy for the Labeonini within the Cyprinidae (Reid, 1982; Yang *et al.*, 2012; Zheng *et al.*, 2012).

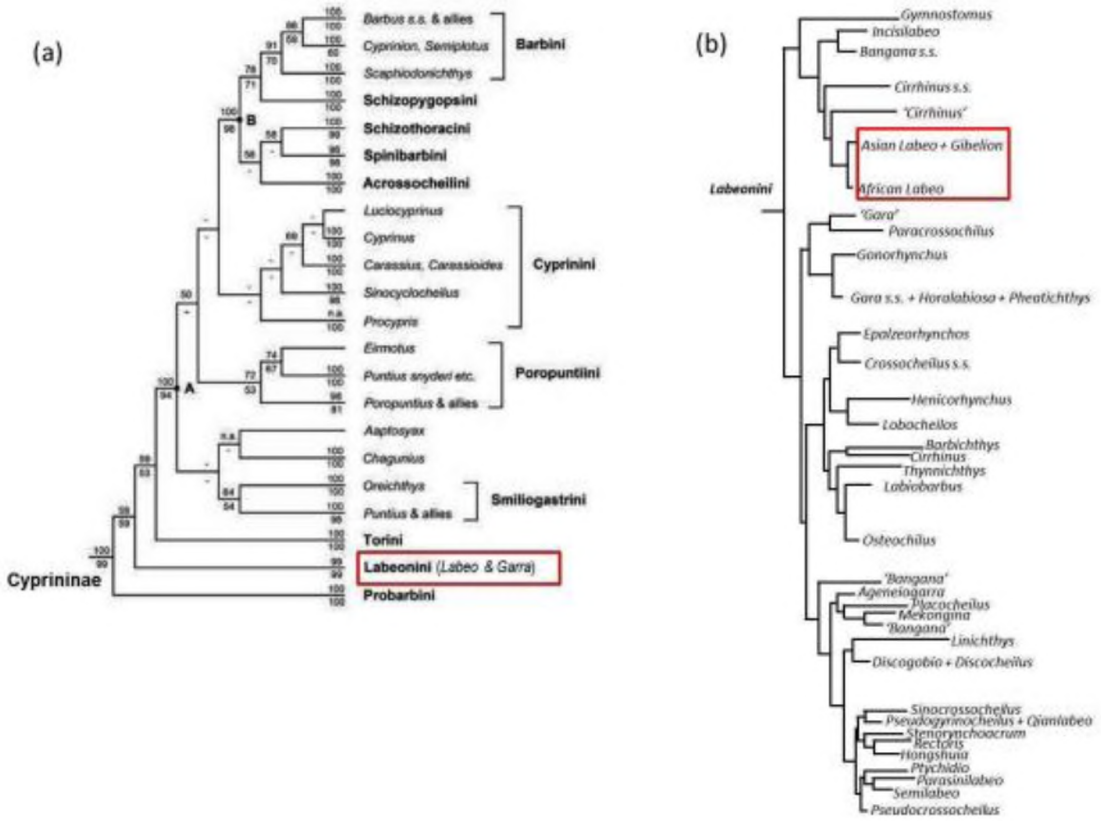


FIG. 1.1. (a) Maximum likelihood phylogram showing relationships among the tribes within the subfamily Cyprininae. The tribe of interest in the current study, the Labeonini, is highlighted within a red rectangle (source: Yang *et al.*, 2015). (b) Maximum likelihood phylogram showing relationships among genera within the tribe Labeonini. The genus of interest in the present study, *Labeo* Cuvier, 1816, is highlighted within a red rectangle (modified from Yang *et al.* 2015).

The tribe Labeonini Bleeker, 1859 [Fig. 1.1(b)] contains about 34 genera and 400 species (Yang & Mayden 2010). The high taxonomic diversity of the tribe is due to the recent description of many new species (e.g., Kottelat, 2000; Su *et al.*, 2000; Vishwanath & Kosygin, 2000; Zhang & Chen, 2002; Su *et al.*, 2003; Kullander & Fang, 2004; Zhang & Chen, 2004; Zhang & Fang, 2005; Zhou *et al.*, 2005; Li *et al.*, 2006; Tshibwabwa *et al.*, 2006; Zhang & Kottelat, 2006; Zhang *et al.*, 2006; Stiassny & Getahun, 2007; Kottelat & Hui, 2008; Li *et al.*, 2008; Krupp & Budd, 2009; Zhang & Zhou, 2012; Arunachalam *et al.*, 2013; Lothongkham *et al.*, 2014; He *et al.*, 2015).

The genera *Labeo* and *Garra* represent almost half of the total number of species in the tribe (Yang & Mayden, 2010). The *Labeo* and *Garra* are predominantly indigenous to Africa, whereas the other genera are distributed in East Asia (China) and Southeast Asia (Yang & Mayden, 2010).

THE GENUS *LABEO*

About 105 valid species of *Labeo* (example Fig. 1.2) are currently recognised. The species are widely distributed throughout the major river systems of Africa, South Asia (India and Sri Lanka) and Southeast Asia (including Taiwan) (Houde & Zastrow, 1993; Froese & Pauly, 2000). At least 72 valid *Labeo* species are indigenous to Africa, which constitute 16.5% of the overall African cyprinid species diversity (Skelton *et al.*, 1991; Houde & Zastrow, 1993; Froese & Pauly, 2000).

Labeo species occupy a variety of habitats (lakes, ponds, swamps, rivers and swift streams) (Reid, 1985) and some species (e.g., *L. capensis* and *L. umbratus*) have the ability to survive under low oxygen (hypoxic) conditions (Hattingh, 1972; Pletzen &

FIG. 1.2. Two representative *Labeo* spp. (*L. umbratus* Smith, 1841 above and *L. capensis* Smith, 1841 below).

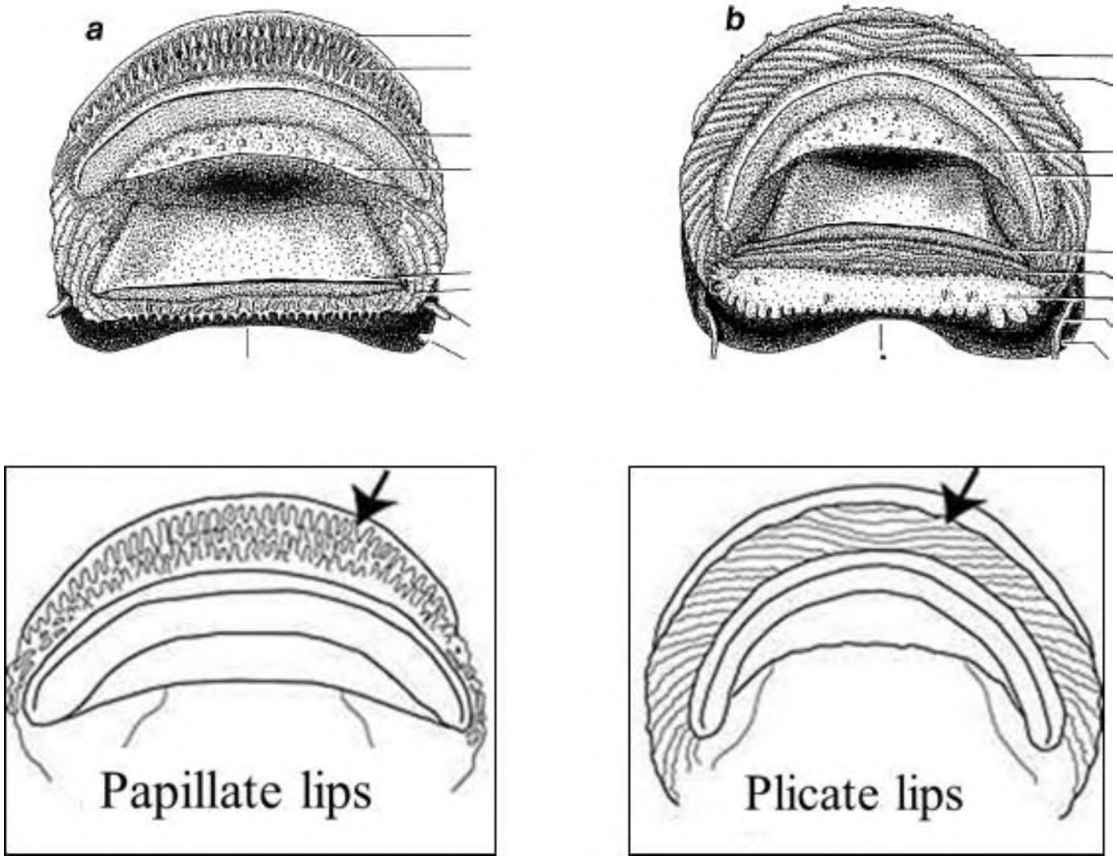
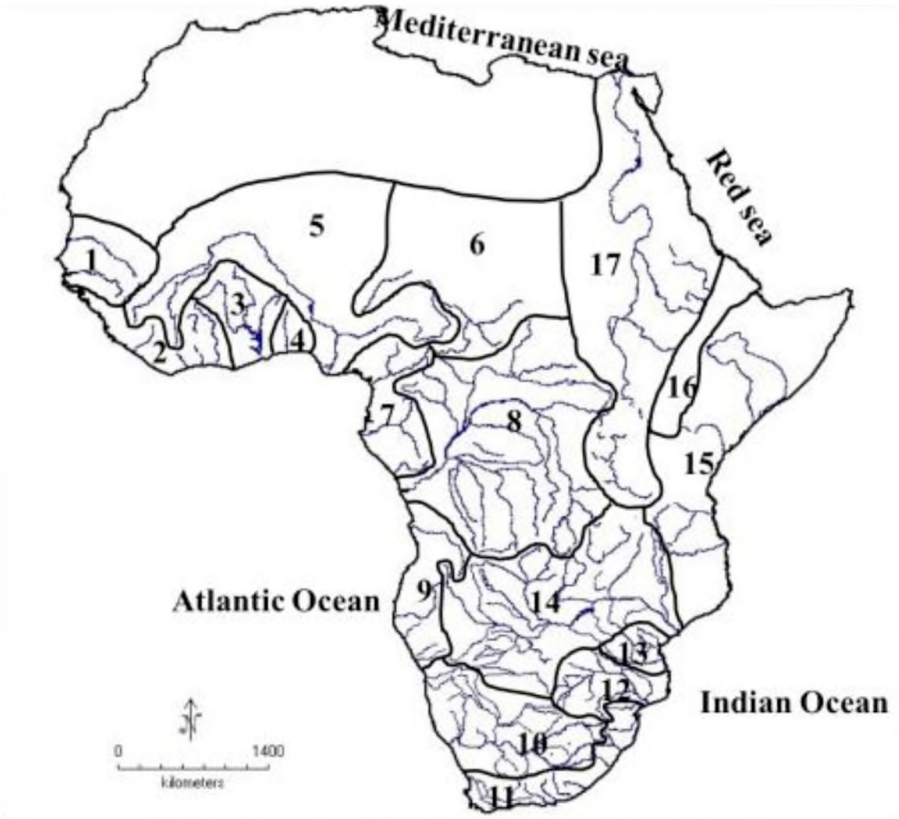


FIG. 1.3. *Labeo* generalised ventral view of the mouth. a) Papillate lips. b) Plicate lips. Images reproduced from Reid (1985).

Hattingh, 1975). Members of the genus *Labeo* have evolved a specialised suctorial mouth (Fig. 1.3) and pharyngeal apparatus used for grazing on benthic algae and aufwuchs on any firm surface (e.g., rocks and woody debris) (Reid, 1985). *Labeo* spp. are caught as a food source in developing countries by commercial and subsistence fisheries, and are occasionally used for angling (Skelton *et al.*, 1991; Ellender *et al.*, 2010).

Labeo spp. are generally polyandrous but differentiation of male and female fish is difficult (Reid, 1985). Most species breed during the rainy season by undergoing lateral migration into the shallow floodplains bordering the riverbeds, but some migrate upstream (Cambray, 1990). Females produce more than 100 000 adhesive eggs, which are scattered (Reid, 1985) on a variety of substrates depending on species. *Labeo umbratus* for example scatters its eggs on flooded vegetation (Jackson & Coetzee, 1982).

Reid (1985) divided the African *Labeo* spp. into six species groups [*Labeo forskalii* (LFG), *Labeo umbratus* (LUG), *Labeo coubie* (LCG), *Labeo macrostoma* (LMG), *Labeo niloticus* (LNG) and *Labeo gregorii* (LGG)] on the basis of morphology and anatomy, but he indicated that these groups may not be monophyletic and that five of the species groups may be more closely related to Asian species than to each other. The LFG is the largest with an estimated 24 valid species and is considered not to be closely related to Asian species. The LNG (12 species) and LCG (15 species) are widespread and almost pan-African (Reid, 1985; Houde & Zastrow, 1993; Lowenstein *et al.*, 2011) (Fig. 1.4).



	Species Group	Distribution on the map
	Umbratus	11, 10
	Forskalii	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17
	Niloticus	7, 8, 9, 12, 13, 14, 15, 17
	Gregorii	15
	Coubie	1, 2, 3, 4, 5, 6, 7, 8, 12, 13, 14, 15, 17
	Macrostoma	2, 6, 7 and 8

FIG. 1.4. Map of Africa showing the major ichthyological regions of inland waters inhabited by *Labeo* spp. and the regional distribution of the African *Labeo* species groups delimited by Reid (1985) (redrawn from Reid, 1985). River Basins: 1) Senegal/Gambia, 2) Western coast (Liberian and Ivory coast), 3) Volta, 4) Benin coast, 5) Niger/Benue, 6) Chad, 7) Cameroon coast, 8) Congo basin, 9) Angolan coast, 10) Orange, 11) Cape, 12) Limpopo, 13) Mozambique coast, 14) Zambezi/Okavango, 15) Eastern coast, 16) Rift valley, 17) Nile. Illustrations of *Labeo* spp. representative of each species group on the left of the table are reproduced from Reid (1985).

The other three species groups show restricted distributions: the LUG (4 species) to southern Africa, the LGG (3 species) to the East Coast, and the LMG (6 species) from Congo to West Africa (Fig. 1.4) (Reid, 1985; Houde & Zastrow, 1993; Lowenstein *et al.*, 2011). Skelton *et al.* (1991) suggested that these species groups may warrant recognition as subgenera or genera.

The validity of Reid's (1985) species groupings have been challenged by several authors (Roberts, 1986; Thys van den Audenaerde, 1987; Tshibwabwa, 1997). Tshibwabwa & Teugels (1995) alternatively grouped the African *Labeo* spp. into two groups based on the anatomy of the inner surface of the lips, which is either papillate or plicate (Fig. 1.3). Lowenstein *et al.* (2011) also revealed that one clade was paraphyletic as it contained species that had both mouth forms. However, the studies by Tshibwabwa & Teugels (1995) and Lowenstein *et al.* (2011) did not include representatives of the southern African species. This represents a notable gap in understanding the phylogenetic relationships among African *Labeo* spp. The inclusion of southern African *Labeo* spp. in molecular phylogenetic analyses is therefore crucial in order to understand the evolutionary history of African *Labeo* spp.

SOUTHERN AFRICAN REGION

Southern Africa is considered to be the portion of the African continent bounded by the Cunene River in the north-west, the Zambezi River in the north-east, and the Cape region in South Africa (Skelton, 2001; Fig. 1.5). Southern Africa is surrounded by oceans on three sides: the Indian Ocean on the eastern and southern coasts, and the Atlantic Ocean on the western coast. The region has a narrow coastal plain in the south that becomes broader in the north-east (Mozambique). Three major river systems have

scoured deep valleys by draining the interior of the region, namely the Orange River, which flows to the west into the Atlantic Ocean, and the Limpopo and Zambezi river systems, which flow into the Indian Ocean. The mouths of these major river systems are separated by shorter succession rivers that drain the coast (Fig. 1.5).

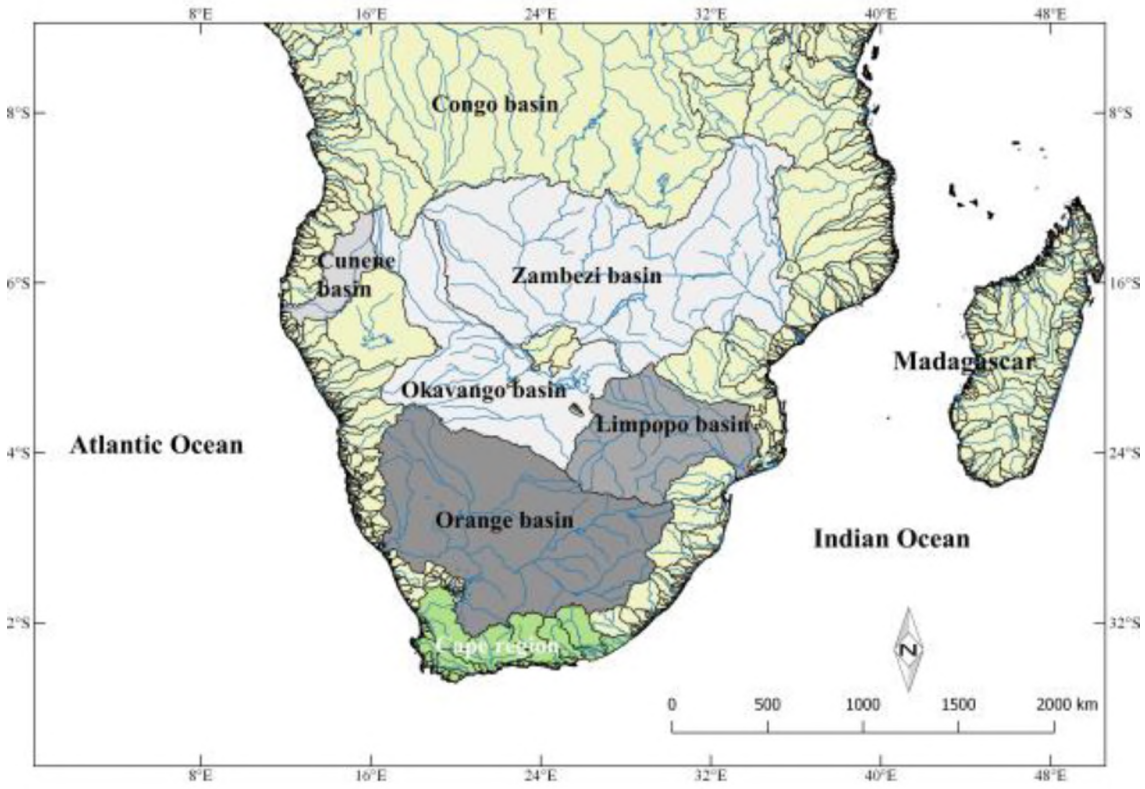


FIG. 1.5. Map of southern Africa showing the location and drainage patterns of the major river systems and the shorter coastal drainage systems.

Southern African *Labeo* spp.

Twelve *Labeo* spp. are indigenous to southern Africa (Skelton, 2001). Four species have been placed in the LUG (*L. umbratus*, *L. capensis*, *L. seeberi* and *L. rubromaculatus* Gilchrist and Thompson, 1913), three in the LNG (*L. altivelis* Peters, 1852, *L. rosae* Steindachner, 1894 and *L. ruddi* Boulenger, 1907), one in the LCG (*L.*

congoro Peters, 1852) and four in the LFG (*L. cylindricus* Peters, 1852, *L. molybdinus* Du Plessis, 1963, *L. lunatus* Jubb, 1963 and *L. ansorgii* Boulenger, 1907). Other than Reid's (1985) morphological and anatomical investigation, the phylogenetic relationships among southern African *Labeo* spp. are poorly understood. Thus, there is a need for a molecular study to: 1) resolve the evolutionary relationships within the genus *Labeo* and 2) allow for the species relationships hypothesised by Reid (1985) to be tested objectively. The focus of this Thesis is on the LUG, which is restricted to the Orange, Tugela and Cape coastal drainages (Skelton, 2001).

The group includes some species, e.g., *L. umbratus* and *L. capensis*, that have been translocated to other river systems, both intentionally (for angling purposes) (Du Plessis, 1963; Skelton, 2001) and unintentionally through inter-basin water transfer schemes (IBWTs) (Cambray & Jubb, 1977). *Labeo umbratus* was translocated intentionally into the Olifants River, which is a tributary of the Limpopo river system (Coetzee *et al.*, 2002). The same species, together with *L. capensis*, has also been translocated to the Great Fish and Sundays river systems via the Orange-Fish and the Cookhouse tunnels, respectively (Cambray & Jubb, 1977; Van Rensburg *et al.*, 2011). *Labeo capensis* were transferred from the upper Vaal catchment into the Tugela River system via an Orange-Thukela IBWT (Karssing, 2008, Van Rensburg *et al.*, 2011). It is also possible that *L. rubromaculatus* reached the Orange river system via the same IBWT as the scheme pumps water 470 m up the Drakensberg Escarpment from the Kilburn Dam into the Driekloof Dam, and at times the water is allowed to return down in order to generate electricity (Department of Water Affairs, Tugela-Vaal project, 1978). Translocated species, in turn, hybridise or compete with the native species (Ramoejane, 2010; Van Rensburg *et al.*, 2011).

Morphologically, the size of the scales of species in the LUG is relatively small compared to those in other species in the genus, the number of lateral-line scales ranges between 42 and 82, and the lowest count of gill rakers (26–43, on the outer margin of the ceratobranchial of the first gill arch) of all *Labeo* spp. groups (Reid, 1985). Most species grow to more than 150 mm (standard length) (Skelton, 2001). The fish migrate upstream during spring and summer to reproduce. *Labeo umbratus* and *L. capensis* are the most-studied species, with greatest research emphasis on the biology, feeding, reproduction and larval development, population structure, migration, parasites and hybridisation (Mulder, 1973; Jackson & Coetzee, 1982; Reid, 1985; Potts *et al.*, 2005). *Labeo umbratus* is naturally distributed across river systems that are currently isolated and could have been isolated for a long enough period to have started to differentiate genetically. However, taxonomic delimitation of such populations is strongly dependent on the species concept that is applied.

SPECIES CONCEPTS

According to Mayr (1982) and De Queiroz (2005) the species is a fundamental unit in biology. However, conceptualisation of the species as a taxonomic unit remains controversial among evolutionary biologists, systematists and ecologists (Mayden, 1999; De Queiroz, 2007). A variety of species concepts have been proposed (Table 1.2). These concepts are incompatible to varying degrees (Mayden, 1997; De Queiroz, 1998), but share an underlying conceptual unity that can be treated as a general or primary concept. The primary defining property of the species category is existence as a separately evolving metapopulation lineage (De Queiroz, 2007). Individual species

concepts also incorporate diverse properties that arise during the speciation process, which may be considered secondary (operational) criteria relevant for assessment of lineage separation (De Queiroz, 2007). All published species concepts have perceived flaws or disadvantages, hence a consensus has not been achieved and different concepts are more compatible with specific methodological approaches. A detailed appraisal of all published species concepts is beyond the scope of the present work, but comparison of the four concepts most frequently applied in ichthyology is pertinent. According to the isolation version of the biological species concept (BSC) the lineage must be intrinsically reproductively isolated from other lineages. Under the ecological species concept (ESC) the lineage occupies a distinct niche. The phenetic species concept (PhSC) prescribes that the lineage is phenetically distinguishable. Under the phylogenetic species concept (PSC; monophyly version) the lineage must be monophyletic in terms of its component genes, taxa or subpopulations (Mayden, 1999; De Queiroz, 2007).

The species concept acts as a guide to lines of evidence relevant to the fundamentally different methodological approaches used to assess the separation of lineages (De Queiroz, 2007). Species concepts can be applied based on the type of questions the investigator seeks to answer and the methodological approach followed (Mayden, 1999). For example, a study that uses a species-level phylogeny to make inferences on historical biogeography might be better served using a species concept that incorporates monophyly (i.e., PSC) (De Queiroz, 2007).

TABLE 1.2. Contemporary species concepts, their abbreviations and properties modified from De Queiroz (2007).

Species concept	Properties	Advocates/references
Biological (BSC)	Interbreeding (natural reproduction resulting in viable and fertile offspring).	Wright (1940); Mayr (1942); Dobzhansky (1950)
Isolation	Intrinsic reproductive isolation (absence of interbreeding between heterospecific organisms based on intrinsic properties, as opposed to extrinsic [geographic] barriers).	Mayr (1942); Dobzhansky (1970)
Recognition	Shared specific mate recognition or fertilisation system (mechanisms by which conspecific organisms or their gametes recognise one another for mating and fertilisation).	Paterson (1985); Masters *et al.* (1987); Lambert & Spencer (1995)
Ecological (ESC)	Occupy the same niche or adaptive zone (all components of the environment with which conspecific organisms interact).	Van Valen (1976); Andersson (1990)
Evolutionary	Unique evolutionary role, tendencies and historical fate.	Simpson (1951); Wiley (1978); Mayden (1997)
Cohesion (CSC)	Phenotypic cohesion (genetic or demographic exchangeability).	Templeton (1989, 1998)
Phylogenetic (PSC)	Heterogeneous (see the following four versions).	
Hennigian	Ancestor becomes extinct when lineage splits.	Hennig (1966); Ridley (1989); Meier & Willmann (2000)
Monophyletic	Monophyly (consisting of an ancestor and all of its descendents; commonly inferred from possession of shared derived character states).	Rosen (1979); Donoghue (1985); Mishler (1985)
Genealogical	Exclusive coalescence of alleles (all alleles of a given gene are descended from a common ancestral allele not shared with those of other species).	Baum & Shaw (1995); Avise & Ball (1990)
Diagnosable	Diagnosability (qualitative, fixed difference).	Nelson & Platnick (1981); Cracraft (1983); Nixon & Wheeler (1990)
Phenetic (PhSC)	Form a phenetic cluster (quantitative difference).	Michener (1970); Sokal & Crovello (1970); Sneath & Sokal (1973)
Genotypic cluster (definition) (GCC)	Form a genotypic cluster (deficits of genetic intermediates; e.g., heterozygotes).	Mallet (1995)
Morphological (MSC)	Similar body shape (morphometrics) and some other structural features (meristics).	Cain (1963)

Some species concepts can only be applied to organisms that reproduce sexually, others to asexual organisms, while others can be applied to both but still with restrictions (Mayden, 1999; De Queiroz, 2007). Species delimitation based on the PSC seems to be superior to the BSC, ESC and PhSC because genetic changes within a lineage occur before morphological and behavioural changes (Wheeler & Meier, 2000).

Some investigators label different populations as "evolutionarily significant units" (ESUs) if the species status is uncertain (Barlow, 2002). The ESU is a concept used when certain populations within a species need to be treated separately for conservation and management purposes as they have been historically isolated from other conspecific populations and are likely to have the potential to become taxonomically distinct in the future (Moritz, 1994). The diverging populations are still at an early stage of speciation and their distinctiveness is measurable in terms of ecological and genetic exchangeability (Crandall *et al.*, 2000). The ESU concept is useful when studying the phylogeography of a species. In the present study the PSC is used for the phylogenetic analysis, ESUs for the phylogeographic analysis, and both PSC and MSC in hybridisation assessments to delimit species.

MOLECULAR SYSTEMATICS

Molecular systematics is the study of patterns and evolutionary relationships of organisms, including identification and delineation of groups of taxa, using genetics and molecular information (Schwartz, 2005). Advances in this field have helped in understanding the underlying evolutionary process and patterns of biodiversity (Barraclough & Nee, 2001; Hammer *et al.*, 2013). Genetic methods such as karyotype

analysis, allozyme electrophoresis, amplified fragment length polymorphisms (AFLPs), mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) sequencing, microsatellites and single nucleotide polymorphisms (SNPs) are used to obtain the raw genetic data (Hammer *et al.*, 2013). Karyotype analysis (karyotype variation) and allozyme electrophoresis (variation within proteins) were the first methods to be used, but now have been largely replaced by DNA-based methods. Two of the DNA-based methods (mtDNA and nDNA sequencing) are mostly used in phylogenetic studies (Patwardhan *et al.*, 2014).

Mitochondrial DNA

Mitochondrial DNA has been used more frequently than other DNA-based methods for phylogenetic and population studies of animals since the 1990s (Hammer *et al.*, 2013; Patwardhan *et al.*, 2014) because the raw genetic data is readily accessed from most organisms owing to the availability of universal and taxon-specific primers. Mitochondria, and consequently mtDNA copies, are present in high numbers within a cell, thus contributing to the ease of detection of amplification products and the cost-efficiency of mtDNA sequencing. Mitochondrial DNA is mostly, if not exclusively, maternally inherited and has been used increasingly in the field of phylogeography (Moritz *et al.*, 1987; Pereira, 2000; Avise, 2009; Guo & Chen, 2010; Hammer *et al.*, 2013). Cytochrome *b* (Cyt *b*) and cytochrome *c* oxidase subunit 1 (COI) are two mtDNA markers that have been used widely in phylogenetic studies (Patwardhan *et al.*, 2014).

The cytochrome *b* gene is fast evolving and is mostly used to elucidate phylogenetic relationships between closely related taxa, as resolution of deeper relationships is lost (Esposti *et al.*, 1993; Patwardhan *et al.*, 2014). Cytochrome *c* oxidase subunit 1 evolves slowly and can be used to infer phylogenetic relationships across a wide range of hierarchical levels (Hebert *et al.*, 2003; Patwardhan *et al.*, 2014). The use of mtDNA in phylogenetic studies is not without problems as mtDNA is subject to incomplete lineage sorting (a problem not unique to mtDNA), does not allow identification of hybrids, and is maternally inherited. These drawbacks make it difficult for investigators using mtDNA to correctly infer phylogenies (Maddison, 1997). The use of both mtDNA and nDNA for phylogeny reconstructions counters these specific problems (Patwardhan *et al.*, 2014).

Nuclear DNA

Nuclear DNA, like mtDNA, is also subject to specific genetic problems (e.g., paralogous gene copies, incomplete lineage sorting and natural selection) (Hammer *et al.*, 2013). Nuclear DNA has been used less widely than mtDNA in molecular systematics because of the lack of suitable primers and their occasional need for gene cloning, and it is less readily amplified because of the low number of gene copies within a cell (Chen *et al.*, 2008; Hammer *et al.*, 2013). However, use of nDNA in molecular systematics has increased recently, partly because of the need to complement mtDNA data (Guo & Chen, 2010). Unlike mtDNA, nDNA is bi-parentally inherited, evolves more slowly than mtDNA, protein-coding genes are partitioned into exons and introns, is generally subject to recombination and is useful for identification of hybrids (Pacheco *et al.*, 2002; Avise, 2009; Hammer *et al.*, 2013). Nuclear introns, according to Guo & Chen (2010), provide an alternative to using coding sequences as genetic

markers. Introns are non-functional and therefore accumulate mutations at a faster rate than coding regions (Friesen, 2000). In addition, introns harbour a much greater degree of genetic polymorphism within and among species than exons (Chow & Hazama, 1998). The latter characteristics have led to the recent increase in the use of introns in population genetic and phylogenetic studies (Friesen, 2000). The different characteristics of nDNA and mtDNA are an indication that they are suitable for different applications within molecular systematics (Hammer *et al.*, 2013). In the present study, mtDNA (Cyt *b* and COI) and nDNA (S7 and Rag1 intron genes) together were used to answer phylogenetic, phylogeographic and hybridisation questions.

Phylogenetic Assessment

Phylogenetics is the study of the evolutionary history of organisms, and typically relationships are represented in the form of a phylogenetic tree (Nei & Kumar, 2000). Phylogenetic inference methods define ancestor and desendant relationships and provide a hierarchical foundation for taxonomic classification (Hou *et al.*, 2007). Phylogenetic trees are an estimate of the evolutionary relationships between individuals or taxa and their hypothetical common ancestor (Nei & Kumar, 2000; Felsenstein, 2004; Hall, 2011). Phylogenies also aims to show the pattern of diversification (Suárez-Díaz & Anaya-Munoz, 2008), which occurs when one ancestral population divides into two or more subpopulations, and which then follow different evolutionary trajectories (Hammer *et al.*, 2013). The subpopulations evolve as a result of accumulation of mutations, natural selection, adaptation and genetic drift (Hammer *et al.*, 2013). According to Hammer *et al.* (2013), populations become distinct species when cross-breeding no longer produce fertile offspring (reproductive isolation). This isolation leads to two types of speciation processes (geographic and ecological speciation)

(Hammer *et al.*, 2013). Geographic speciation is a process under which populations are separated by a barrier (allopatry): physical (e.g., catchment divide), chemical (e.g., seawater), biotic (e.g., predators) and behavioural (e.g., unsuitable habitats and specialisation for a particular habitat) (Losos & Glor, 2003; Hammer *et al.*, 2013). Ecological speciation is the process by which the gene pool is divided within a geographic area of overlap (sympatry) (Rundle & Nosil, 2005; Nosil, 2009; Hammer *et al.*, 2013). The division can be caused by sexual selection, hybridisation, spawning grounds and habitat specialisation (Nosil, 2009; Hammer *et al.*, 2013).

Phylogenetic relationships are mostly inferred using molecular data (Hall, 2013). Molecular phylogenetic analysis involves the use of approaches such as maximum likelihood and Bayesian inference to construct phylogenetic trees, and evolutionary distance and a molecular clock to estimate the timing of speciation (Hall, 2011, 2013). Phylogenetic analytical methods have been applied previously to southern African fishes to address questions about their evolution, divergence times and phylogenetic relationships (e.g. Swartz *et al.*, 2008; Goodier *et al.*, 2011).

Swartz *et al.* (2008) resolved the phylogenetic relationships among members of the genus *Pseudobarbus* Smith 1841, but incongruence between morphological and molecular data was observed, possibly because of convergent evolution and homoplasy in some of the morphological characters. Swartz *et al.* (2008) determined that the earliest divergence among the extant taxa was between *Pseudobarbus quathlambae* (Barnard 1938), which is restricted to Lesotho, and the remaining species of the *Pseudobarbus* genus from the Cape Floristic Region. The Cyt *b* phylogenetic tree of

Goodier *et al.* (2011) revealed five previously unknown *Hydrocynus* Cuvier 1816 lineages. Molecular dating of the phylogeny indicated that estimated divergence times of the 10 lineages based on mtDNA data were consistent with Neogene geological events that modified drainage in Africa (Goodier *et al.*, 2011). The current study will investigate the evolutionary and phylogenetic relationships among southern African *Labeo* spp. using molecular phylogenetic analysis.

Phylogeography

Phylogeography is the field of study concerned with the historical processes governing the contemporary geographical distribution of organisms, especially at the intraspecific level (Avise, 1998; Hickerson *et al.*, 2010). The field was first developed to combine phylogenetic and population genetic studies (Avise *et al.*, 1987), but has evolved as an interactive discipline between all fields of biology (Hickerson *et al.*, 2010). Phylogeographic studies can be based on single or multiple species (for comparative purposes) (Hickerson *et al.*, 2010); in analyses of single species, phylogeography considers multiple populations distributed across the landscape (Avise, 2009).

Phylogeographic analysis initially focused on the use of mtDNA data (Avise, 1998; Hickerson *et al.*, 2010), but subsequently has expanded to include nDNA data (Avise, 1998). Mitochondrial DNA was preferred mostly because of its rapid evolutionary rates and maternal inheritance (Avise, 1998), as this allows detection of mutations at the population level (Avise, 2000). The disadvantage of using only mtDNA is that it represents only a minuscule fraction of the total historical record within a sexual organismal pedigree (Avise, 1998). Much of that history can be retrieved by inclusion

of nDNA, and especially fast-evolving nDNA regions such as introns of the S7 ribosomal protein gene (Morrison *et al.*, 2006). A phylogeographic study of isolated populations of *Crystallaria asprella* (Jordan 1878) in eastern United States rivers (Ohio River Basin, upper Mississippi River, Gulf Coast and lower Mississippi River) utilising mtDNA Cyt *b* sequence data indicated the presence of four distinct populations from the latter United State rivers, with the most divergent population identified within the Ohio River Basin (Elk River) (Wood & Raley, 2000). Morrison *et al.* (2006), using sequence data for the mtDNA control region and nDNA S7 gene for a population genetic analysis of *C. asprella*, observed that population relationships indicated by the control region were consistent with the Cyt *b* data of Wood & Raley (2000), whereas the Elk River population was the only monophyletic population retrieved with the S7 data.

A few phylogeographic studies of southern African fishes have been conducted using mtDNA data (e.g. Swartz *et al.*, 2007; Chakona *et al.*, 2013a), but no phylogeographic study of *Labeo* spp. has been undertaken. A phylogeographic analysis of southern African *Labeo* using both mtDNA and nDNA will contribute to an improved understanding of the history and biodiversity patterns observed among southern African river systems. This study will therefore also contribute to the elucidation of the evolutionary relationships and genetic diversity among the species studied, and will also assist in setting population conservation priorities.

MORPHOLOGY

Morphology is considered to be a science of organismal form (Camardi, 2001). It is a descriptive discipline and its analytical methods or principles can be used as the underlying principles of taxonomy (Camardi, 2001). Morphology was mostly used in traditional taxonomy, but often has proved to be inadequate for reliable phylogenetic reconstruction (Hou *et al.*, 2007). Morphology has provided a limited number of phylogenetically informative characters, especially among organisms with reduced or conserved body forms (Hou *et al.*, 2007). Morphology has been used for over a century in systematics to classify fish diversity by the study of external and internal morphology (Stepien & Kocher, 1997). Counts (for meristic analysis), measurements (for morphometrics) and dissections of fish by systematists have successfully enabled identification of groups of evolutionarily related species (Stepien & Kocher, 1997).

Morphological data also have been used successfully to identify hybrids. Godbout *et al.* (2009), for example, formulated a dichotomous key to identify hybrids between Largemouth Bass [*Micropterus salmoides* (Lacepède 1802)] and Spotted Bass [*Micropterus punctulatus* (Rafinesque 1819)] in the field from morphological characters. It is generally assumed that hybrids are phenotypically intermediate to the two parental species (Smith, 1992). However, this is not always the case and also not true for second-generation hybrids (F_2 hybrids and backcrosses), as such hybrids may be morphologically indistinguishable from one of the parental species owing to dominance (Campton, 1987). Backcrosses or advanced-generation hybrids can also possess novel traits or heterosis (have stronger or bigger traits than both parental species) (Stokes *et al.*, 2007). Morphological characters, in combination with genetic data, enables more reliable detection of hybrids (Stepien & Kocher, 1997).

Behaviour has been found to assist in species delimitation and to resolve taxonomic problems (Barlow, 2002; Stauffer *et al.*, 2002). An example is the diagnoses of the Lake Malawi cichlids. Behaviour as expressed by mate choice based on colour patterns or bowel shape has been used for species discrimination among both rock-dwelling and sand-dwelling cichlid species. However, collection of behavioural data is more difficult and time consuming (Barlow, 2002). This type of data is typically only collected in crucial cases. The disadvantage of behavioural data collection is that the data can mostly only be collected from live specimens and, because colour is an important character for mate choice, data cannot be collected from museum specimens. This highlights the importance of recording photographs or videos of live specimens (Barlow, 2002).

THESIS OUTLINE

The primary objective of this thesis is to contribute towards the knowledge required for developing startegies for better conservation of genetic diversity in *L. umbratus.* To do this, the thesis is divided into five chapters. Chapter 1 introduces the thesis and outlines the purpose of the research, explains important concepts, provides an overview of previous relevant work in the subject area, and provides a contextual background to the research. In Chapter 2 the phylogenetic relationships among the southern African *Labeo* spp. are analysed and the divergence times between species and species groups are estimated. The main aim of this chapter is to reconstruct the phylogenetic relationships among the southern African *Labeo* species to determine: (1) the monophyly of the *Labeo umbratus* group using a molecular approach; (2) the timing of *Labeo* species diversification using divergence-time estimates; (3) whether divergence times between *Labeo* spp. could be used to predict the vulnerability of populations to hybridisation,

which could lead to overall loss of genetic integrity, assuming that recently diverged species hybridise readily. In Chapter 3 the geographic distribution of *Labeo umbratus* genetic lineages is mapped, the evolutionary processes that may be responsible for the contemporary genetic diversity patterns are assessed, and the reconstructed population history is related to known climatic and geological events. The chapter seeks to determine: (1) whether populations of *L. umbratus* constitute historically isolated lineages distributed among different river systems; (2) how long the river systems have been isolated and whether the known drainage history matches genetic structuring in the *Labeo* spp.; and (3) what evolutionary processes have played a role in the genetic differentiation or lack thereof among river systems. Chapter 4 presents a reassessment of hybridisation between *L. umbratus* lineages and *L. capensis*, firstly by assessing whether the two species can be differentiated using 33 morphometric and four meristic characters in conjunction with mtDNA and nDNA data, and secondly by determining if morphological characters can be used to identify hybrids in the field. This chapter attempts to answer the questions: (1) what type of hybridisation (hybridisation without introgression or wide spread hybridisation or complete admixture) is occurring in man-made impoundments; (2) whether morphological and/or genetic data are suitable for identification of hybrids between *L. capensis* and *L. umbratus*; and (3) whether hybrids can be identified reliably in the field. Finally, Chapter 5 presents a synthesis of the overall thesis findings, discussion of the conservation implications and future research opportunities.

CHAPTER TWO: PHYLOGENETIC RELATIONSHIPS OF SOUTHERN AFRICAN *LABEO* SPECIES

INTRODUCTION

Phylogenetic analysis of molecular data has been used to help resolve uncertainties concerning evolutionary species relationships, diversity and taxonomy (Hou *et al.*, 2007; Yang *et al.*, 2012). In recent years fish classification has changed extensively with the aid of molecular systematics, with the description of many new taxa at species and higher taxonomic levels, clarification of species delimitation and synonymisation of previously recognised taxa (Tshibwabwa *et al.*, 2006; Stiassny & Getahun, 2007; Yang *et al.*, 2012; Chakona & Swartz, 2013). More than 100 species of *Labeo* in Africa are recognised, but the monophyly of the genus and interspecific relationships within the genus remain uncertain (Reid, 1985; Yang *et al.*, 2012). Reconstruction of the phylogeny of *Labeo* will aid in understanding the diversity and interrelationships of the African members of this genus. Previous taxonomic (Reid, 1985; Tshibwabwa & Teugels, 1995; Tshibwabwa, 1997; Tshibwabwa *et al.*, 2006) and molecular phylogenetic studies on *Labeo* species from the Congo basin have been undertaken (e.g., Lowenstein *et al.*, 2011). In the present study molecular phylogenetic analyses were performed as part of an assessment of phylogenetic relationships among southern African species of *Labeo*.

Species Groups among Southern African *Labeo*

Of the six species groups within *Labeo* proposed by Reid (1985), four are represented by species in southern Africa, namely the *Labeo forskalii* Group (LFG), *Labeo niloticus* Group (LNG), *Labeo coubie* Group (LCG) and *Labeo umbratus* Group (LUG).

Labeo forskalii group

The *Labeo forskalii* Group (hereafter LFG) comprises more than 40 species and is thus the largest *Labeo* species group (Reid, 1985). Members of the LFG are distinguished from other *Labeo* species groups by the slender, streamlined body form and the large, fleshy snout (Reid, 1985). Morphologically, members of the LFG more closely resemble members of the LCG than any other African or Asian *Labeo* species groups with regard to general oromandibular (mouth, tongue and jaw) characteristics (Reid, 1985).

Four species of the LFG are present in southern Africa (Reid, 1985; Skelton, 2001). These are *Labeo cylindricus* Peters 1852, *Labeo molybdinus* Du Plessis 1963, *Labeo lunatus* Jubb 1963 and *Labeo ansorgii* Boulenger 1907 (Reid, 1985). *Labeo cylindricus* is the most widespread of these species (Fig. 2.1), occurring in the Okavango, Zambezi, Pungwe, Buzi, Save, Limpopo, Incomati and Pongola river systems, in parts of the Congo Basin (Lualba River and Zambian Congo) and in the East African Rift Valley Lakes as far north as Ethiopia (Reid, 1985; Tshibwabwa, 1997; Skelton, 2001; Bills *et al.*, 2010b). In the field, *L. cylindricus* is sometimes confused with *L. molybdinus*, with which it co-occurs.

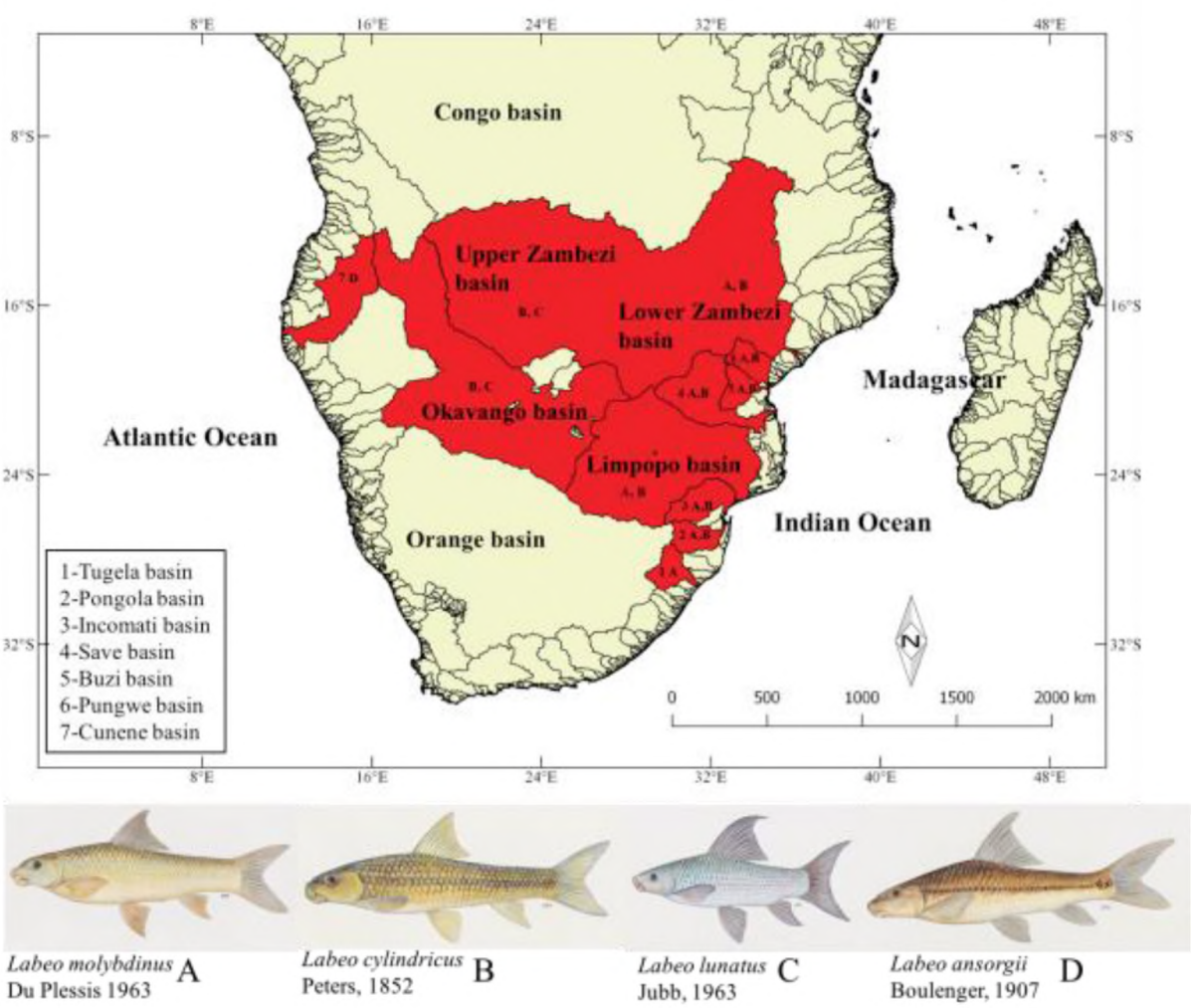


FIG. 2.1. Southern African distribution of the *Labeo forskalii* species group indicated with red colour. Distribution data are based on Skelton (2001). Localities for each species are indicated by alphabets (A-D). Fish illustrations for each species are produced from (Skelton, 2001) with permission from the South African Institute for Aquatic Biodiversity (SAIAB).

However, the two species differ in eye colour (grey in *L. molybdinus* and bright red in *L. cylindricus*) and lateral scale count (Du Plessis, 1963; Reid, 1985). The distribution of *L. molybdinus* extends from the middle Zambezi River south to the Tugela River system (Bills & Cambray, 2007). *Labeo lunatus* closely resembles *L. molybdinus* and *L. forskalii* (Reid, 1985). It occurs in the upper Zambezi and Okavango rivers (Skelton,

2001; Marshall & Tweddle, 2007). According to Reid (1985), *L. lunatus* can be regarded as a form of *L. molybdinus* with a sail-like dorsal fin, as there is otherwise little morphological differentiation between the two species, and a form of *L. molybdinus* with a sail-like dorsal fin has been described from the middle Zambezi and middle Buzi rivers (Jubb, 1961, 1964; Bell-Cross, 1976). *Labeo ansorgii* occurs in West Coast river systems (Bengo, Cunene and Kwanza) and Lake Kilunda in Angola and Namibia (Bell-Cross, 1976; Da Costa, 2007). The species closely resembles *L. forskalii* and *L. cylindricus*, but differences in colour patterns may enable identification of these species (Reid, 1985).

Labeo niloticus group

The *Labeo niloticus* Group (hereafter LNG) is a pan-African species group of nine species that can be distinguished from other *Labeo* species by their small mouths. According to Reid (1985) members of the LNG more closely resemble Asian *Labeo* species, such as *Labeo gonius* (Hamilton 1822), than the other African *Labeo* species. In southern Africa the LNG is represented by three species: *Labeo altivelis* Peters 1852, *Labeo rosae* Steindachner 1894 and *Labeo ruddi* Boulenger 1907 (Reid, 1985; Skelton, 2001). *Labeo rosae* occurs in the Lowveld reaches of the Limpopo, Incomati and Pongola river systems (Bell-Cross, 1976; Bills *et al.*, 2007; Van Steenberge *et al.*, 2014), which flow eastward into the Indian Ocean (Fig. 2.2). There is evidence that this species has been translocated to the Congo Basin based on the re-identification of two museum specimens previously identified as *Labeo mesops* Günther 1868 as (Steenberge *et al.*, 2014).

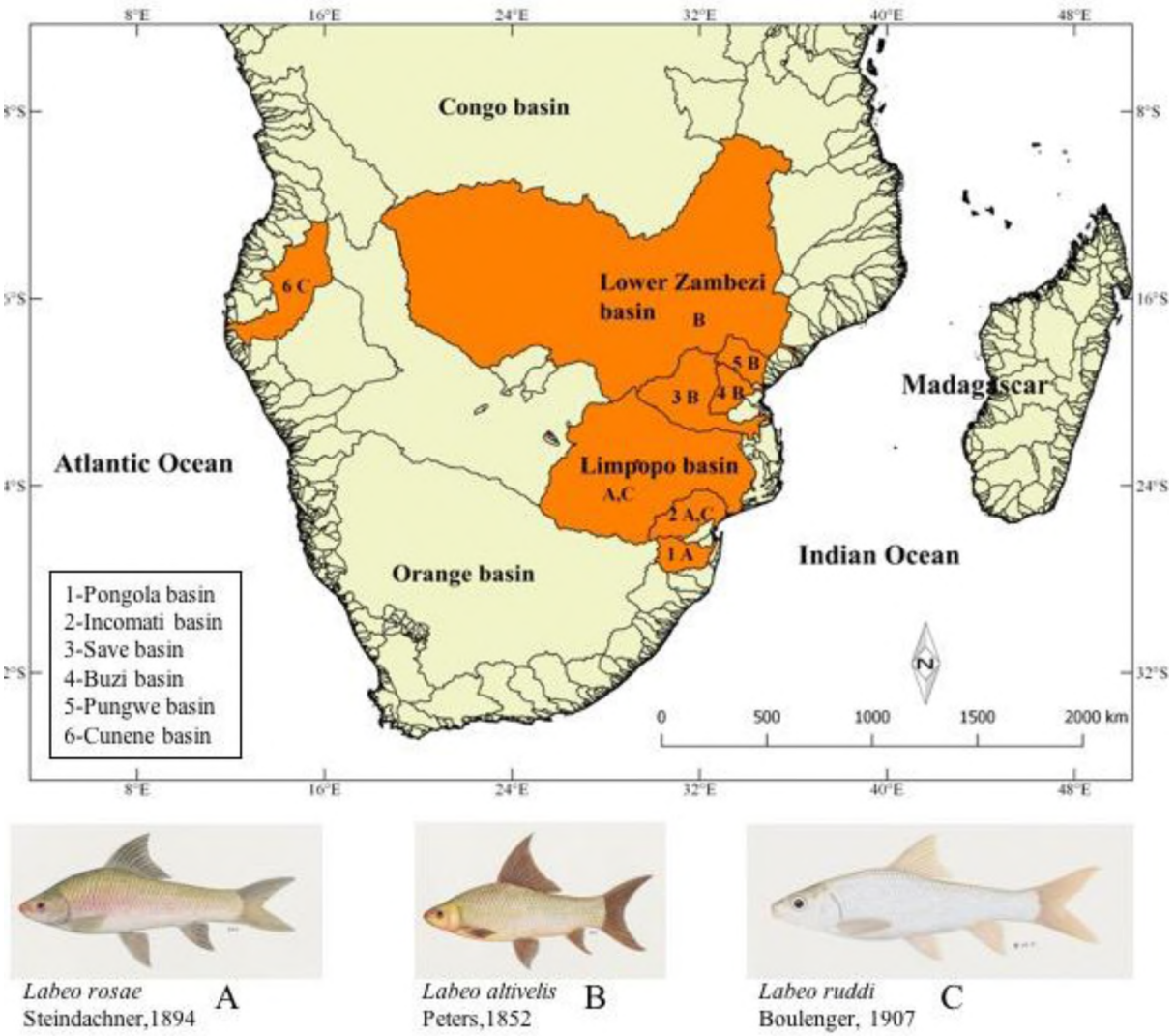


FIG. 2.2. Southern African distribution of the *Labeo niloticus* species group indicated with orange colour. Distribution data are based on Skelton (2001). Localities for each species are indicated by alphabets (A-C). Fish illustrations for each species are produced from (Skelton, 2001) with permission from the ©SAIAB.

The dorsal fin of *Labeo rosae* is very low (DF >25%) and its posterior margin is concave in shape. *Labeo altivelis* specimens from the middle Zambezi River have, on average, a very high dorsal fin (DF>28%), which decreases in size in specimens from river systems southwards towards the Save River (Jubb, 1961; Bell-Cross, 1976).

The posterior margin of the dorsal fin of *L. altivelis* may be convex, straight or concave (Jubb, 1961; Bell-Cross, 1976). *Labeo altivelis* occurs in the Buzi, Save and Zambezi river systems, north of the latter eastern systems (Bell-Cross, 1976; Bills *et al.*, 2010a). *Labeo ruddi* is unique among species in the *L. niloticus* group. *Labeo ruddi* possesses a distinct caudal peduncle spot in preserved sexually mature individuals, it lacks longitudinal stripes, its lateral line is more ventral than dorsal, the gill raker (outer margin of the ceratobranchial of the first gill arch) count is low (43–48) compared with those of *L. rosae* (56) and *L. altivelis* (60), and *L. ruddi* has a short blunt snout (Reid, 1985). *Labeo ruddi* consists of two disjunct populations, one in the Cunene River system on the Angolan–Namibian border, and the second (over 1000 km distant from the Cunene River) in the lower reaches of the Limpopo and Incomati river systems (Engelbrecht *et al.*, 2007; Van Steenberge *et al.*, 2014). This species is absent from the Okavango Delta and Zambezi river systems that separate the two populations (Engelbrecht *et al.*, 2007). Currently, the two populations are considered to be conspecific (Engelbrecht *et al.*, 2007).

Labeo coubie group

The *Labeo coubie* Group (hereafter LCG) comprises eight species, of which only *L. congoro* Peters 1852 occurs in southern Africa. *Labeo congoro* is widespread occurring from Katanga and Lake Mweru in the Democratic Republic of Congo through to the eastward-flowing Malagarasi, Rufiji, Ruaha, Zambezi and Phongolo river systems (Bell-Cross, 1976; Bayona *et al.*, 2010) (Fig. 2.3).

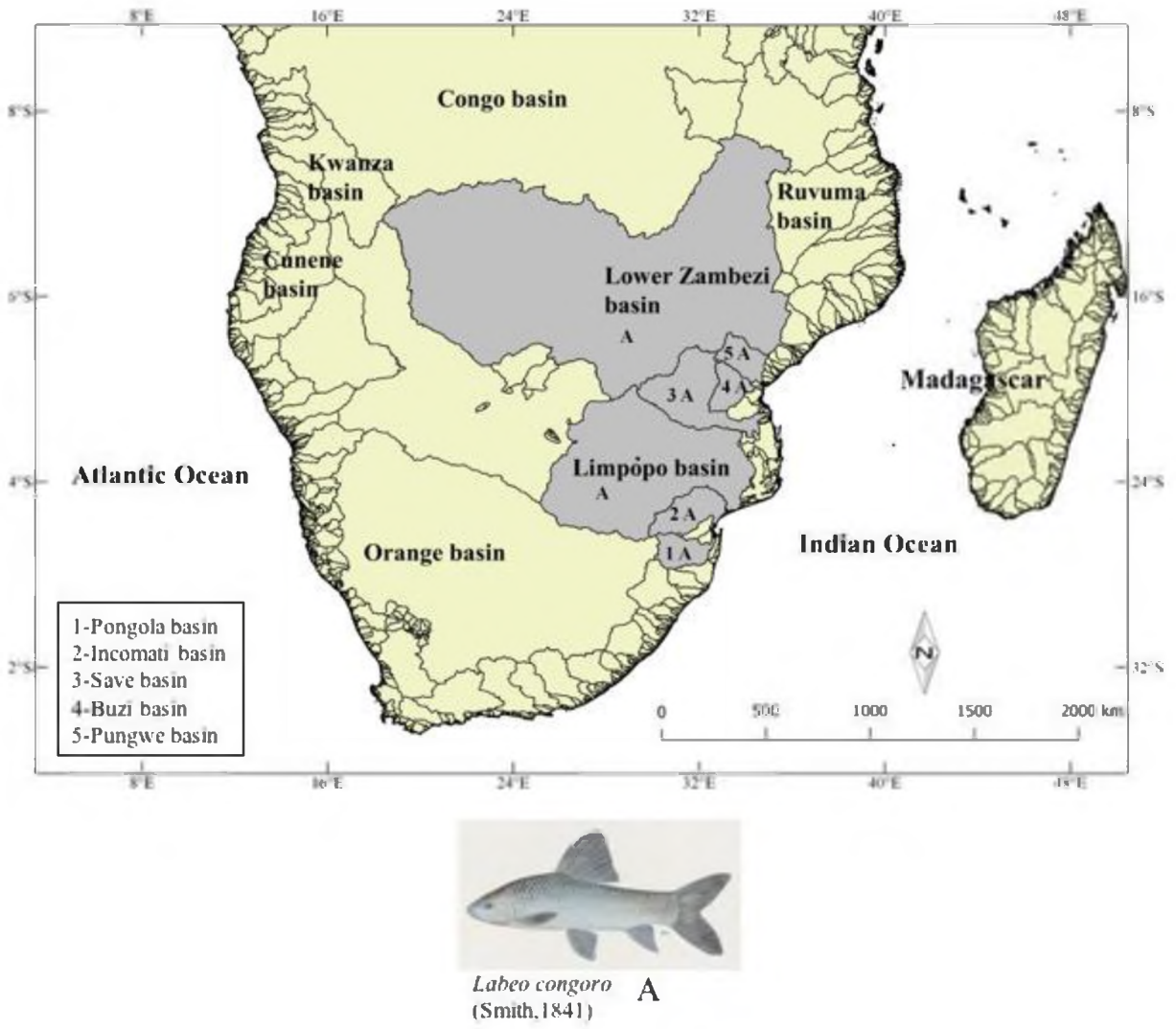


FIG. 2.3. Southern African distribution of *Labeo congoro* indicated with grey colour. Distribution data are based on Skelton (2001). Fish illustrations for each species are produced from (Skelton, 2001) with permission from the ©SAIAB.

Labeo umbratus group

The *Labeo umbratus* Group (hereafter LUG) contains four species: *Labeo capensis* (Smith 1841), *Labeo umbratus* (Smith 1841), *Labeo seeberi* Gilchrist and Thompson 1911 and *Labeo rubromaculatus* Gilchrist and Thompson 1913. The LUG is restricted to southern Africa (Fig. 2.4) and its distribution does not overlap with that of other species groups, except in the Tugela River (Reid, 1985) where *L. rubromaculatus* co-

occurs with *L. molybdinus* of the LFG. *Labeo seeberi* possesses the highest number and the smallest scales of any African *Labeo* species (Reid, 1985). This species occurs in the Olifants River system, which flows westward into the Atlantic Ocean (Reid, 1985; Lubbe *et al*., 2015). *Labeo seeberi* is morphologically similar to *L. umbratus* and *L. capensis* (Reid, 1985). *Labeo umbratus* is more widely distributed than the other species in the LUG, with its distribution extending to the southward-flowing Gourits, Gamtoos, Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon, and Orange river systems (Reid, 1985; Swartz & Impson, 2007).

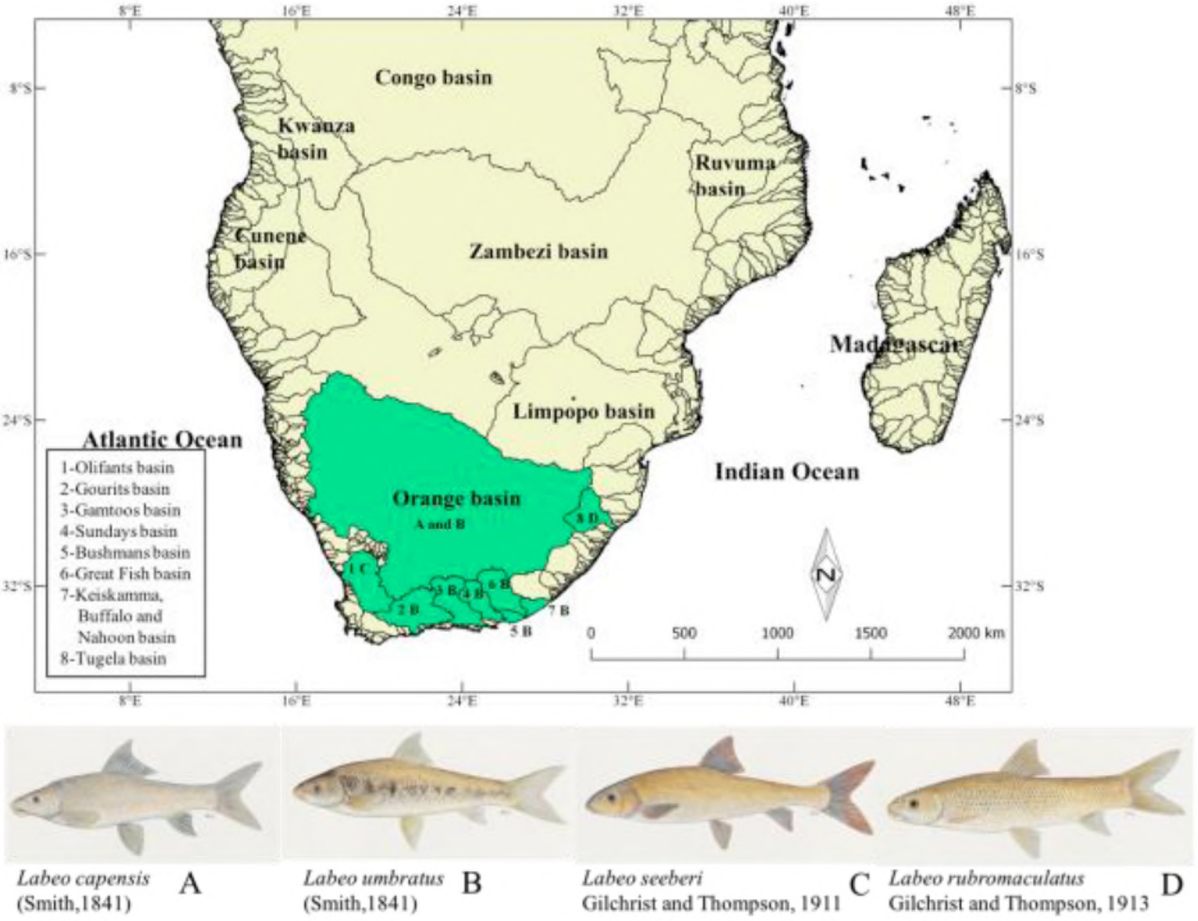


FIG. 2.4. Southern African distribution of the *Labeo umbratus* species group indicated with green colour. Distribution data are based on Skelton (2001). Localities for each species are indicated by different coloured circles. Fish illustrations for each species are produced from (Skelton, 2001) with permission from the ©SAIAB.

In the Orange River system, *L. umbratus* co-occurs with *L. capensis*, from which it differs morphologically by possessing a larger head (HL >24% SL in *L. umbratus* vs <23% SL in *L. capensis*), a higher number of lateral line scales (57 vs 43) and a smaller dorsal fin (DF>22% in *L. capensis* vs <20% in *L. umbratus*) (Reid, 1985). The natural distribution of *L. capensis* is restricted to the Orange River system but the species has been translocated to the Great Fish and Sundays river systems by means of an inter-basin water transfer scheme (Cambray & Jubb, 1977). *Labeo rubromaculatus* occurs in the Tugela River together with *L. molybdinus* from the LFG (Du Plessis, 1963). Reid (1985) suggested that the anatomy of the mouth region of *L. rubromaculatus* is plesiomorphic compared with that of other species within the LUG, but *L. rubromaculatus* is otherwise morphologically similar to *L. umbratus* and *L. capensis*. *Labeo rubromaculatus* is distinct from the other species in the LUG in that its flanks are golden-red in live adult specimens and this species also has a higher gill raker count (43 vs 38–42 in the other LUG species; Reid, 1985).

Members of the LUG more closely resemble the Asian species *Labeo porcellus* (Heckel 1844), in terms of general morphology and small scale size, than they do African species (Reid, 1985). Skelton (1991) suggested that, on the basis of morphological differences and geographical distribution, the LUG could be assigned to a separate genus if, in addition to its morphological and biogeographical distinctiveness and delimited synapomorphies, it proves to be monophyletic (Gill *et al.*, 2005). Smith (1841) described *L. capensis* and *L. umbratus* as members of the genus *Abrostomus* Smith 1841. Smith may have considered that the two species were sufficiently distinct in morphology to represent a separate genus as this was after Cuvier (1816) described the genus *Labeo*. Boulenger (1909) subsequently transferred the two species to *Labeo*.

The latter grouping has been accepted by subsequent researchers (Gilchrist & Thompson, 1913; Barnard, 1943; Du Plessis, 1963; Reid, 1985). In light of Skelton's (1991) suggestion, a reassessment of the generic affinities of the LUG is therefore warranted.

Aims and Objectives

The main aim of this chapter was to reconstruct the phylogenetic relationships between the southern African *Labeo* species to (1) assess the monophyly of the LUG using a molecular phylogenetic approach, (2) estimate the timing of *Labeo* species diversification using a molecular dating approach and (3) determine whether estimated divergence times between *Labeo* spp. could be used as an indicator of vulnerability to interspecific hybridisation.

MATERIALS AND METHODS

Sampling

One hundred and nine samples representing 45 ingroup species and *Gara rufa* Heckel 1843 as an outgroup species, were used for DNA extraction (Table 2.1, 2.2 and 2.3). When available, samples from the South African Institute for Aquatic Biodiversity (SAIAB) National Fish Collection were used (Table 2.1). Samples collected at or close to the type locality of a species were favoured as these samples were assumed to be representative of the holotypes. DNA could not be amplified from *Labeo rosae* specimens lodged in the SAIAB National Fish Collection, probably due to the age of the tissue samples.

In addition to the use of museum specimens, surveys were conducted to collect specimens of *L. capensis*, *L. umbratus*, *L. seeberi*, *L. rubromaculatus*, *L. cylindricus* and *L. molybdinus* (see Fig 2.5 and Table 2.1 for details). Fish were sampled using a variety of methods, including electrofishing, seine netting, gill netting and fyke netting. Collection of fresh specimens of *L. rosae* and from the eastern population of *L. ruddi* was not possible owing to logistical constraints and were therefore not included in the analysis.

TABLE 2.1. *Labeo* spp. samples used for genetic analysis obtained from the South African Institute for Aquatic Biodiversity (SAIAB) National Fish Collection. Species names and groups (*sensu* Reid, 1985) to which they belong, number of specimens, locality and river system where collected, and field identification numbers and SAIAB accession numbers are given. Field IDs are indicated by letters and numerals, SAIAB accession numbers by numerals only.

Species	No.	Locality	River system	SAIAB or Field ID
L. altivelis LNG	1	Mozambique	Buzi	61605
	1	Mozambique	Zambezi	97228
	1	Mozambique	?	AC13A087
L. ansorgii LFG	1	Namibia	Kunene	78470
	1	Namibia	Kunene	78477
L. capensis LUG	1	Vaal Dam	Orange	MR09A071
	1	Vaal Dam	Orange	MR09A072
L. cf. annectens LFG	2	Angola	Lucala	84710
L. cf. mesops LNG	2	Mozambique	Lugenda	73897
L. congoro LCG	1	Mozambique	Muarazi	97324
	1	Mozambique	Muarazi	97064
L. cylindricus LFG	1	Limpopo	Crocodile	PM11A007
	1	Limpopo	Sundays	78512
	1	Limpopo	Groot Letaba	78716
	1	Botswana	Okavango delta	66522
	3	Malawi	Bua	118769
	2	Zambia	Zambezi	72655
	1	Angola	Cuele	186808
	1	Angola	Cuele	186797
L. lunatus LFG	1	Zambia	Kafue	85247
	1	Botswana	Okavango Delta	87197
	1	Botswana	Okavango Delta	87198
L. molybdinus LFG	1	Limpopo	Crocodile	PM11A008
	1	Limpopo	Crocodile	78497

AC= Albert Chakona, MR= Mpho Ramoejane, PM= Pholoshi Maake; Species groups: LFG= *Labeo forskalii* group, LNG, *Labeo niloticus* group, LCG= *Labeo coubie* group, LUG= *Labeo umbratus* group.

TABLE 2.1. continued

Species	No	Locality	River system	SAIAB or field numbers
L. molybdinus LFG	1	Mpumalanga	Incomati	PM11A
	1	Mozambique	Revue	67721
	1	KwaZulu-Natal	Pongola	PM11A037
	1	KwaZulu-Natal	Tugela	187470
L. rubromaculatus LUG	2	Ekuthokozen	Tugela	187470
L. ruddi LNG	1	Namibia	Kunene	78568
	1	Namibia	Kunene	78766
L. seeberi LUG	1	Doring	Olifants	MR11A068
	1	Doring	Olifants	MR11A091
Labeo spp.	2	Angola	Lucala	85101
	1	Angola	Luando	85189
	2	Angola	Kwanza	85508
	1	Angola	Kwanza	85336
	1	Angola	Kwanza	85157
	1	Angola	Kwanza	85370
	1	Central African Republic	Mbourou	77595
	1	Central African Republic	Oubangui	77992
L. umbratus LUG	1	Vaal Dam	Orange	MR09A088
	1	Vaal Dam	Orange	MR09A089
	1	Perdegat Pool near Steytlerville	Gamtoos	MR08J006
	1	Perdegat Pool near Steytlerville	Gamtoos	MR08J007
	1	Stompdrift Dam	Gourits	AC08A015
	1	Stompdrift Dam	Gourits	AC08A016
L. weeksii LNG	1	Angola	Chicapa	99148
G. rufa	2	KwaZulu-Natal Aquarium mortality		186130

Field samples collectors: AC= Albert Chakona, MR= Mpho Ramoejane, PM= Pholoshi Maake; Species groups: LFG= *Labeo forskalii* group, LNG, *Labeo niloticus* group, LMG= *Labeo macrostoma* group, LCG= *Labeo coubie* group, LUG= *Labeo umbratus* group.

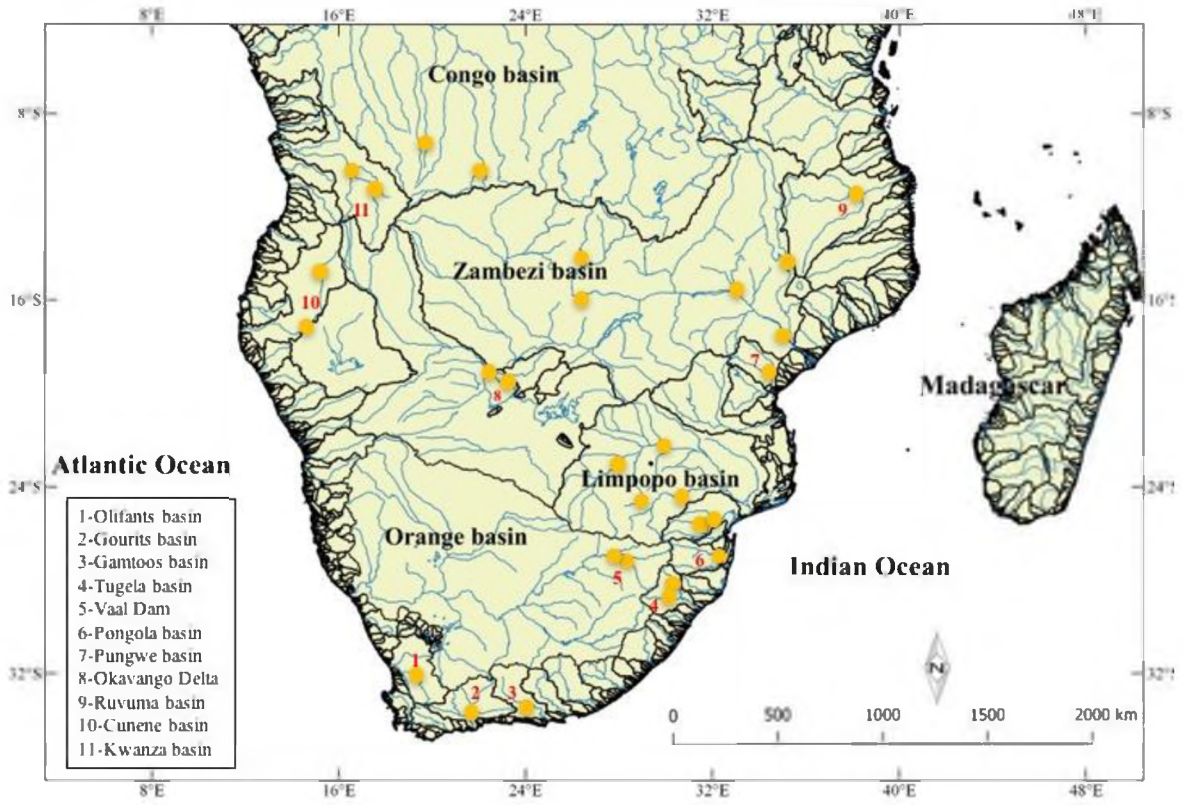


FIG. 2.5. Localities from where all the southern African *Labeo* specimens used for DNA extraction were collected. 1 = Olifants basin, 2 = Gourits basin, 3 = Gamtoos basin, 4 = Tugela basin, 5 = Vaal Dam in Orange basin, 6 = Pongola basin, 7 = Pungwe basin, 8 = Okavango Delta, 9 = Ruvuma basin, 10 = Cunene basin, 11 = Kwanza basin, Zambezi basin, Congo basin . Species link to localities are shown in Table 2.2.

Upon capture, each specimen was killed with an overdose of the anaesthetic eugenol (clove oil) and a portion of muscle tissue or a fin clip was sampled from each individual and preserved in 99% ethanol. The tissue samples were later transferred to a −80 ºC freezer for long-term storage after replacing the ethanol with fresh 99% ethanol. Voucher specimens in the field, were tagged with labels that corresponded to the genetic samples and then fixed in 10% formalin. The body cavity was injected with 10% formalin to improve preservation of the whole specimen.

Samples were subsequently transferred through an ethanol gradient series (to prevent rapid dehydration and to minimise changes in body shape) to a final concentration of 70% ethanol for long-term preservation and donated to the SAIAB National Fish Collection. In addition, sequence data for *Labeo* spp. published by Lowenstein *et al.* (2011) and Yang *et al.* (2012) for the cytochrome *c* oxidase subunit I gene (COI), cytochrome *b* gene (Cyt *b*) and Recombination activating gene 1 (Rag1) were downloaded from the GenBank (https://www.ncbi.nlm.nih.gov/genbank) and the Barcode of Life Database (BOLD; http://www.boldsystems.org) databases, respectively (Tables 2.2 and 2.3). These sequences were included in the data set to explore the phylogenetic relationships of southern African *Labeo* spp. in relation to other African *Labeo* species. The *Labeo* spp. included from the studies by Lowenstein *et al.* (2011) and Yang *et al.* (2012) represent four of the six African species groups (LFG, LCG, LMG and LNG) (Tables 2.2 and 2.3). The only species group that was not represented in this study was the *L. gregorii* group.

The sequences downloaded from BOLD were mislabelled (i.e., had different labels to those specified in the published study by Lowenstein *et al.*, 2011), as were some of their copies in GenBank. This problem was solved by reconstructing the COI maximum likelihood tree using only the sequences from BOLD and comparing that tree with the one constructed by Lowenstein *et al.* (2011). The COI sequences where then labelled accordingly and the corrected COI labels were used to correct those for the Rag1 sequences. Specimens (*L.* aff. *rectipinnis*, *L.* cf. *coubie*, *L.* cf. *cyclorhynchus*, *L.* cf. *maleboensis*, *L.* cf. *parvus*, *L. chariensis*, *L. lukulae*, *L.* sp. nov. and *L. sorex*) that could not be matched to sequences were excluded from the present analysis. Sequence labels were identical to those used in the published paper Lowenstein *et al.* (2011).

TABLE 2.2. *Labeo* species, locality and GenBank accession numbers for DNA sequence data from the study by Yang *et al.* (2012). African species group (*sensu* Reid, 1985) abbreviations are shown following the species names. COI = Cytochrome *c* oxidase subunit I gene, Cyt *b* = Cytochrome *b* gene, Rag1 = Recombination activating gene 1

Species	Locality	GenBank accession no.		
		COI	Cyt *b*	Rag1
L. angra (Hamilton 1822)	Asia	AP011329	AP011329	JX074473
L. barbatulus (Sauvage 1878)	Asia	KC631197	KC631289	KC631222
L. bata (Hamilton 1822)	Asia	JX074181	JX074260	JX074473
L. batesii Boulenger 1911 LMG	Africa	AB238967	AB238967	EU711150
L. calbasu (Hamilton 1822)	Asia	AP012143	AP012143	GQ913472
L. coubie Rüppell 1832 LCG	Benin: Pendjari National Park	JX074182	JX074261	GQ913473
L. chrysophekadion (Bleeker 1850)	Cambodia: Market, Ta Khmau, Kandal	AP011199	AP011199	EU409622
L. cyclorhynchus Boulenger 1899 LMG	Africa	AP011359	AP011359	JX074474
L. dussumieri (Valenciennes 1842)	India	JX074168	JX074250	GQ913453
L. dyocheilus (McClelland 1839)	Asia	JX074183	JX074262	GQ913474
L. forskalii Rüppell 1835 LFG	Ethopia: Alwero River	JX074210	JX074287	JX074491
L. horie Heckel 1847 LNG	Ethopia: Alwero River	JX074211	JX074288	JX074492
L. parvus Boulenger 1902 LFG	Ethopia: Baro River	JX074209	JX074286	JX074490
L. pierrei (Sauvage 1880)	Cambodia: Landing port, Kampong Chhnang	AP011200	AP011200	GQ913475
L. rohita Hamilton 1822	Cambodia: Landing port, Kampong Chhnang	AP011201	AP011201	GQ913476
L. senegalensis Valenciennes 1842 LNG	Benin: Queme and Iguidi Rivers	AB238968	AB238968	EU711151
L. stolizkae Steindachner 1870	China: Ruili, Yannan	GU086536	GU086574	GU086522
L. vulgaris Heckel 1847 LNG	Ethopia: Welkite + Gojeb Rivers tributary of Gibe River	JX074222	JX074298	JX074497
L. weeksii Boulenger 1909 LNG	Africa	JX074184	JX097079	GQ911680
L. yunnanensis Chaudhuri 1911	Asia	JX074205	JX074282	JX074486

Accession number prefixes: AB (DNA data bank of Japan); AP, EU, GQ, GU, JX, KC (Genbank); Species groups: LFG= *Labeo forskalii* group, LNG, *Labeo niloticus* group, LMG= *Labeo macrostoma* group, LCG= *Labeo coubie* group, LUG= *Labeo umbratus* group.

TABLE 2.3. *Labeo* species, locality and BOLD ID code for DNA sequence data from the study by Lowenstein *et al.* (2011). The left code is the BOLD process ID that may be used to view specimen records and GenBank accession numbers, and the right identification code is the AMNH tissue accession. African species group abbreviations are shown following the species names. COI = Cytochrome *c* oxidase subunit I gene, Rag1 = Recombination activating gene 1

Species	Locality	BOLD ID code for COI and Rag1
L. barbatus Boulenger 1898 LCG	Lower Congo River	t-022-2149 AMNHI-385
L. fulakariensis Tshibwabwa, Stiassny & Schelly 2006 LCG	Lower Congo River: Maiko River	t-062-6122 AMNHI-129
	Lower Congo River: Bulu	t-030-2985 AMNHI-407
	Lower Congo River: Mbelo	t-027-2629 AMNHI-408
	Lower Congo River: Lufula River	C08-760 AMNHI-410
L. greenii Boulenger 1902 LMG	Lulua River	t-075-7407 AMNHI-414
	Kisangani: Maiko River	t-062-6121 AMNHI-411
	Lower Congo River: Maiko River	t-062-6123 AMNHI-412
L. lineatus Boulenger 1898 LNG	Lower Congo River: Ntsele River	t-069-6876 AMNHI-212
	Lower Congo River: Mosolo River	t-039-3878 AMNHI-424
	Lower Congo River: Foulakari River	t-020-1994 AMNHI-421
L. lividus Roberts & Stewart 1976 LCG	Lower Congo River: Kinsuka	t-067-6671 AMNHI-206
L. longipinnis Boulenger 1898 LCG	Lower Congo River: Louzi	C08-54 AMNHI-429
	Lower Congo River: Lufula River	t-050-4913 AMNHI-428
L. nasus Boulenger 1899 LFG	Lower Congo River: Mbelo	t-027-2620 AMNHI-430
	Lower Congo River: upstream Luozi	t-033-3231 AMNHI-431
	Lower Congo River: below Bulu	t-032-3147 AMNHI-432
	Lower Congo River: Bulu	t-030-2946 AMNHI-433

t-/C- number= BOLD process ID; AMNHI-number= American Museum of Natural History tissue ID; Species groups: LFG= *Labeo forskalii* group, LNG, *Labeo niloticus* group, LMG= *Labeo macrostoma* group, LCG= *Labeo coubie* group.

TABLE 2.3. (continued)

Species	Locality	Sample code for COI and Rag1
L. nunensis Pellegrin 1929 LFG	Lower Congo River: Cameroon: Ebebda	t-064-6381 AMNHI-161
	Lower Congo River: Cameroon: Ebebda	t-064-6382 AMNHI-162
L. parvus LFG	Lower Congo River: Lulua River	t-055-5405 AMNHI-440
	Lower Congo River: Lulua River	t-055-5450 AMNHI-441
L. quadribarbis Poll & Gosse 1963 LFG	Kisangani: Mpozo River	t-033-3252 AMNHI-447
L. sengaensis Valenciennes 1842 LCG	Cameroon: Ebebda	t-064-6380 AMNHI-449
L. simpsoni Ricardo-Bertram 1943 LFG	Lower Congo River: downstream Luozi	t-052-5119 AMNHI-461
	Lower Congo River: Luozi	t-030-2905 AMNHI-463
L. weeksii LNG	Lower Congo River: Lenga River	CO8-928 AMNHI-474
	Lower Congo River: Nsele River	t-068-6798 AMNHI-475
	Lower Congo River: Nsele River	t-068-6791 AMNHI-468
	Lower Congo River: Ndjili River	t-001-0031 AMNHI-470
	Lower Congo River: Fouta	t-039-3831 AMNHI-471

t-/C- number= BOLD process ID; AMNHI-number= American Museum of Natural History tissue ID; Species groups: LFG= *Labeo forskalii* group, LNG, *Labeo niloticus* group, LCG= *Labeo coubie* group.

Two *Garra rufa* specimens available in the SAIAB collection were selected as an outgroup, because the subtribe Garraina Bleeker 1863 is indicated to be the sister subtribe to subtribe Labeoina Bleeker 1859, in which *Labeo* is classified (Yang *et al.*, 2012).

DNA Extraction, Amplification and Sequencing

Genomic DNA for the 58 samples was extracted using Promega DNA purification kits (Madison, New York, USA). The primer pairs used to amplify two mitochondrial genes (COI and Cyt *b*) and one nuclear gene region (Rag1) from the extracted DNA are listed

in Table 2.4. These genes were preferred because they were used successfully in previous studies to resolve phylogenies of fishes that included *Labeo* spp. (Mayden *et al.*, 2009; Tang *et al.*, 2009; Yang *et al.*, 2012). In addition, the COI gene marker is widely used for genetic barcoding, thus, there are ample sequence data available and it is sufficiently variable to be able to discriminate between two closely related species (Hebert *et al.*, 2003a, 2003b).

TABLE 2.4. List of primers and the respective nucleotide sequences used in the study. Primer designers are listed under Reference.

Primer name	Primer sequence	Reference
Cytochrome *c* oxidase subunit I gene (COI)		
VF2_tl forward	5′-TGT AAA ACG ACG GCC AGT CAA CCA ACC AAG ACATTG GCA C-3′	(Ivanova *et al.*, 2007).
VR1_tl reverse	5′-CAG GAA ACA GCT ATG CTT CTG GGTGGC CAA AGA ATC A-3′	(Ivanova *et al.*, 2007).
Cytochrome *b* gene (Cyt *b*)		
GluF forward	5′-AAC CAC CGT TGT ATT CAA CTA CAA-3′	(Machordom & Doadrio, 2001)
ThrR reverse	5′-ACC TCC GAT CTT CGG ATT ACA AGA CCG-3′	(Machordom & Doadrio, 2001)
Gcyt-Glu forward	5′-GAA AAA CCA CCG TTG TTG TTA TTC A-3′	(Waters & Wallis, 2001)
Gcyt-Thr reverse	5′-CGA CTT CCG GAT TAC AAG ACC-3′	(Waters & Wallis, 2001)
Recombination activating gene 1 (Rag1)		
Rag1-F1	5′-CTG AGC TGC AGT CAG TAC CAT AAG ATG T-3′	(López *et al.*, 2004)
Rag1-R3	5′-GTC TTG TG(CG) AGG TAG TTG GT-3′	(López *et al.*, 2004)

Initially, the GluF forward and ThrR reverse primers were used to amplify the Cyt *b* gene, but subsequently the Gcyt-Glu forward and Gcyt-Thr reverse primer pair was used when the former primer combination failed to amplify the target region. All PCR reactions were performed in a final volume of 25 µl, consisting of 2.5 µl of 1× buffer, 2.5 µl of 2 mM $MgCl_2$, 2.5 µl of 0.2 mM dNTPs, 0.5 µl each 20 mM primer, 0.1 µl of 5 U/µl Taq polymerase, DNA and double-distilled water. The volume of water depended on the amount of DNA used (2–10 µl). The PCR conditions for amplification of each DNA region are listed in Table 2.5.

TABLE 2.5. PCR protocols for amplification of one nuclear and two mitochondrial DNA regions. COI = Cytochrome *c* oxidase subunit I gene, Cyt *b* = Cytochrome *b* gene, Rag1 = Recombination activating gene 1

Gene region	Initial denaturation		Denaturation		Annealing		Extension		Final extension	
			35 cycles							
	Temp	Time	Temp	Time	Temp	Time	Temp	Time	Temp	Time
Cyt *b*	94 °C	2 min	94 °C	45 s	52 °C	60s	72 °C	60 s	72 °C	5 min
COI	95 °C	60 s	94 °C	30 s	54 °C	45s	72 °C	60 s	72 °C	10 min
RAG1	95 °C	2 min	94 °C	30 s	52 °C	30s	72 °C	1 min	72 °C	10 min

The PCR products of the three gene regions were purified with the QIAquick PCR Purification Kit (QIAGEN, Valencia, California, USA), and sequenced (in both forward and reverse directions) using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California, USA). The end products of cycle-sequencing were screened on an ABI 3730xl automated DNA Genetic Analyser (Applied Biosystems) by Macrogen Inc. (Seoul, South Korea), and on an ABI 3100 Genetic Analyser at Rhodes University (Grahamstown, South Africa).

Phylogenetic Analysis

Sequences for COI, Cyt *b* and Rag1 were edited manually in SeqMan (Lasergene v.7.2.0, DNA Star, Inc., Madison, Wisconsin, USA) to check for gaps or uncertain nucleotides. Consensus sequences were obtained from alignment of the forward and reverse sequences to check for discrepancies. For the Rag1 nuclear sequences, nucleotide ambiguity codes were assigned to heterozygotes. DnaSP 5.10 v5 (Librado & Rozas, 2009) was used to phase Rag1 genotypes (identification of the two alleles from the sequenced haplotype) for better identification of hybrids and to identify the unique haplotypes of Cyt *b* and COI, and alleles of Rag1.

The three sequence data sets (COI, Cyt *b* and Rag1) were first analysed separately and then were combined (COI + Cyt *b* + Rag1 and COI + Rag1) to assess the robustness of relationships across data sets. The COI and Rag1 data sets comprised the highest number of *Labeo* species, thus the two regions were combined into a single data set (Table 2.6). The combined sequence data sets were further edited if necessary in Molecular Evolutionary Genetics Analysis (MEGA) version 6 (Tamura *et al.*, 2013). Edited sequence data sets were then aligned further using ClustalX 2.1 (Larkin *et al.*, 2007). Models of nucleotide substitution that best fit each of the five data sets were selected from among 1624 models using the Akaike information criterion (AIC) (Akaike, 1974) with jModelTest 2.1.7 (Guindon & Gascuel, 2003; Darriba *et al.*, 2012). With the same software, base frequencies, Ti:Tv ratio, proportion of invariable sites (I) and the α value of the gamma distribution (rate variation among sites) were estimated for each gene region.

Sequence analysis

The number of taxa, number of base pairs per sequence, and the optimal evolutionary model selected for each of the five data sets are presented in Table 2.6. A maximum likelihood (ML; Felsenstein, 1981) phylogenetic approach was used to reconstruct phylogenetic trees for each gene and the two combined data sets with MEGA version 6 and PAUP* 4.0 beta (Swofford, 2002). Maximum likelihood (ML) is a phylogenetic method that infers an evolutionary tree by finding the tree that maximises the probability of observing the data (Hall, 2011).

TABLE 2.6. Number of taxa, number of bases per sequence and optimal evolutionary model selected using the Akaike information criterion for each of the five data sets

Data set	No. of taxa	No. of bases	Selected models and references
Cyt *b*	67	948	GTR+I+G
			Lanave *et al.* (1984), Tavare (1986), Rodriguez *et al.* (1990)
COI	73	570	TrN+I+G
			Tamura & Nei (1993)
Rag 1	71	974	TrNef+G
			Tamura & Nei (1993)
COI+Rag 1	102	1544	TIM2ef+I+G
			Posada (2004)
COI+Cyt *b*+Rag 1	67	2490	TIM2+I+G
			Posada & Buckley (2004)

GTR, general time reversible; TrN, Tamura & Nei (1993); TIM, transitional model; I, proportion of invariable sites; G, rate heterogeneity; ef, equal base frequencies.

The parameters of the optimal evolutionary model selected with jModelTest were implemented in heuristic searches with the tree-bisection-reconnection (TBR) branch swapping method and 10 random addition replicates were used to find the best ML tree-topology for each of the five data sets.

In addition, Bayesian inference (BI) (Laplace, 1812) was performed with MrBayes v3.0B (Huelsenbeck & Ronquist, 2001), using the evolutionary models selected with jModelTest. This approach was used to estimate the phylogenetic relationships between southern African *Labeo* spp., their placement among the African *Labeo* spp. and their relationships to Asian species.

As with ML, BI is a powerful and widely used method for estimating phylogenetic trees (Hall, 2011). The approach uses the log likelihood as a criterion for choosing among trees (Hall, 2011). One cold and three heated Monte Carlo Markov chains (MCMC) were run simultaneously for 10 million generations. Parameter values and trees were sampled every 2000 generations. The first 20% generations were discarded as "burn-in" to be confident that the MCMC chains were only sampling optimal trees. The likelihood scores for the trees were then examined to determine that stationarity had been attained. The 50% majority-rule consensus tree was calculated from the remaining 4000 trees for each of the analyses from which the posterior probabilities were estimated. These probabilities in percentage are used to estimate the probability that a particular relationship is retrieved. Probabilities ≥ 0.95 are considered significant (Ronquist & Huelsenbeck, 2003).

Genetic distance estimates

Estimates of genetic distance among sequences were analysed with MEGA. The number of nucleotide substitutions per site among sequences was calculated using the models selected by jModelTest. The genetic distances within lineages were calculated for each individual data set and between lineages for the COI+Rag1 and COI+Cyt b+Rag1 combined data sets. The mean genetic distance among species and between sister species within lineages should be less than the average genetic distance among lineages (Hebert *et al.*, 2003a). The genetic distances were used to evaluate differences among species and species groups.

Divergence time estimates

The mitochondrial Cyt *b* data set was used to estimate approximate divergence times between *Labeo* species. Mitochondrial DNA was chosen because it is most often used for estimation of divergence events, thus there are more data for comparison (Galtier *et al.*, 2009; Hedges & Kumar, 2009). Cytochrome *b* was preferred over COI because it is the most used gene for this purpose and the most variable mitochondrial marker with which to estimate divergence times for intra- and interspecific comparisons (Tobe *et al.*, 2010). The nuclear Rag1 intron was not used because it was less variable.

A Bayesian MCMC algorithm, implemented in BEAST 1.7.1 (Drummond *et al.*, 2012), was used to estimate divergence times between lineages. A relaxed molecular clock method that allows a branch-specific rate of variation, drawn from a log-normal distribution, was employed to co-estimate the tree and date the divergences of the lineages under the chosen evolutionary model. Given the lack of a calibrated mutation rate for *Labeo* species, a range of published Cyt *b* substitution rates (slow and fast) of 0.76% and 2.2% per million years was used here (see Chakona *et al.*, 2013b). This range includes the 2% mean rate for vertebrate mtDNA that has been used in a number of studies of fishes for which the true mutation rate is unknown (Brown *et al.*, 1979; Goodier *et al.*, 2011). Tree priors were computed according to a Yule speciation process (Gernhard, 2008) and all other priors were set to the default values. The MCMC analyses were run for 60 million generations and sampled every 2000th generation, with the first 10% of samples discarded as burn-in. TRACER v1.6 (Rambaut *et al.*, 2014) was used to inspect for and confirm stationarity, and that effective sample sizes (ESS > 100) were adequate for all parameters. FigTree v1.4.2 (Rambaut, 2009) was used to build the phylogenetic tree containing the divergence time estimates.

Species delimitation

Species were delimited using the phylogenetic species concept (monophyly), which considers species as a group of organisms that descended from a common ancestor (Rosen, 1979; Donoghue, 1985; Mishler, 1985). The concept is not restrictive in that it allows breeding between different species and takes into account even the slightest genetic difference between organisms (Giraud *et al.*, 2008). Sister species or groups were identified as being the most closely related to a given group or species and share a common ancestor. Divergence of putative unnamed species was compared with that of closely related species delimited by Reid (1985) on the basis of morphological discontinuities. This approach helped to identify groups of individuals that would need further investigation.

RESULTS

Phylogenetic Relationships

Maximum likelihood and BI phylogenies showed similar topologies for all data sets analysed (Cyt *b*, COI, Rag1, COI+Rag1 and COI+Cyt *b*+Rag1), thus the ML phylogenetic tree with Bayesian posterior probabilities (PP) is presented as representative of the phylogenetic relationships resolved from each data set (see Figs. 2.6–2.10). Only significant (PP $\geq$ 0.95) Bayesian posterior probabilities are shown on all trees. The five data sets differed in taxa representation, hence ML trees derived from each data set are presented.

All African species formed a monophyletic well-supported group (PP > 98%) for all data sets except Cyt *b* (Fig. 2.6) and were shown to be distinct from Asian *Labeo* species. This differentiation was supported by mean percentage genetic distances of 1.3% (COI+Rag1) between Asian *Labeo* spp. and the genetically closest African *Labeo* species, *L. batesii* (Table 2.7). Six major African lineages (designated A to F) were revealed, as represented by the COI+Rag1 phylogenetic tree, which was robust (clades were well resolved, nodes were statistically well supported, and the data set had good taxonomic coverage) (Fig. 2.9). The monophyly of the designated lineages mostly received good Bayesian posterior probabilities support (PP > 95%) in all data sets where they were recovered, except for lineage B for Cyt *b* (Fig. 2.6), lineage C for COI+Rag1 (Fig. 2.9), and lineages D and F for Rag1 (Fig. 2.8).

However, the relationships among the lineages were not well resolved, as evidenced by the non-significant posterior probabilities on deeper branches (PP < 95%). The poor resolution of major clades was most evident in the phylogeny derived from the Rag1 data set (Fig. 2.8). The Rag1 sequence data also showed the lowest within-lineage percentage distances (0.4–1.2) (Table 2.8). The southern African *Labeo* species were placed in four lineages (A, B, E and F) in all majority-rule consensus phylogenies retrieved. Most species were monophyletic where the data sets contained more than one individual per taxon (PP > 95%) (Figs. 2.6–2.10).

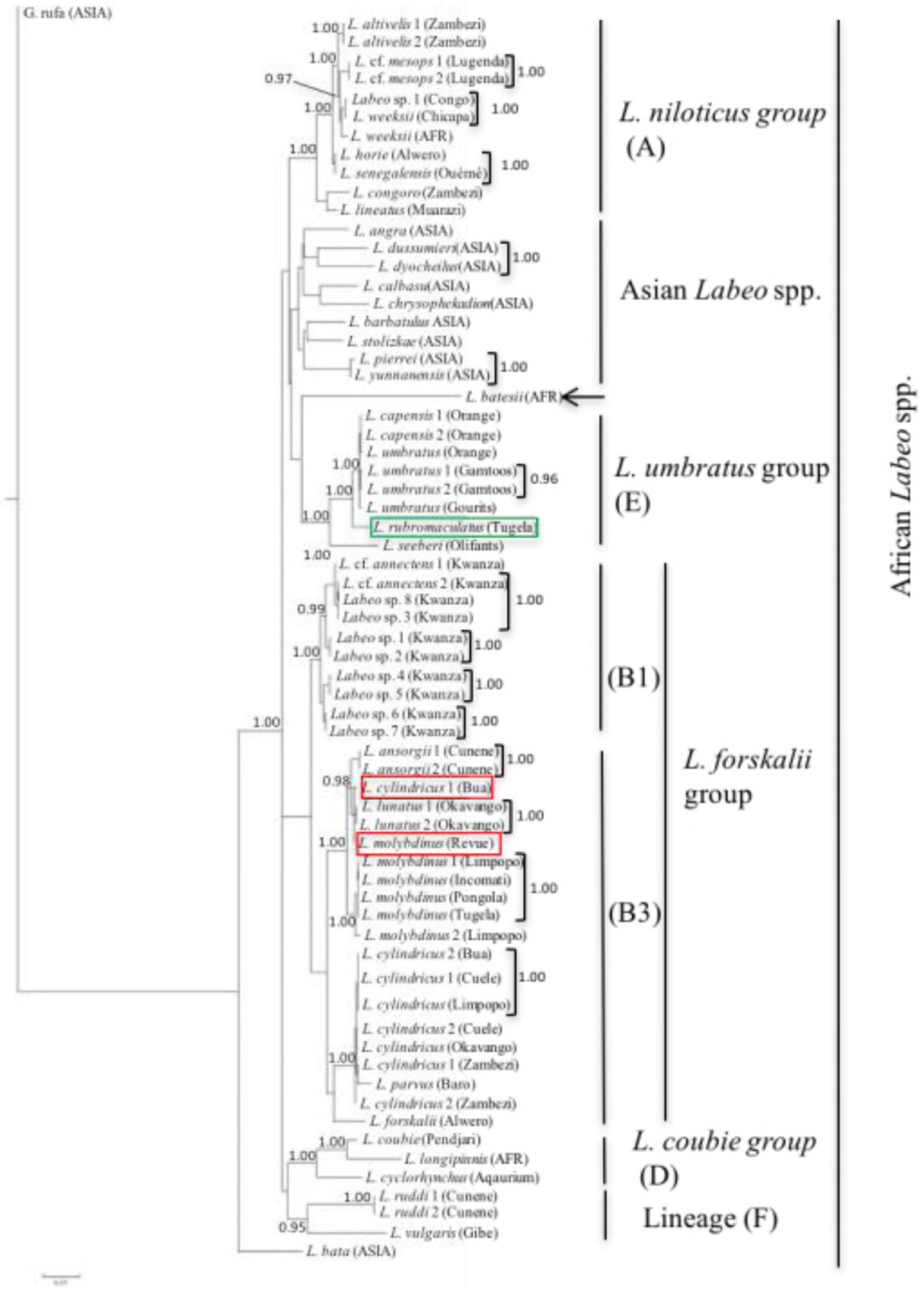


FIG. 2.6. Maximum likelihood phylogenetic tree derived from cytochrome *b* sequence data showing the relationships among southern African *Labeo* species. Values above the branches and next to a square parenthesis are significant Bayesian posterior probabilities ≥ 0.95. Values next to species names indicate the sample number. Lineages are indicated by an upper-case letter in parentheses. Possible hybrids are indicated by a red rectangle; a possible hybrid that was indicated by COI sequence data was not included in the data set due to amplification difficulty. The sample in the green rectangle was identified as a possible hybrid from Rag1 sequence data (see Fig. 2.8).

TABLE 2.7. Mean percentage genetic distances between lineages. Values above the diagonal are for COI+Cyt *b*+Rag1 and those below the diagonal are for COI+Rag1. Lineage A represents the *Labeo niloticus* group, lineage B represents the *L. forskalii* group, lineage C represents *L. macrostoma* group, lineage D represents the *L. coubie* group, lineage E represents the *L. umbratus* group, lineage F and *L. batesii* represent novel lineages, Asian represents Asian *Labeo* spp. and *G. rufa* is an outgroup.

Groups	Lineage A	Lineage B	Lineage C	Lineage D	Lineage E	Lineage F	*L. batesii*	Asian	*G. rufa*
Lineage A		4.0	-	4.5	4.7	4.5	7.3	3.2	13.5
Lineage B	2.5		-	3.9	4.1	3.6	6.7	2.4	13.2
Lineage C	2.5	2.3			-	-	-	-	-
Lineage D	2.9	2.4	2.5		4.6	4.5	7.6	3.1	13.9
Lineage E	3.1	2.2	2.6	2.8		4.5	7.5	3.1	14.1
Lineage F	3.1	2.4	2.6	3.1	2.8		7.2	3.2	13.2
L. batesii	2.2	0.6	2.0	2.1	1.9	2.2		6.3	16.2
Asian	2.6	1.7	2.3	2.1	2.2	2.3	1.3		11.2
G. rufa	11.4	11.0	11.5	11.5	11.1	10.7	10.6	9.5	

Most species were placed in consistent lineages in the phylogenies derived from the different data sets. However, incongruence between data sets was noted for certain species. Specimens identified as *L. rubromaculatus* in the Cyt *b* and COI trees (Figs. 2.6 and 2.7) had nuclear DNA alleles similar to *L. capensis* but with an extra mutational step (Fig. 2.8). Two individuals (*L. cylindricus* from the Bua River and *L. molybdinus* from the Revue River system) clustered with *L. lunatus* in the mtDNA Cyt *b* tree (Fig. 2.6), whereas in the Rag1 tree (Fig. 2.8), for each of *L. cylindricus* and *L. molybdinus* all alleles were clustered together within lineage B. The COI gene could not be amplified for these individuals, thus they are not included in the COI tree (Fig. 2.7). All sequences of possible hybrids were excluded from the data sets used in the combined analyses (Figs. 2.9 and 2.10) except for *Labeo* sp. 2 alleles.

TABLE 2.8. Mean percentage genetic distances within lineages for each of the five data sets.

Groups	COI	Cyt *b*	Rag1	COI+Rag1	COI+Cyt*b*+Rag1
Lineage A	2.4	3.3	0.5	1.5	2.1
Lineage B	5.8	6.3	0.6	2.2	3.2
Lineage C	6.6	–	0.9	3.3	–
Lineage D	5.0	9.3	0.4	2.1	3.3
Lineage E	2.0	3.1	0.6	1.3	2.4
Lineage F	5.3	8.5	1.2	2.3	4.4

Lineage A consisted of species that Reid (1985) grouped into the LNG (*L. altivelis*, *L.* cf. *mesops*, *L. horie*, *L. senegalensis* and *L. weeksii*), and included an unidentified specimen (*Labeo* sp. 1) from the Congo Basin; thus, the lineage was considered to represent the LNG. *Labeo congoro* (a southern African species) and *L. lineatus* were considered by Reid (1985) to belong to the LCG, but were shown here to group with the LNG. The unidentified specimen *Labeo* sp. 1 was resolved to be closely related to *L. weeksii* [PP > 95% for Cyt *b* (Fig. 2.6), COI+Rag1 (Fig. 2.9) and COI+Cyt *b*+Rag1 (Fig. 2.10)]. The *Labeo* sp. haplotype and *L. weeksii* were mostly resolved to be closely related to *L.* cf. *mesops* in certain trees [PP > 95% for Cyt *b* (Fig. 2.6) and COI+Cyt *b*+Rag1 (Fig. 2.10)], but not in the phylogenies derived from the Rag1 (Fig. 2.8) and COI+Rag1 (Fig. 2.9) data sets. *Labeo weeksii*, *L.* cf. *mesops* and *L.* sp. 1 were resolved to be sister to *L. altivelis* (a southern African species) in some trees [PP > 97% for Cyt *b* (Fig. 2.6) and COI+Cyt *b*+Rag1 (Fig. 2.10)], except for the phylogenies derived from the COI (Fig. 2.7), COI+Rag1 (Fig. 2.9) and Rag1 (Fig. 2.8) data sets, which resolved *L. altivelis* as the earliest divergence within the lineage. *Labeo horie* and *L. senegalensis* were resolved as sister species, which was identical to the finding by Yang *et al.* (2012), and were shown to share an identical Rag1 allele (Fig. 2.8).

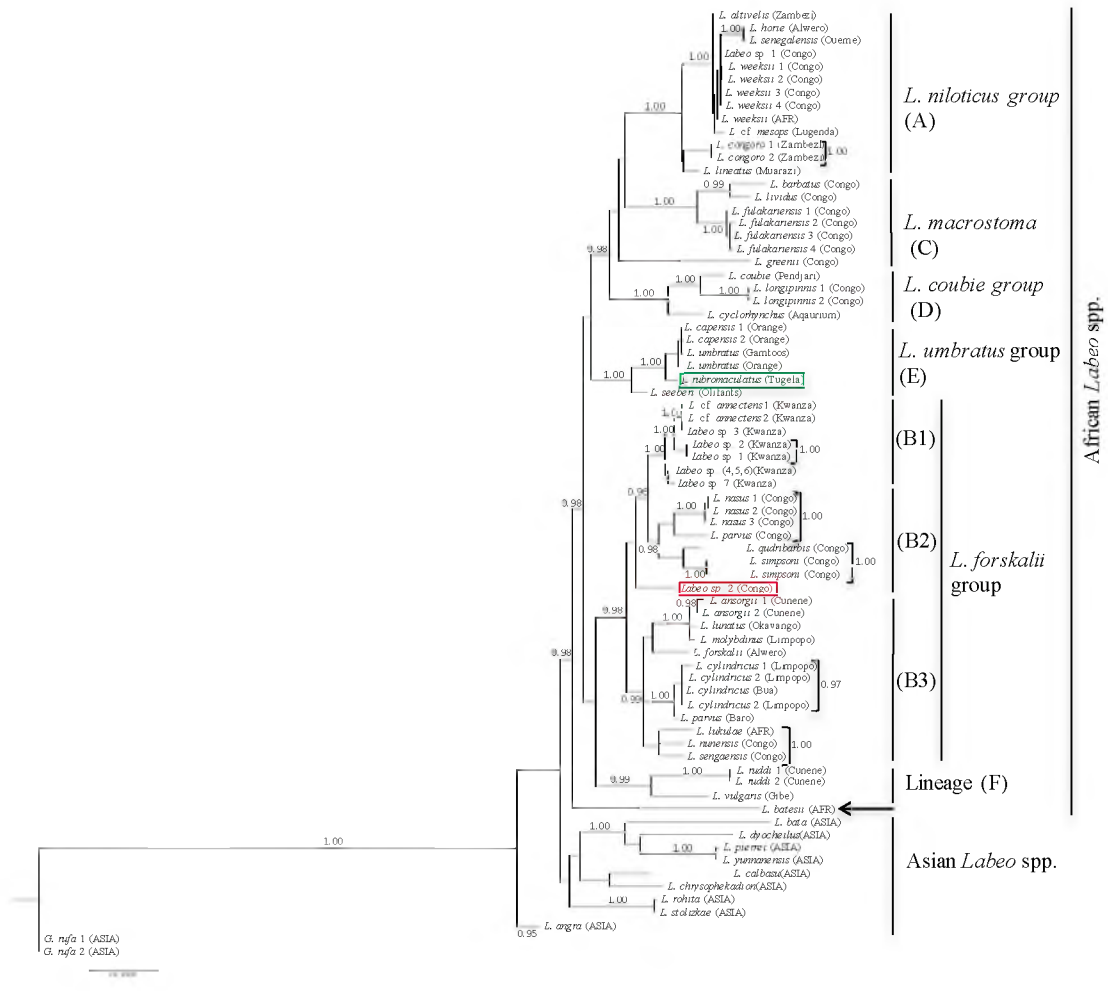


FIG. 2.7. Maximum likelihood phylogenetic tree derived from cytochrome *c* oxidase subunit I sequence data showing relationships among southern African *Labeo* species. Values above the branches and next to a square parenthesis are significant Bayesian posterior probabilities ≥ 0.95. Values next to species names indicate the sample number. Lineages are indicated by an upper-case letter in parentheses. A possible hybrid is indicated by a red rectangle; possible hybrids that were indicated by Cyt *b* sequence data were not included in the data set due to amplification difficulty. The sample indicated by the green rectangle was identified as a possible hybrid from Rag1 sequence data (see Fig 2.8).

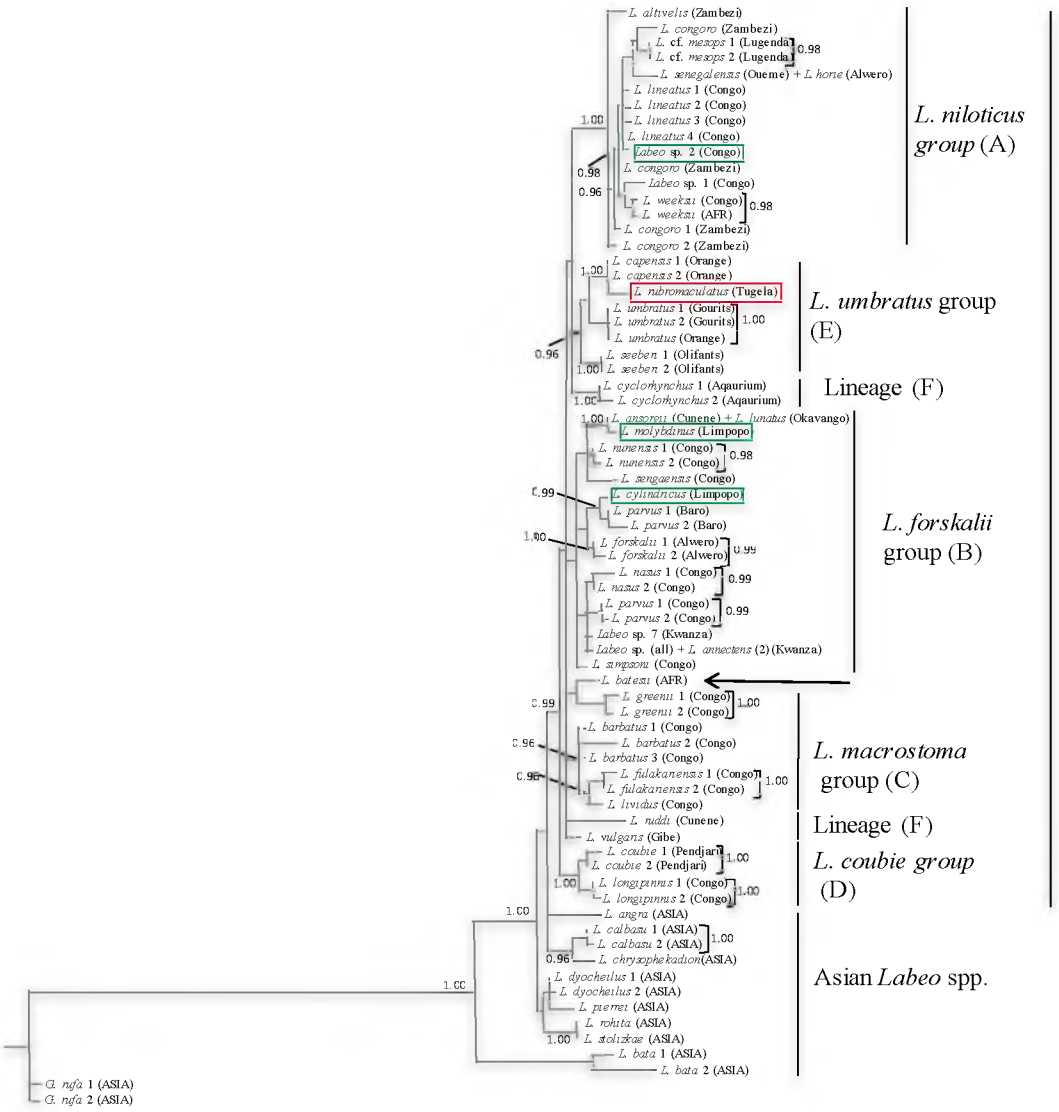


FIG. 2.8. Maximum likelihood phylogenetic tree derived from recombination activating gene 1 sequence data showing relationships between alleles among southern African *Labeo* species. Values above the branches and next to a square parenthesis are significant Bayesian posterior probabilities ≥ 0.95. Values next to a species name indicate the sample number. Lineages are indicated by an upper-case letter in parentheses; their presentation was adopted from COI+Rag1 data. A possible hybrid is indicated by a red rectangle; the alleles indicated by a green rectangle were identified as possible hybrids in trees derived from other data sets (see Figs. 2.6 and 2.7).

The latter two species were resolved to be sister species to the four above-mentioned species (*L. altivelis*, *L.* cf. *mesops*, *Labeo* sp. 1 and *L. weeksii*) in some trees [PP = 100% for Cyt *b* (Fig. 2.6) and COI+Cyt *b*+Rag1 (Fig. 2.10)], but not in the COI (Fig. 2.7), Rag1 (Fig. 2.8) and COI+Rag1 (Fig. 2.9) trees. *Labeo congoro* and *L. lineatus* were resolved as sister species in the COI+Rag tree [PP = 100% (Fig. 2.9)] but not the Rag1 tree (Fig. 2.8). *Labeo congoro* and *L. lineatus* were generally resolved as sister to a clade of six species (*L. altivelis*, *L.*cf. *mesops*, *L. horie*, *Labeo* sp. 1, *L. senegalensis* and *L. weeksii*) [PP = 100% for Cyt *b* (Fig. 2.6), COI (Fig. 2.7) and COI+Rag1 (Fig. 2.9) trees] except in the Rag1 tree (Fig. 2.8). The *Labeo* sp. 2 individual from the Congo Basin was resolved as sister to all species of the LNG in the COI+Rag1 tree (PP = 97%) (Fig. 2.9), but was also nested with species from the LFG [COI (Fig. 2.7)] and with species from the LNG [Rag1 (Fig. 2.8)]. The Cyt *b* gene could not be amplified for this individual and thus was also not included in the combined COI+Cyt *b*+Rag1 analysis.

Lineage B comprised species that Reid (1985) grouped into the LFG (*L. ansorgii*, *L.* cf. *annectens*, *L. cylindricus*, *L. forskalii*, *L. lunatus*, *L. molybdinus*, *L. nasus*, *L. nunensis*, *L. quadribarbis*, *L. parvus*, *L. sengaensis* and *L. simpsoni*) and included unidentified specimens (*Labeo* spp 1–8.) from the Kwanza River system (over all tree). The group therefore was considered to represent the LFG. The genetic distance within this group was slightly higher (0.6–6.3%) compared with that of the LNG (lineage A)(0.5–3.3%) (Table 2.8). The two groups were separated by 4.0% mean genetic distance for all data sets combined and 2.5% for the COI+Rag1 combined data set (Table 2.7). The monophyly of lineage B was well supported (PP > 98%) in all trees except the Cyt *b*

tree (PP < 95%) (Figs. 2.6–2.10). Overall, the lineage could be subdivided into three monophyletic sublineages (B1, B2 and B3). Sublineage B1 comprised a group of closely related unidentified *Labeo* specimens from the Kwanza River in Angola and two individuals of uncertain identity that showed an affinity to *L. annectens* (*L.* cf. *annectens*). The monophyly of sublineage B1 was well supported (PP = 100%) except in the Rag1 tree (Fig. 2.8), in which they formed a monophyletic lineage B with other species alleles that belong to the LFG. The unidentified *Labeo* spp. and *L.* cf. *annectens* individuals could be grouped into five possible forms: (a) *Labeo* sp. 1 and 2, (b) *Labeo* sp. 4 and 5, (c) *Labeo* sp. 6 and 7, (d) *L.* cf. *annectens* 1, and (e) *L.* cf. *annectens* 2 and *Labeo* sp. (3 and 8) [PP > 97% for COI+Rag1 (Fig. 2.9), COI+Cyt *b*+Rag1 (Fig. 2.10)] trees. This arrangement reflects comparison of the sequence divergence of two groups of closely related, recently diverged and morphologically distinct species [*L. capensis* and *L. umbratus* (0.15), *L. horie* and *L. senegalensis* (0.14)] with the forms that showed the lowest sequence divergence [(d) and (e) (0.21)] (Fig. 2.11).

The sublineage B1 was resolved to be sister to sublineage B2 [PP > 95% for the COI (Fig. 2.7) and the COI+Rag1 (Fig. 2.9) trees]. Sublineage B2 consisted of four species from the Congo River system (*L. nasus*, *L. parvus*, *L. quadribarbis* and *L. simpsoni*; PP > 97% in the trees in which all of these species were present) except in the Rag1 tree (Fig. 2.8). The relationship of these four species was consistent with the results of Lowenstein *et al.* (2011). Excluding the potecial hybrid (Labeo sp. 2 Congo), the sublineages B1 and B2 were mostly resolved to be a sister clade to sublineage B3 where all of the sublineages were recovered [PP > 97% for the COI (Fig. 2.7) and the COI+Rag1 (Fig. 2.9) trees].

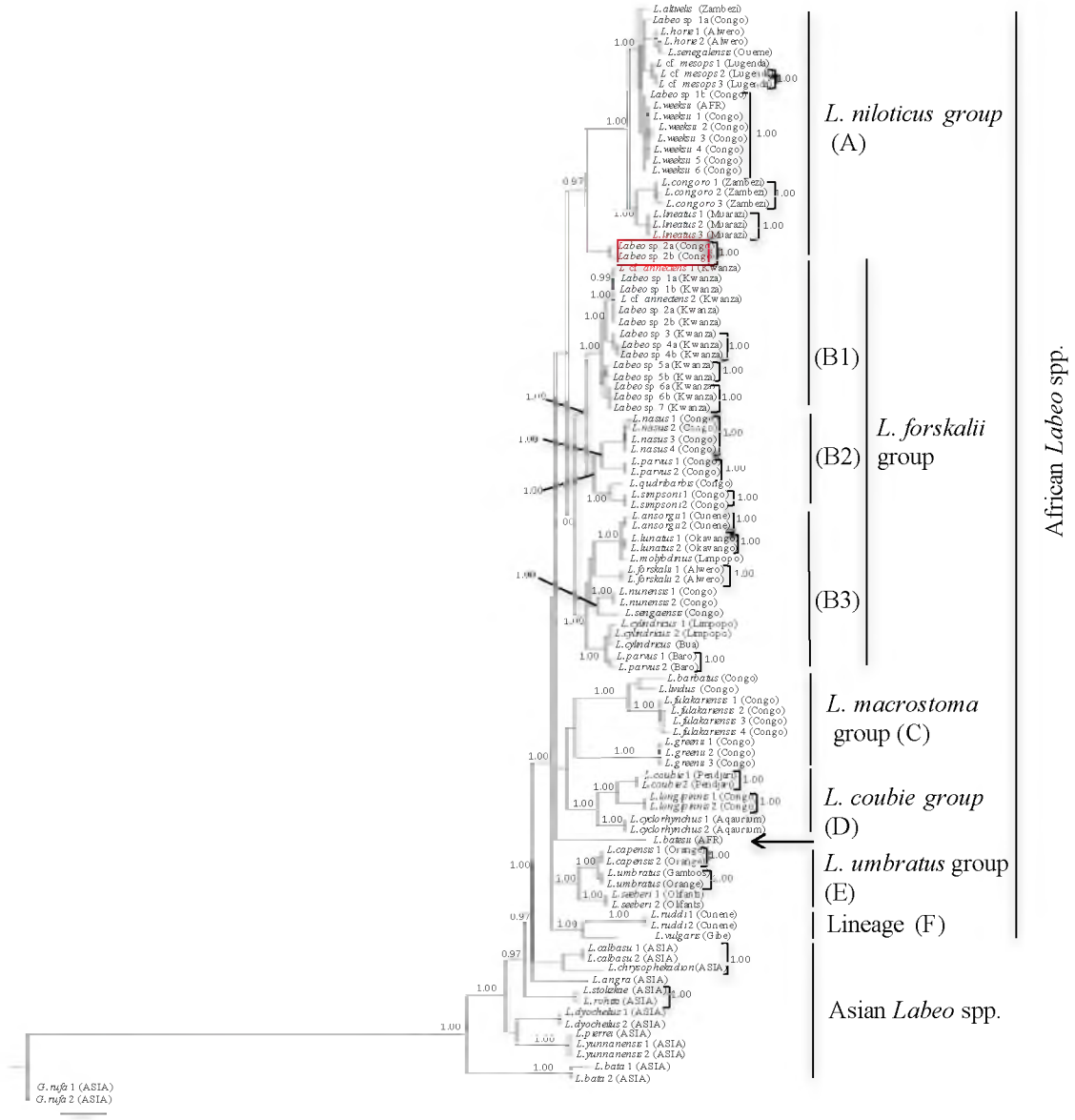


FIG. 2.9. Maximum likelihood phylogenetic tree derived from the combined cytochrome *c* oxidase subunit I gene + recombination activating gene 1 sequence data showing the relationships among southern African *Labeo* species. Values above the branches and next to a square parenthesis are significant Bayesian posterior probabilities ≥ 0.95. Values next to a species name indicate the sample number; numerals followed by a lower-case letter indicate distinct alleles. Lineages are indicated by an upper-case letter in parentheses. Possible hybrids were not included in the data set except *Labeo* sp. 2 alleles (indicated by a red rectangle), which were suggestive of ancient hybridisation between species from the *L. forskalii* group and *L. niloticus* group.

Sublineage B3 consisted of the remaining species placed in the LFG (*L. ansorgii*, *L. cylindricus*, *L. forskalii*, *L. lunatus*, *L. molybdinus*, *L. nunensis*, *L. parvus* (Baro) and *L. sengaensis*). Ignoring potential hybrids [*L. cylindricus* 1 (Bua)], *L. ansorgii* (a southern African species) was resolved to be sister to *L. lunatus* (southern African) with good support [PP > 97% for the Cyt *b* (Fig. 2.6) and the COI+Cyt *b*+Rag1 (Fig. 2.10) trees]. In addition, the two species shared an identical Rag1 allele. Ignoring the potential hybrids [*L. cylindricus* 1 (Bua) and *L. molybdinus* (Revue)],, *L. ansorgii* and *L. lunatus* were mostly resolved to be sister to *L. molybdinus* [PP > 98% for the Cyt *b* (Fig. 2.6), Rag1 (Fig. 2.8) and COI+Cyt *b*+Rag1 (Fig. 2.10) trees] and formed a monophyletic group. Disregarding potential hybrid [*L. molybdinus* (Revue)], *L. molybdinus* individuals from different river systems across its distribution were generally shown to share alleles (over all trees) and all individuals formed a monophyletic group [PP = 100% in the Cyt *b* tree (Fig. 2.6)]. Where recovered, *L. nunensis* was generally resolved as sister to *L. sengaensis* [PP = 100 for the COI+Rag1 tree (Fig. 2.9)], which consistent with the finding of Lowenstein *et al.* (2011), and *L. lukulae* was resolved to be closely related to *L. nunensis* and *L. sengaensis* (PP = 100% for the COI tree). *Labeo parvus* (Baro) was resolved to be sister to *L. cylindricus* [PP > 98% in all trees (Figs. 2.7–2.10)] except in the Cyt *b* tree (Fig. 2.6) (PP = 100) in which it was nested among the *L. cylindricus* alleles. Ignoring the potential hybrid [*L. cylindricus* 1 (Bua)], *L. cylindricus* individuals from different river systems across its distribution were generally shown to share alleles and all individuals formed a monophyletic group [PP = 97% in the COI tree (Fig. 2.7)].

Where included in the data set (Figs. 2.7–2.9), lineage C consisted of three species from the Congo Basin that Reid (1985) grouped into the LMG (*L. barbatus*, *L. lividus* and *L.*

fulakariensis). Thus, the lineage was considered to represent the LMG. The genetic distance within this lineage was low (0.9–6.6%) compared with that of the LCG (0.4–9.3%) (Table 2.8). Lineage C was resolved as sister to the LCG, but with poor support (PP < 95), in the COI+Rag tree (Fig. 2.9)], whereas in the COI tree (Fig. 2.7) it was the sister group to the LNG and in the Rag1 tree (Fig. 2.8) its relationship to other clades was unclear. *Labeo barbatus* was resolved as sister group to *L. lividus* in the COI tree (PP = 99%) (Fig. 2.7), but not in the Rag1 tree (Fig. 2.8). The relationship among the LMG was similar to that reported by Lowenstein *et al.* (2011). Lowenstein *et al.* (2011) retrieved *L. greenii* in this lineage (73% ML bootstrap support), but in the present study the placement of *L. greenii* in this group received poor support [PP < 95% in the COI+Rag tree (Fig. 2.9)] or was placed outside the group [PP < 95% in the COI (Fig. 2.7) and Rag1 (Fig. 2.10) trees].

Lineage D comprised three species considered by Reid (1985) to belong to the LCG (*L. coubie*, *L. cyclorhynchus* and *L. longipinnis*). Thus, this lineage was considered to represent the LCG. *Labeo coubie* was resolved to be sister to *L. longipinnis* where the two species were recovered [PP = 100% in all trees (overall)]. The latter species was resolved to be sister to *L. cyclorhynchus* where the three species were recovered (PP = 100%) in all trees except for the Rag1 tree (Fig. 2.8), in which *L. cyclorhynchus* was not associated with the group, thus rendering the LCG polyphyletic.

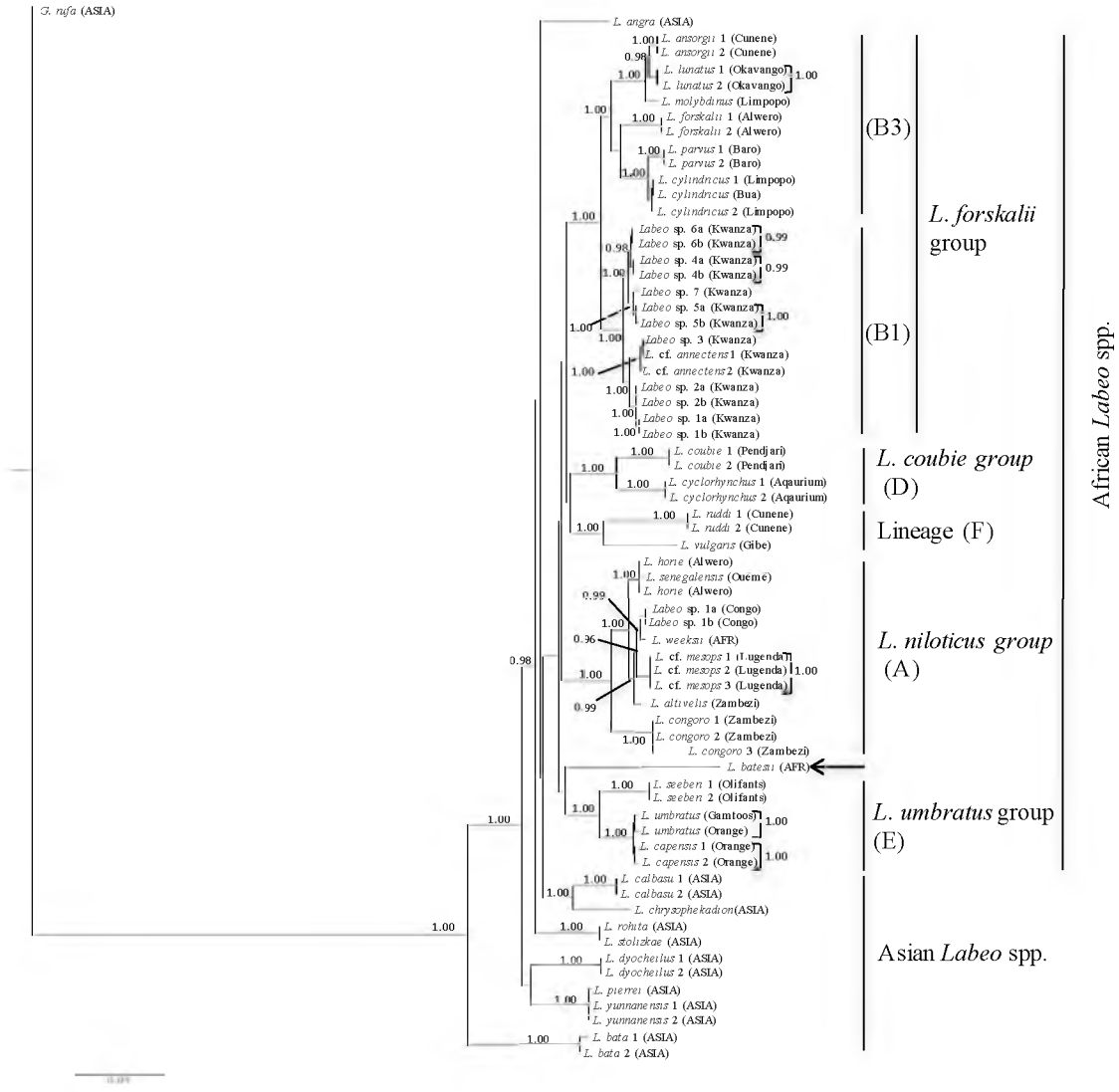


FIG. 2.10. Maximum likelihood phylogenetic tree derived from cytochrome *c* oxidase subunit I gene + cytochrome *b* + recombination activating gene 1 combined data set showing the relationships among southern African *Labeo* species. Values above the branches and next to a square parenthesis are significant Bayesian posterior probabilities ≥ 0.95. Values next to a species names indicate the sample number; numerals followed by a lower-case letter indicate distinct alleles. Lineages are indicated by an upper-case letter in parentheses. Possible hybrids were not included.

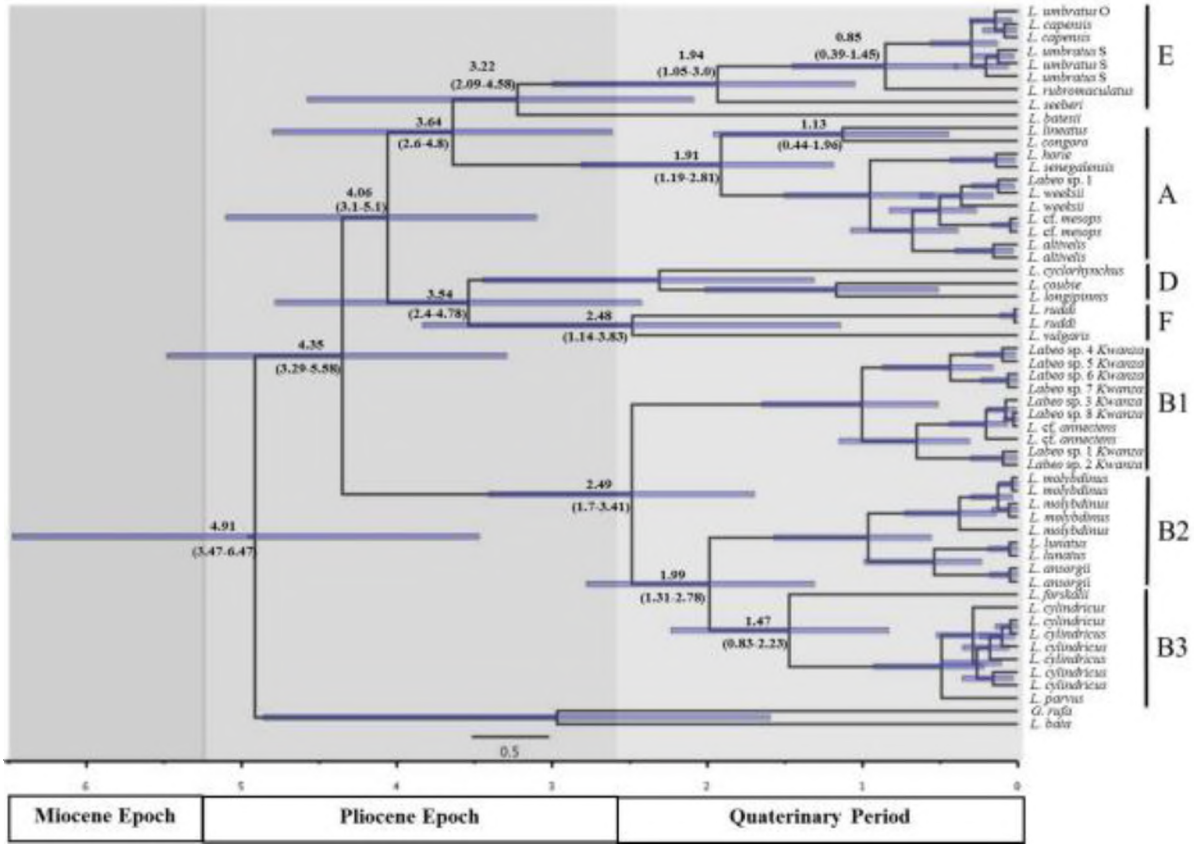


FIG. 2.11. Dated Bayesian tree of the Cyt *b* sequence data of *Labeo* species. The tree was produced with BEAST (Drummond *et al*., 2012) using the GTR parameters specified by jModelTest (Darriba *et al*., 2012). Values above the nodes represent the estimated mean divergence time (million years ago) and 95% highest posterior density ranges are presented in parentheses below each node. Bars represent the 95% confidence interval for the divergence estimates.

Lineage E comprised the four species (*L. capensis*, *L. rubromaculatus*, *L. seeberi* and *L. umbratus*) considered by Reid (1985) to form the southern African LUG. This lineage, together with lineage A, showed the lowest within-lineage genetic distance (0.6–3.1%) (Table 2.8). *Labeo capensis* was generally resolved to be sister to the two *L. umbratus* specimens from the Orange and southern-flowing river systems [PP = 100% in the combined-data trees (Figs. 2.9 and 2.10)]. This relationship is further assessed in Chapter 4. *Labeo capensis* and *L. umbratus* were resolved to be sister to *L. rubromaculatus* (PP = 100%) in the Cyt *b* (Fig. 2.6) and COI (Fig. 2.7) trees, but not

the Rag1 tree (Fig. 2.8). *Labeo seeberi* was resolved to be the sister species to *L. capensis* and *L. umbratus* [PP = 100% in all trees (overall)].

Lineage F consisted of two species (*L. ruddi* and *L. vulgaris*). This lineage and lineage D showed the highest within-lineage genetic distance (1.2–8.5%) (Table 2.8), which is an indication that the two species are highly divergent from each other. The monophyly of the group was well supported [PP ≥ 95% in all trees (Figs. 2.6, 2.7, 2.9 and 2.10)], except for the Rag1 tree (Fig. 2.8) in which lineage F was paraphyletic.

Divergence Time Estimates

Using Cyt *b* data, all of the six groups of *Labeo* spp. were estimated to have diverged during the Pliocene epoch (1.8–5.3 million years ago [mya]) (Fig. 2.11). The LFG (lineage B) diverged from the remainder of the African lineages around 4.35 mya (95% highest posterior density [HPD]: 3.29–5.58 mya), followed by the split between lineage F and the LCG (lineage D) and LNG (lineage A) + LUG (lineage E) + *L. batesii* around 4.06 mya (HPD: 3.1–5.1). The split between the LNG and the LUG + *L. batesii* was estimated at 3.64 mya (HPD: 2.6–4.8), followed by a split between the LUG and *L. batesii* around 3.22 mya (HPD: 2.09–4.58). The split between the LCG and lineage F happened around 3.54 mya (HPD: 2.4–4.78). The split within the LFG between sublineages B1 and B2 + B3 was estimated at around 2.49 mya (HPD: 1.7–3.4 mya). The split between sublineages B2 and B3 was estimated at about 1.99 mya (HPD: 1.31–2.78 mya), which also fell within the Pliocene epoch. The divergence between most species within lineages was estimated to have occurred during the Holocene period (0.01–1.8 mya).

DISCUSSION

The current molecular-based analysis used new sequence data for southern African *Labeo* spp. together with sequences from Lowenstein *et al.* (2011) that could be confidently assigned to species and the *Labeo* sequence data from Yang *et al.* (2012). In concordance with Reid (1985), the resulting ML phylogenetic trees (Figs. 2.6–2.10) resolved six *Labeo* lineages plus a divergent species identified as *L. batesii*. Five of the lineages (A, B, C, D and E) represented the five African groups (LNG, LFG, LMG, LCG and LUG, respectively) proposed by Reid (1985). The deeper relationships between the different African *Labeo* groups were unresolved, which might be the result of gene saturation from which mtDNA suffers (Ho *et al.*, 2005).

The present results also supported Tshibwabwa & Teugels (1995) and Lowenstein *et al.* (2011) in dividing the African *Labeo* spp. on the basis of the morphology of the inner surface of the lips (Fig. 2.12). The LFG and LCG have the papillate mouth form, whereas the LNG have a plicate mouth form (Tshibwabwa & Teugels, 1995; Lowenstein *et al.*, 2011). The results from the present study supported the latter finding on the LFG, but in addition LUG and lineage F possess plicate mouth forms. The conclusion by Lowenstein *et al.* (2011) that the plicate group is paraphyletic (the LCG clustered together with the papillate group) and the papillate group is polyphyletic was supported by the topology of the dated BI tree (Fig. 2.11). The BI tree showed that the LFG group diverged from the ancestor of the remaining African species groups around 4.35 mya. Lowenstein *et al.* (2011) provisionally found that African *Labeo* spp. were monophyletic (see Chapter 1). The results of all phylogenetic analyses presented in this chapter were concordant with the Lowenstein *et al.* (2001) study except for the Cyt *b* data (Fig. 2.6), which placed most of the Asian *Labeo* spp. sister to the LUG and *L. batesii*.

This latter finding, however, also might be the result of mtDNA saturation because other phylogenetic reconstructions presented in this chapter and previous studies (Lowenstein *et al.*, 2011; Yang *et al.*, 2012) placed the Asian lineage as distinct from the African lineage. Cytochrome *b* is the most variable gene compared with COI and Rag1, yet the deeper branches were poorly supported.

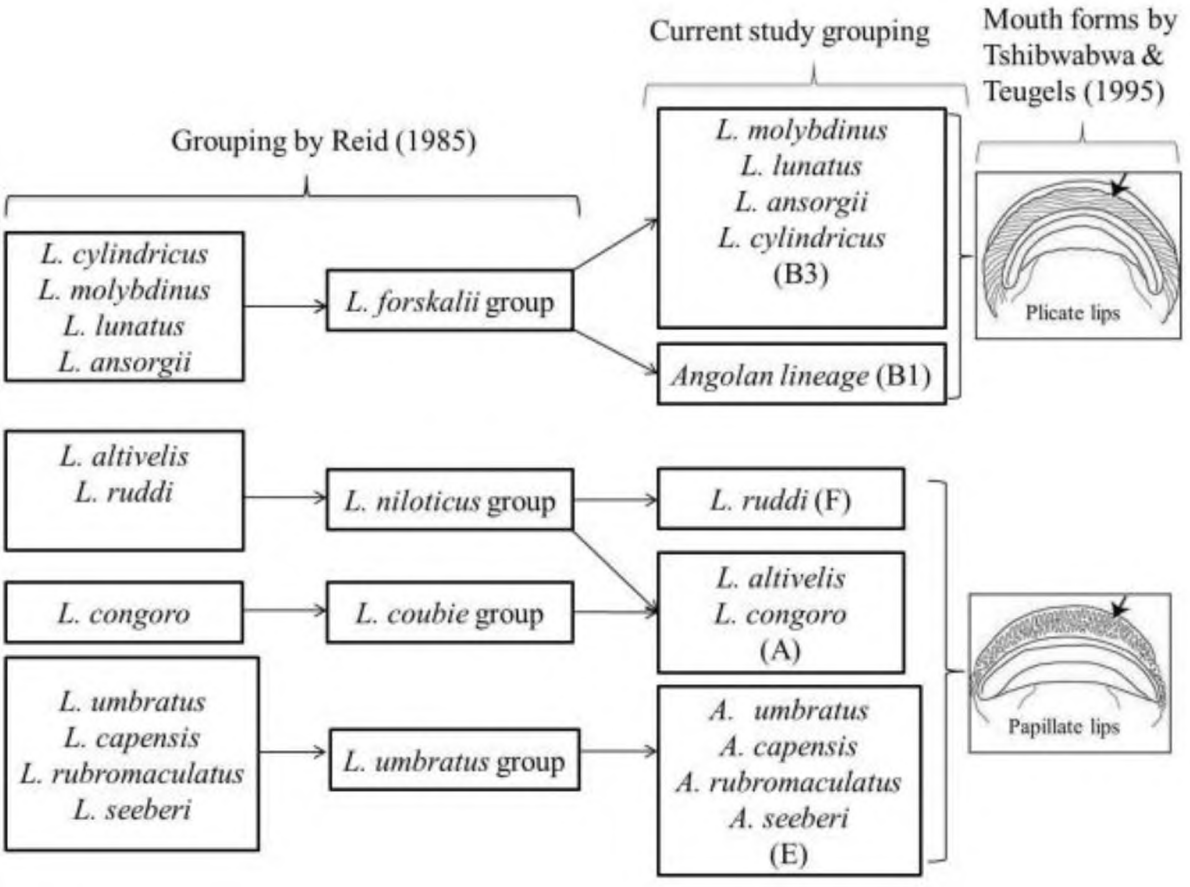


FIG. 2.12. Correspondence of *Labeo* species groups proposed by Reid (1985) for southern African species, species groups supported by molecular phylogenetic analyses in the current study, and groups based on the morphology of the inner surface of the lips by Tshibwabwa & Teugels (1995).

Within southern African *Labeo* species, four monophyletic lineages (A, B, E and F) were apparent (Fig. 2.12). These lineages did not fully represent the four *Labeo* spp.

groups that were proposed by Reid (1985), the exception being that an LCG species (*L. congoro*) was placed in the LNG.

Phylogenetic Relationships of Southern African Lineages

Labeo niloticus group sensu lato (Lineage A) and Lineage F

All of the *Labeo* spp included in this chapter that were proposed to belong to the LNG grouped together in the ML phylogenies (Figs. 2.6–2.10) with the exception of *L. ruddi*. Phylogenetic analyses showed that *L. ruddi* is not closely related to any of the species in the LNG as proposed by Reid (1985). Instead, the analyses showed *L. ruddi* to be closely related to *L. vulgaris* Heckel 1847 and that both species formed a separate clade (lineage F). This implied close relationship between the two species is surprising, as *L. vulgaris* occurs within the Nile River system located in North Africa (Yang *et al.*, 2012), approximately 3000 km distant from the Cunene River system where *L. ruddi* occurs, and the two species diverged about 2.48 mya (Fig. 2.11). Between the two river systems lies the East African Rift System and the sickle-shaped Congo basin, which were formed 3.1 and 2.5–1.5 mya, respectively (Goodier *et al.*, 2011; Prüfer *et al.*, 2012). The formation of the latter systems could have acted as a geographical barrier and led to the divergence of the most recent common ancestors of the two species. Inclusion of the *L. ruddi* population from the Limpopo River system in future studies is important to fully understand relationships within this newly discovered lineage F.

The present results indicate that *L. lineatus* is not closely related to *L. weeksii*, with which it co-occurs in the upper and lower Congo Basin as proposed by Reid (1985), but instead it is closely related to *L. congoro*, which was proposed to belong to the LCG *sensu lato* (Reid, 1985) (Figs. 2.6, 2.8–2.10). Currently, *Labeo congoro* is known to

occur in Central Africa (Katanga and Lake Mweru) separated by the Congo Basin. The close phylogenetic relationship between *L. lineatus* and *L. congoro* might indicate that the ancestor of the two species occurred throughout their distributions. According to Goodier *et al.* (2011) a barrier in the form of a high-lying landscape (the Kundululungu Plateau 1000 m above sea level; Flügel *et al.*, 2015) formed about 1.5 mya when Lake Tanganyika became isolated from the Congo River system and might have led to the split between *L. lineatus* and *L. congoro*. The plateau, located in the southern part of the Congo Basin, is part of the high landscape that separates the Congo Basin from the East African Rift System. The results of the dated BI analysis (Fig. 2.11) indicated that *L. lineatus* and *L. congoro* may have diverged after the formation of the latter barrier about 1.13 mya. Inclusion of *L. rosae* in the analysis would have allowed evaluation of its placement within the LNG and to test the validity of its hypothesised close (morphological) relationship with *L. altivelis* (Reid, 1985).

An unidentified specimen (*Labeo* sp. 1; SAIAB no. 77992) from Oubangui River, which forms part of the Congo Basin, was shown in all phylogenetic trees (Figs. 2.6–2.10) to be closely related to *L. weeksii*. This specimen might represent an additional form of *L. weeksii*, as both forms occur in the same drainage system and are indicated to have diverged recently (Fig. 2.11). However, a morphological comparison between the forms is needed to clarify the identity of the unidentified specimen.

Relationships within the Labeo forskalii Group

Sublineage B1

The unidentified specimens [*Labeo* sp. 1 and 2 (SAIAB no. 85508), *Labeo* sp. 3 (85336), *Labeo* sp. 4 (85157) and 5 (85189), *Labeo* sp. 6 and 7 (85101) and *Labeo* sp. 8 (85370)] from the Kwanza River system in Angola were shown to be monophyletic and formed a clade with two taxonomically indeterminate individuals [*L.* cf. *annectens* (84710)] (see Table 2.1 and Fig. 2.11). *Labeo annectens* occurs in the Congo River Basin and in most basins of the lower Guinea region (Moelants, 2010), thus the *L.* cf. *annectens* individuals included in the present study may represent an unnamed species. Using the phylogenetic species concept (monophyly), two of the closely related unidentified forms [(d) and (e)] with the lowest sequence divergence (0.2%) (Fig. 2.11), were compared to two groups of closely related, recently diverged and morphologically distinct species [*L. capensis* and *L. umbratus* (sequence divergence of 0.15%); *L. horie* and *L. senegalensis* (0.14%)]. I then concluded that, the unidentified individuals form five monophyletic forms, which may warrant recognition as distinct species. The forms are: (a) *Labeo* sp. 1 and 2, (b) *Labeo* sp. 4 and 5, (c) *Labeo* sp. 6 and 7, (d) *L.* cf. *annectens* 1, (e) *L.* cf. *annectens* 2 and *Labeo* sp. (3 and 8). Thorough taxonomic investigation of these forms is needed.

Sublineage B3

Reid (1985) proposed that *L. cylindricus* closely resembles *L. ansorgii*. The present results show that *L. cylindricus* is closely related to *L. parvus* (Baro), whereas *L. parvus* (Congo) clustered with *L. nasus*, *L. quadribarbis* and *L. simpsoni* to form the monophyletic sublineage B2 (Figs. 2.7 and 2.9). *Labeo parvus* is widely distributed within the Congo River Basin, Chad, Senegal, Gambia, Volta and Niger basins as well

as in the Ouémé and Mono rivers and other West African coastal basins (Lévêque 1990). The species is also present in Lake Tanganyika and the Malagarasi River Basin in East Africa (Hanssens *et al.*, 2010). Contrary to the provenance of the *L. parvus* specimen stated by Yang *et al.* (2012), there seems to be no prior record of *L. parvus* in the Baro River (Lévêque 1990). Yang *et al.* (2012) might have confused *L. parvus* with a newly discovered unnamed form or a divergent form of *L. cylindricus*. The *L. parvus* sample analysed by Lowenstein *et al.* (2011) may be correctly identified as it was collected within the Congo Basin. A morphological comparison of *L. cylindricus* and *L. parvus* (Yang *et al.*, 2012) is needed to resolve the uncertainties between these species. Samples of *Labeo cylindricus* populations from different river systems were shown to share alleles (Figs. 2.6 and 2.7). A population genetics study across its entire distribution might reveal how this species managed to attain its wide distribution.

Reid (1985) also proposed that *L. ansorgii* closely resembles *L. forskalii* and *L. cylindricus*. The results from the current study showed *L. forskalii* to be the sister species to *L. cylindricus* and *L. parvus sensu* Yang *et al.* (2012) (Fig. 2.11). The finding is sensible as *L. forskalii* occurs in the Nile River system (Getahun & Twongo, 2010), which is where *L. cylindricus* occurs, and is close to Baro River where the *L. parvus* specimen of Yang *et al.* (2012) was collected. *Labeo forskalii* may have diverged from the common ancestor of *L. cylindricus* and *L. parvus sensu* Yang *et al.* (2012) within the Nile River system (possibly by sympatric speciation) about 1.47 mya as indicated by estimated divergence times (Fig. 2.11).

The present analyses showed *Labeo ansorgii* to be closely related to *L. lunatus* (Fig. 2.11). *Labeo ansorgii* occurs in the Cunene and Kwanza river systems that flow westward into the Atlantic Ocean (Reid, 1985), whereas *L. lunatus* occurs within the drainage basin of the upper Zambezi/Okavango rivers, which are in close proximity to the Kwanza and Cunene rivers but flow eastward into the Indian Ocean (Marshall & Tweddle, 2007). The two species were indicated to have diverged recently in the late Holocene period (Fig. 2.11). At this time a single, widespread ancestral population of the two species may have diverged into two populations, possibly because of isolation of the two river systems (i.e., geographic speciation) (Hammer *et al.*, 2013).

The results from the current study indicated *L. molybdinus* to be a sister species to both *L. lunatus* and *L. ansorgii* (Fig. 2.11). This relationship is contrary to Reid's proposal (1985) that *L. lunatus* is a form of *L. molybdinus* with a sail-like dorsal fin (Reid, 1985). *Labeo molybdinus*, as with *L. lunatus*, occurs within the Zambezi River system, but the distribution of the former starts from the middle Zambezi down to the lower Zambezi (Bills *et al.*, 2007), whereas *L. lunatus* occurs in the upper Zambezi and Okavango rivers (Skelton, 2001; Marshall & Tweddle, 2007). The upper and middle Zambezi are separated by a barrier in the form of a waterfall (Victoria Falls). The upper and middle Zambezi were once separate river systems, according to Balon (1974), and the Victoria Falls were formed about 0.50 mya when the two parts of the Zambezi joined. The divergence time estimates (Fig. 2.11) indicate that *L. molybdinus* diverged from the recent common ancestor of *L. lunatus* and *L. ansorgii* about 0.97 mya, which was before the formation of the Victoria Falls. Thus, *L. molydinus* and *L. lunatus* may have speciated before formation of the Victoria Falls. Balon (1974) indicated that the two parts of the Zambezi harbour different species and proposed that the "pre-Upper

Zambezi River" provided a variety of habitats, which promoted high fish species diversity.

Labeo umbratus group

As expected, the subpopulations of *L. umbratus* in the Orange River and two southern-flowing river systems (Gamtoos and Gourits) were shown to be closely related (Figs. 2.8–2.10). A phylogeographic analysis of these populations is presented in Chapter 3. *Labeo capensis* was shown to be more closely related to Orange River *L. umbratus* in the Cyt *b* tree (Fig. 2.6). The relationship between these two species is explained in detail in Chapter 4. Both *L. umbratus* and *L. capensis* were shown to be closely related to *L. rubromaculatus*, which occurs in the Tugela River system that flows south-eastward into the Indian Ocean (Figs. 2.6 and 2.7) (Cambray, 2007). The Orange and Tugela river systems, in which the latter two *Labeo* species occur, are separated by the Drakensberg Mountains. The estimated divergence time of the recent common ancestor of *L. umbratus* and *L. capensis* from *L. rubromaculatus* was about 0.85 mya (Fig. 2.11). The ancestral population of *L. rubromaculatus* must have found a means of reaching the Tugela River system but the mountains acted as a barrier, blocking the species from colonising the Orange River until the development of the Tugela-Vaal inter-basin water transfer scheme. The phylogenetic analyses estimated the earliest split in the LUG to be between *L. seeberi* and the recent common ancestor of the other three species, rather than between *L. rubromaculatus* and the remaining species as Reid (1985) suggested. *Labeo seeberi* occurs in the Olifants River system (Lubbe *et al.*, 2015) and, according to Dingle and Hendey (1984), the Orange and the Olifants river systems have been connected in the past due to river capture. The Orange River system has changed its exit points four times in the past (Dingle & Hendey, 1984). It was postulated that during the

Late Cretaceous the Orange River used the 28°S exit point, which is where the Western Cape Olifants River system currently exits. During the Palaeogene period, it used the current 31°S exit point. In late Miocene epoch, the Orange River switched back to the 28°S exit point. Further alterations in the drainage patterns were postulated during the late Miocene–Pleistocene epoch. If the divergence time estimates obtained in the current study are accurate, the Orange River may have switched back to the current exit point about 1.94 mya and led to the divergence between *L. seeberi* and the recent common ancestor of the *L. rubromaculatus*, *L. capensis* and *L. umbratus* lineage around that time.

Smith (1841) originally described *L. umbratus* and *L. capensis* as *Abrostomus umbratus* and *A. capensis*. However, Smith's original type specimens cannot be located (Reid, 1985). Boulenger (1909) rediagnosed the species and described them as *L. umbratus* and *L. capensis*. According to Reid (1985), this was accepted by Gilchrist and Thompson (1913) and revised and clarified by Barnard (1943). The four species in the LUG resemble each other but differ in proportions, modal meristics and in colouration (Reid, 1985). These species differ from other African *Labeo* lineages in that they possess the highest scale count, lack tubercles on the snout, are restricted to the southern Africa (Reid, 1985; Skelton, 2001), and are genetically distinct (the current chapter). Because this group is morphologically, genetically and geographically distinct from other *Labeo* species groups, it is concluded that the group could be promoted to genus level with a new name. This would however require additional work on the taxonomy of the full *Labeo* group, which was beyond the scope of this thesis.

Hybridisation potential

Evidence for possible hybridisation between *L. rubromaculatus* and *L. capensis* owing to the presence of *L. capensis* nuclear DNA within *L. rubromaculatus* (Fig. 2.8). Mitochondrial Cyt *b* (Fig. 2.6) and COI (Fig. 2.7) sequence data showed that *L. rubromaculatus* is a species distinct from *L. capensis* and *L. umbratus*, but the nuclear Rag1 tree (Fig. 2.8) resolved *L. rubromaculatus* to be closely related to *L. capensis*. This incongruence could be due to incomplete lineage sorting/random lineage sorting or ancestral polymorphism. An additional potential scenario is ancient introgressive hybridisation because replacement of a *L. rubromaculatus* nDNA gene would be expected after several generations of hybridisation (Allendorf *et al.*, 2001). Koblmüller *et al.* (2008), in a study of the endemic Lake Tanganyika cichlid tribe Tropheini using two mtDNA (ND2 and control region) and AFLP markers, found that the two types of markers lacked congruency and the authors attributed this to incomplete lineage sorting/random lineage sorting or ancestral polymorphism. The hypotheses that inter-basin water transfer schemes that connect the Orange and Tugela rivers systems might lead to hybridisation (and possible backcrossing as suggested above) between *L. rubromaculatus* and *L. capensis* or *L. umbratus* is feasible (Chapter 1). *Labeo capensis* has also being found to hybridise with *L. umbratus* where either species was introduced (Ramoejane, 2010; Chapter 4).

Possible hybridisation between *L. molybdinus* from the Revue River in Mozambique and *L. lunatus* was also indicated because *L. molybdinus* individual's mtDNA (Cyt *b*) was associated with *L. lunatus* (Fig. 2.8). *Labeo cylindricus* may also hybridise with *L. lunatus* because mtDNA of the latter species was detected in one *L. cylindricus* individual from the Bua River (Lake Malawi). These putative hybridisation instances

also may be the result of incomplete lineage sorting/random lineage sorting or ancestral polymorphism (Koblmüller *et al.*, 2008). Divergence time estimates indicated that the above-mentioned species may have diverged recently during the Pleistocene epoch (0.126–2.58 mya) (Lourens *et al.*, 2004), during which most species within each lineage were estimated to have diverged (Fig. 2.11). Thus, reproductive isolation mechanisms may not have evolved completely between the extant species within the lineages, and any translocation or disturbance of the species' natural habitats might lead to hybridisation.

Evidence for hybridisation between species from different species groups was also observed. An unidentified specimen (*Labeo* sp. 2; SAIAB no. 77595) from the Mbourou River (Congo Basin) was placed within the LFG in the COI tree (Fig. 2.7), within the LNG in the Rag1 tree (Fig. 2.8), and between the LFG and LNG in the COI+Rag1 tree (Fig. 2.9). In this instance, ancient hybridisation might have occurred because the individual mtDNA COI gene is divergent from that of other species within the LFG and the possible hybridisation event was between species groups that diverged from each other over 4.3 mya (Fig. 2.11). A hybridisation study including individuals from the LFG, LNG and additional morph types of the unidentified specimen is needed to test these hypotheses. The capability for hybridisation is interesting as some *Labeo* species from different species groups are naturally sympatric. Given the detection of possible hybridisation between two distantly related species groups, a larger set of samples from river systems in which different species co-occur might enable detection of hybrids between other species groups.

Conclusion

While additional work is needed to resolve the taxonomic confusion in this group of fish, the results presented in this chapter clarify the monophyly of African *Labeo* groups as confirmed in the COI+Cyt *b*+Rag tree (Fig. 2.10). Five potential unnamed species and one species group were detected and require further taxonomic investigation. *Labeo* spp within different species groups are prone to hybridisation. Hybridisation is not only limited to within a species group but is possible between species groups. Of particular interest is the LUG, of which all members were analysed in the present study.

CHAPTER THREE: PHYLOGEOGRAPHY OF THE SOUTHERN LINEAGE OF *LABEO UMBRATUS*

INTRODUCTION

Labeo umbratus (Smith 1841) (Fig. 3.1) is closely related to *L. capensis* (Smith 1841) (Reid, 1985; Chapters 2 and 4). Both species co-occur in the Orange River Basin, which drains the Drakensberg Mountains of Lesotho and flows west to Alexander Bay where it joins the Atlantic Ocean (Cambray *et al.*, 1986; Fig. 3.2). *Labeo umbratus* from the Orange River is genetically distinct from individuals from two currently isolated southward-flowing river systems (Gourits and Gamtoos) (Chapter 2). This finding is interesting because *L. umbratus* is found not only in the Orange, Gourits and Gamtoos river systems, but also occurs in other currently isolated southward-flowing river systems, namely the Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon (Fig. 3.2) (Jubb, 1964; Cambray & Jubb, 1977; Cambray, 1990). It would be of interest to determine if similar genetic diversity is reflected in *L. umbratus* populations from other southward-flowing river systems.

FIG. 3.1. *Labeo umbratus* from Brak River (tributary of the Orange River). Photograph by N. Mazungula, SAIAB.

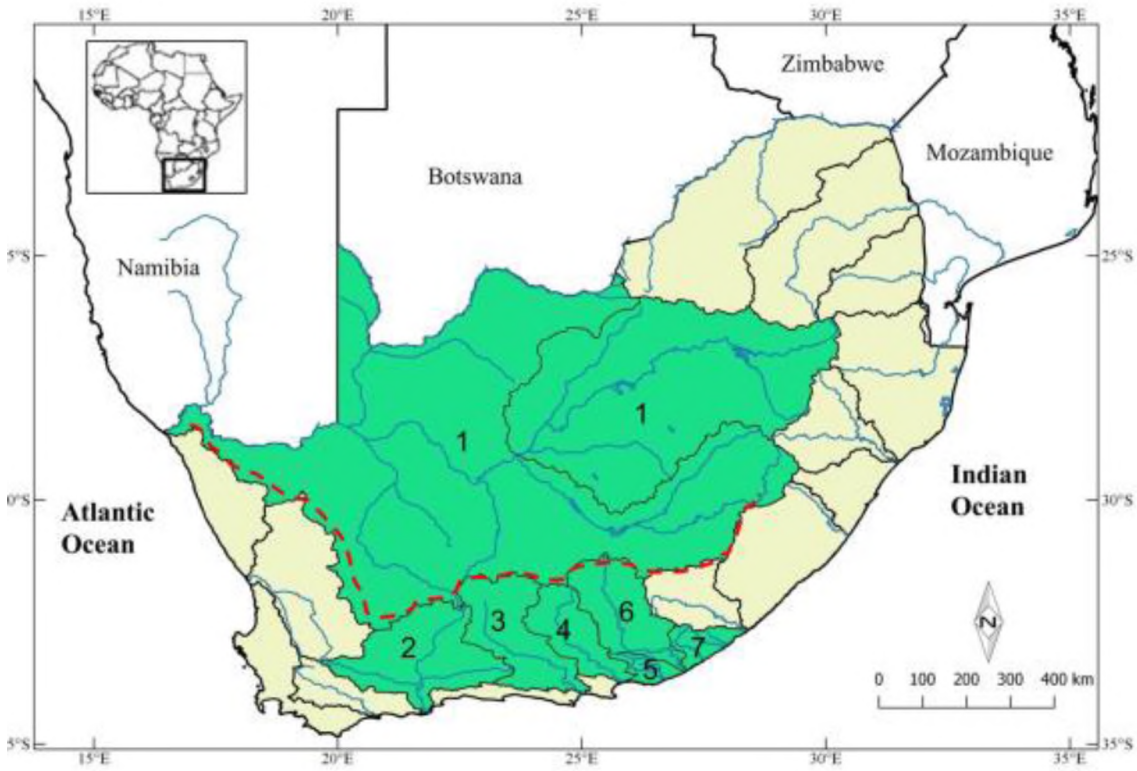


FIG. 3.2. Primary drainage systems in which *Labeo umbratus* occurs. (1) Orange, (2) Gourits, (3) Gamtoos, (4) Sundays, (5) Bushmans and (6) Great Fish. Drainage (7) is shared by the Keiskamma, Buffalo and Nahoon river systems. Broken red lines represent the Great Escarpment.

The above-mentioned southward-flowing river systems are separated from the Orange River system by the ±1200 m Great Escarpment (Fig. 3.2). It is therefore not surprising that *L. umbratus* populations in southward-flowing river systems represent a lineage genetically distinct from the Orange River system (see Chapters 2 and 4). This finding is of relevance to recent research which has demonstrated that several other fishes in southern Africa's temperate region previously thought to be single species are, in fact, species complexes (e.g., Swartz *et al.*, 2007; Chakona *et al.*, 2013a, 2013b). These species are *Pseudobarbus burchelli* Smith 1841 (Fig. 3.3a), *Sandelia capensis* (Cuvier 1831) (Fig. 3.3b) and *Galaxias zebratus* Castelnau 1861 (Fig. 3.3c) (Chakona *et al.*,

2013b). *Pseudobarbus burchelli* consisted of four very distinct (3.8–10.0% genetic distance) phylogroups, *S. capensis* contained two deeply divergent (5.5–5.9%) lineages and seven minor lineages with strong geographical congruence, and *G. zebratus* comprised nine highly divergent lineages (3.5–25.3%). Chakona *et al.* (2013a) attributed the dispersal of these primary freshwater fish species across isolated river systems to the influence of extrinsic factors and intrinsic adaptations.

Extrinsic factors such as lower sea levels (Swartz *et al.*, 2007), whereby the sea level regressed in the past (c. 22,000–18,000 years ago) during the Last Glacial Maximum (LGM) to about 120 ± 5 m below present levels. This allowed adjacent river systems that shared the same paleo-river to connect, thus facilitating movement of fish between systems. Freshwater fishes, such as *Galaxias* sp. 'nebula' (Chakona *et al.*, 2013a), used intrinsic adaptations (e.g., air breathing) to move between river systems that did not coalesce during periods of heavy flooding in the Pleistocene and Holocene epochs.

Understanding the genetic diversity of the southern lineage of *L. umbratus* in these systems is not only important for prioritising conservation management strategies for this fish, but may also help with understanding the historical connections between rivers and dispersal of the species lineages (Zardoya & Doadrio, 1999; Pusey *et al.*, 2004; Unmack, 2013).

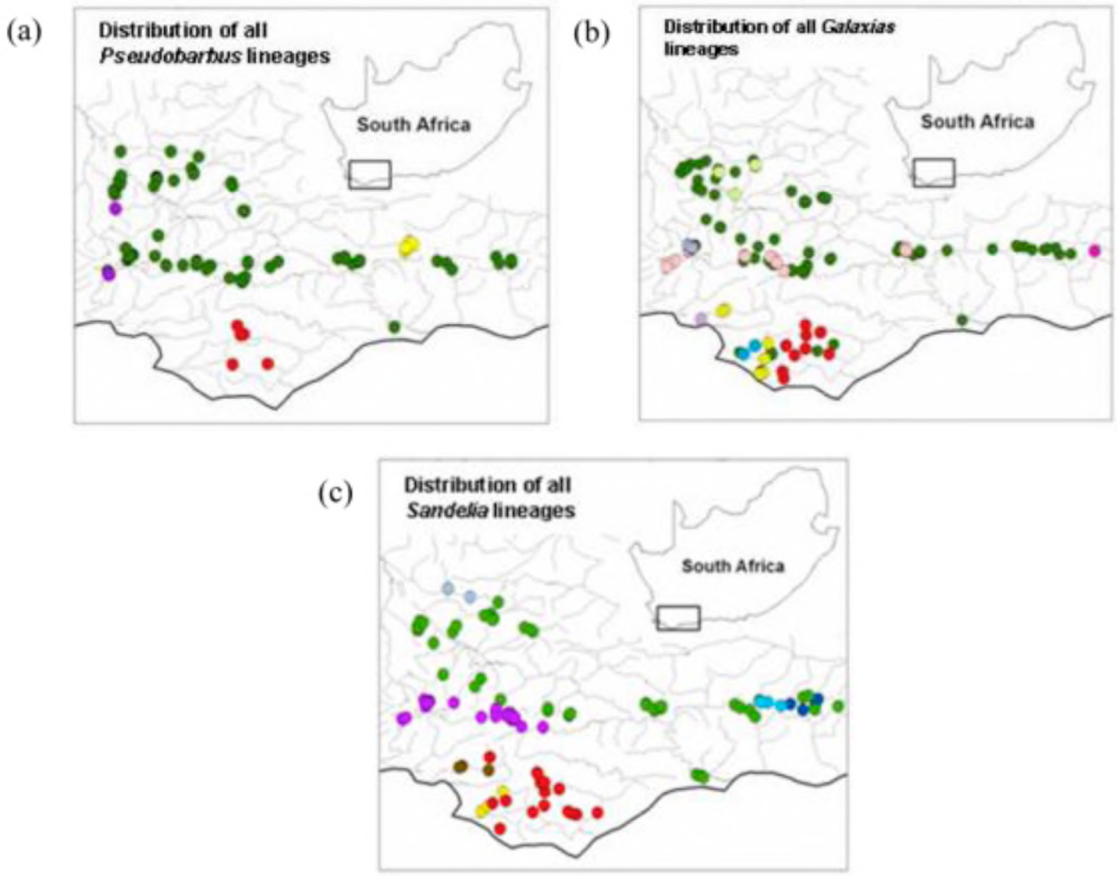


FIG. 3.3. Distribution of (a) *Pseudobarbus*, (b) *Galaxias* and (c) *Sandelia* lineages in the south-western Cape Floristic Region indicated by different colours. Maps reproduced from Chakona *et al.* (2013a).

River Connections

Burridge *et al.* (2008) reviewed previously proposed mechanisms of dispersal of primary freshwater fishes between isolated river systems and categorised these mechanisms into two classes: vicariance and dispersal.

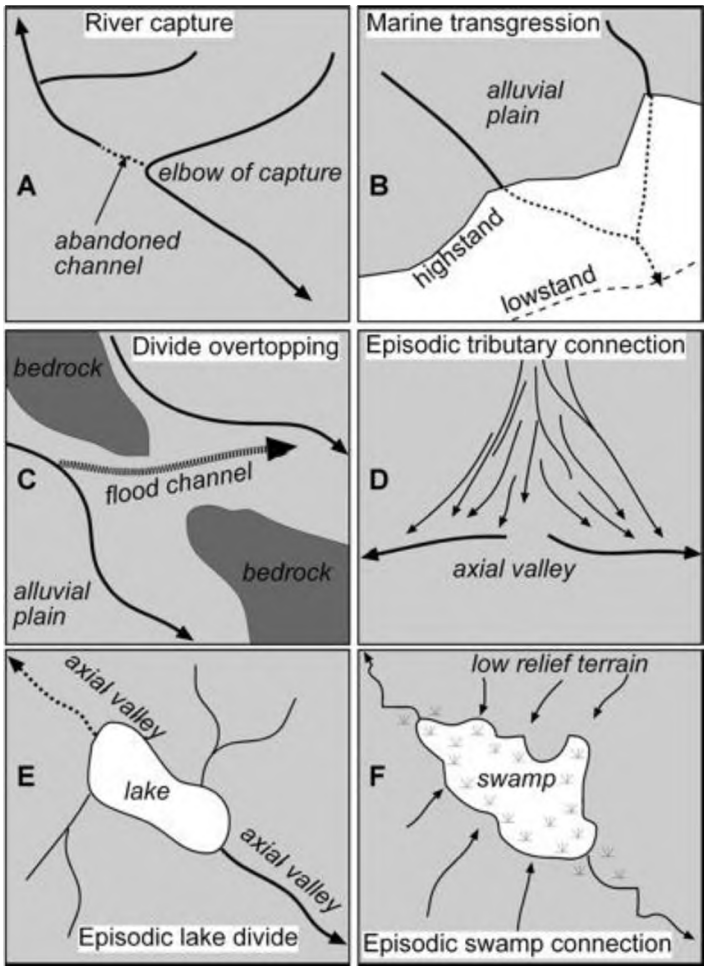


FIG. 3.4. Six geomorphological processes by which freshwater-limited fish taxa could have become distributed in catchments presently isolated from one another by marine and terrestrial barriers. (a) River capture, (b) marine transgression, (c) divide overtopping, (d) episodic tributary connection, (e) episodic lake divided and (f) episodic swamp connection (reproduced from Burridge *et al.*, 2008).

Vicariance is the geographical division of a population, typically by a physical barrier, resulting in a loss of intervening freshwater environment. Vicariance mechanisms comprise (a) river capture, which happens when some parts of the headwater river tributaries are captured by adjacent drainage tributaries, and (b) marine transgression, where the rise in sea level after a glaciation period separates two adjacent systems that had a common confluence during periods of lower sea levels (Fig. 3.4). Glaciation events

can also change the flow direction of streams by depositing huge boulders, rocks, unsorted sand, silts and clay as the glacial ice melts (Ray, 1974).

Dispersal is a chance movement between isolated habitats. Four mechanisms by which primary freshwater fishes disperse are recognised: (1) divide overtopping, where water from one or more river systems overflows, during flooding, into adjacent river systems; (2) episodic tributary connections, by which tributaries close to the drainage divide periodically exchange water with neighbouring river systems; and (3) episodic lake divide or (4) episodic swamp connection, by which a lake or a swamp located on the drainage divide may discharge water between adjacent systems at different times, or continuously flow to both adjacent river systems (Fig. 3.4). Freshwater fish species that live in the vicinity of the connections may use these mechanisms of dispersal (Unmack, 2013).

Dispersal-related behaviour of *L. umbratus*

Labeo umbratus is a primary freshwater fish that breeds in summer by migrating upstream or by undertaking lateral migrations from the river onto floodplains during flooding (Mulder, 1973; Tómasson *et al.*, 1984; Cambray, 1990). Depending on their size, *L. umbratus* spawn by producing 36,500–210,000 small eggs (Potts *et al.*, 2005), with eggs hatching within one to three days (Tómasson *et al.*, 1984). Tómasson *et al.* (1984) suggested that *L. umbratus* were able to disperse during heavy flooding because of their occurrence in isolated pools of water within the vicinity of the river system. Hamman *et al.* (1982) stated that such migration was observed during heavy flooding in summer (1980/81) in the Gourits River system (Hamman *et al.*, 1982, cited in Tómasson *et al.*, 1984). According to Cambray (1990), migration is undertaken not only

to breed, but also to seek suitable habitats for feeding, shelter and colonisation. Cambray (1990) indicated that both adults and juveniles migrate, of which juveniles are considered to show a greater ability for dispersal (Tómasson *et al.*, 1984). The ability to migrate was proposed to have evolved to optimise feeding, to avoid unfavourable conditions and possibly to promote colonisation (Cambray, 1990). Thus, *L. umbratus* could be considered to be a species able to use opportunistic connections between neighbouring drainage systems to disperse. In the southern drainage systems, potential drivers of this dispersal are geological processes, climate change, sea-level changes and human introductions.

Geological Processes

According to Cowling *et al.* (2009), the present landscape and drainage patterns in southern Africa are of relatively recent origin. Cowling *et al.* (2009) came to this conclusion because of the young (early Pliocene, 5.3 mya) geomorphic features in the region that support specialised vegetation types (e.g. renosterveld, succulent karroo, and limestone fynbos). Southern Africa has experienced two tectonic uplifts, which raised the interior and changed the drainage patterns (King, 1963, Partridge & Maud, 1987, 2000). The first uplift (250–300 m in height) was in the early Miocene (22 mya) and the most recent uplift (200–300 m in height), which was much more intense than the first uplift, occurred in the Pliocene (5.3–2.6 mya). This change in landscape rejuvenated the drainage systems in southern Africa (Dollar, 1998). The drainage changes and tectonic uplifts led to the capture of the Orange River by the lower Orange from the Olifants River system, the creation of the Augrabies Waterfalls that now act as a biogeographic barrier between the lower and upper Orange River, and the uplift of the Great Escarpment, which was already formed 180 mya and currently acts as a barrier between the Orange

River and the southward-flowing drainage systems (Cambray *et al.*, 1986). The Great Escarpment was moved inland to the present position over 20 mya by the upliftment of the central plateau (Truswell, 1977).

The confluence between the Orange and Olifants river systems is well documented (Dingle & Hendey, 1984; De Wit, 1993), but the links between the former and southward-flowing river systems of the Western and Eastern Cape provinces are not well understood. Jubb & Farquharson (1965) supported the hypothesis for the link between the Olifants and Orange River systems and also suggested links between the Orange River system and southward-flowing river systems, because of the presence of *L. umbratus* and *Barbus anoplus* Weber 1897, which are found in the Orange and southward-flowing river systems. This connection was hypothesised to have aided the migration of these fishes to the southward-flowing river systems (Jubb & Farquharson 1965).

Climatic Changes

Climate change is perceived to be the major driving force responsible for changes in landforms and drainage re-orientations leading to the present landscape (Craw *et al.*, 2008). Southern Africa experienced a warm tropical climate that led to severe flooding during most of the Pliocene epoch (Maud & Partridge, 1987; Hattingh, 1996). Such flooding may have facilitated connections between isolated river systems by means of divide inundation or water exchange between tributaries close to drainage divides or on the flood plains. By the end of the Pliocene and beginning of the Pleistocene, southern Africa experienced cooler, drier and more acidic conditions (Harwood, 1985; Partridge,

1993; Cowling *et al.*, 2009). The cooling was due to alternating glacial and interglacial conditions (Hattingh, 1996) that led to sea-level changes. These cool conditions were responsible for the mass extinction of molluscs along the South African coast and elsewhere (Raffi *et al.*, 1985; Stanley, 1986). During the Holocene (altithermal 6,000–8,000 years ago), southern Africa experienced wet climatic conditions (high rainfall) in some areas (Partridge *et al.*, 1999). Such conditions could have increased connections between isolated river systems.

In addition, all of the above-mentioned southward-flowing river systems share a continuous floodplain (Fig. 3.5). During very high rainfall that led to high flooding, the southern systems may have connected via these floodplains and facilitated dispersal of *L. umbratus*. However, the continuous floodplains do not imply that *L. umbratus* dispersal was continuous. The distance between the Gourits and Gamtoos river systems floodplain (302 km) is much greater than among other southward-flowing river systems, so the two systems were probably rarely connected in this manner.

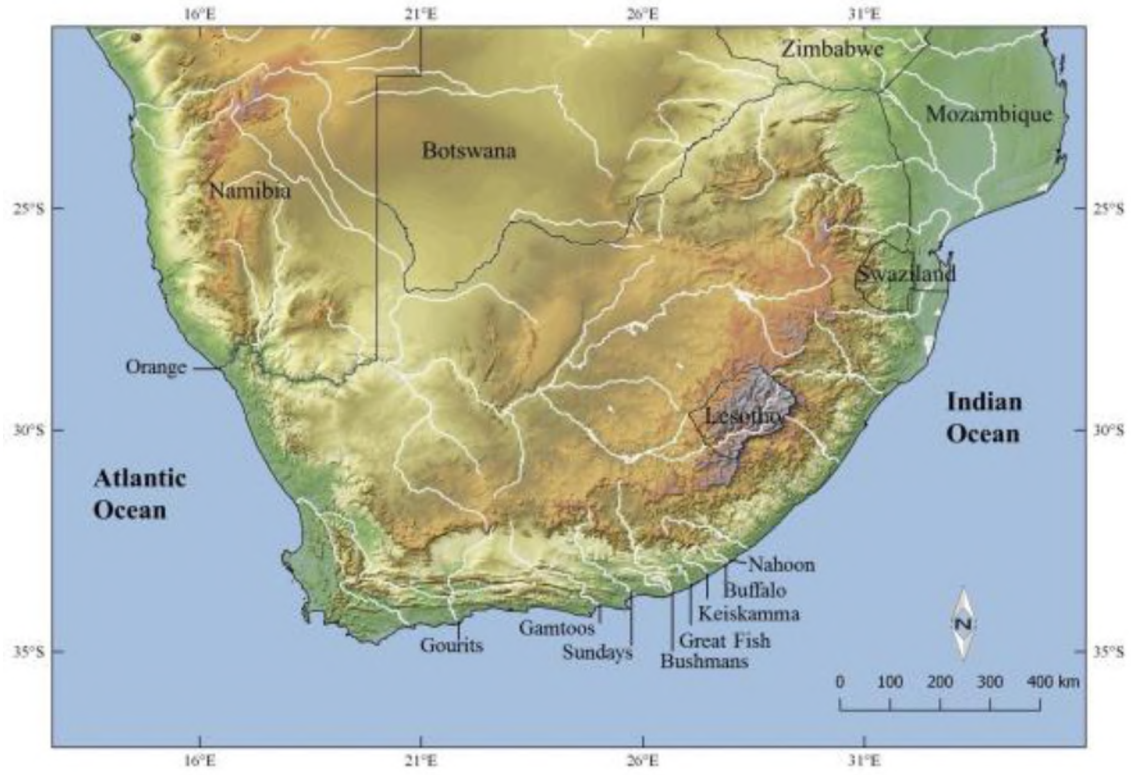


FIG. 3.5. Different altitudinal zone; Green = below 610 m (low-lying areas or floodplains where connection between drainage systems is possible).

Sea-level Changes

The southern African coastal systems experienced transgressions during the early Pliocene epoch (3.4–5.2 mya) (Butzer & Helgren, 1972) and regression during the LGM (18,000 ya) (Ruddock, 1947; Siesser & Dingle, 1981). The sea level regressed to about 120 ± 5 m below the current sea level (Tankard, 1976; Rogers, 1985; Ramsay & Cooper, 2002), which would have allowed several neighbouring river systems to connect through a common confluence and form palaeo-river systems (e.g., Swartz *et al.* 2007). A number of southern African primary freshwater fish species may have used this as a means of dispersal to other currently isolated systems (Swartz *et al.*, 2007, 2009; Chakona *et al.*, 2013a) and it is possible that *L. umbratus* also made use of such

opportunities as it is a primary freshwater species. The transgression phase that followed and currently prevails will have led to the separation of such river systems, thus inhibiting the dispersal of primary freshwater fish between adjacent river systems.

Human Introductions

Fishes can also be moved between systems due to human influence (Ellender & Weyl, 2014). A number of vectors, such as recreational angling, conservation translocations, inter-basin water transfer schemes (IBWTs) and biocontrol, are drivers of recent human-mediated fish introductions in South Africa (Bruton & Van As, 1986). *Labeo umbratus* is thought to have been introduced into the Buffalo and Nahoon rivers, possibly by anglers as bait (Jubb, 1964), as these populations were recently discovered.

Aims and Objectives

The aim of the present study was to assess the genetic variation within the southern lineage of *L. umbratus* and to reconstruct its evolutionary history in relation to the drainage history of the region. Specific hypotheses that will be addressed in this chapter are that: 1) isolation of the southward-flowing river systems led to genetic divergence; 2) southward-flowing river systems were connected in the recent past due to geological, climatic or sea-level changes; and 3) the *L. umbratus* populations in the Buffalo and Nahoon rivers are naturally occurring and not introduced as suggested by Jubb (1964).

MATERIALS AND METHODS

Sampling

A total of 172 samples of *L. umbratus* were collected from the Orange River and the southward-flowing river systems (Gourits, Gamtoos, Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon) and were used in this chapter (Fig. 3.2, Table 3.1). Specimen collection and preservation followed the same procedure outlined in Chapter 2.

TABLE 3.1. Number of specimens of *Labeo umbratus* collected from each locality and geographic coordinates for phylogeographic analysis.

Locality	River system	No. of specimens	Latitude	Longitude
Gariep Dam	Orange	29	30° 38' 38.2" S	25 °33' 50.9" E
Vaal Dam	Orange	10	26° 51' 58.9" S	28 °10' 14.3" E
Brak	Orange	10	31° 32' 26.4" S	22 °20' 35.0" E
Nahoon Dam	Nahoon	18	32° 54' 18.4" S	27 °48' 32.2" E
Kwaklifu	Buffalo	10	32° 56' 03.5" S	27° 26' 25.0" E
Need's Camp	Buffalo	10	32° 59' 30.0" S	27° 38' 25.2" E
Middledrift	Keiskamma	10	32° 49' 08.6" S	26° 59' 39.7" E
Keiskammahoek	Keiskamma	10	32° 41' 12.7" S	27° 09' 09.1" E
Kat River Dam	Great Fish	16	32° 33' 46.5" S	26 °46' 43.0" E
Amakhala Game Reserve	Bushmans	10	33° 31' 02.5" S	26 °07' 29.2" E
Slagboom Dam	Sundays	10	33° 22' 31.1" S	25 °40' 45.4" E
Near Mont Pellier	Gamtoos	7	33° 13' 38.5" S	24 °09' 15.0" E
Perdegat Pool near Steytlerville	Gamtoos	12	33° 18' 41.8" S	24 °20' 50.0" E
Stompdrift	Gourits	5	33° 30' 42.3" S	22 °36' 14.2" E
Die Poort	Gourits	5	33° 58' 34.8" S	21 °39' 19.0" E
Total		172		

DNA Extraction, PCR and Sequencing

The mitochondrial cytochrome *b* (Cyt *b*) and nuclear S7 ribosomal protein gene intron 1 (S7) gene regions were amplified and sequenced from DNA extracted from the 172 samples. The same procedures described in Chapter 2 were followed for DNA extraction and PCR amplification for Cyt *b*. PCR amplification of S7 (using the primers S7RPEX1F [forward] and S7RPEX2R or S7RPEX3R [reverse]) was performed in a final volume of 50 µl consisting of 6 µl DNA, 5 µl of 1× buffer, 5 µl of 2 mM $MgCl_2$, 5 µl of 0.2 mM dNTPs, 1 µl each 20 mM primer, 0.2 of 5 U/µl SuperTherm Taq DNA polymerase (Hoffman-La-Roche, US) and 26.8 µl double-distilled water. The PCR conditions were as follows: initial denaturation 92 ºC for 2 min; followed by 35 cycles of denaturation at 92 ºC for 60 s, annealing at 58 ºC for 60 s and extension at 72 ºC for 60 s; and a final cycle of extension at 72 ºC for 7 min. Fewer samples were analysed for S7 than for Cyt *b* because of amplification difficulties. The sequencing, alignment and editing of the sequences followed the procedure detailed in Chapter 2.

Sequence Variation and Diversity

DnaSP 5.10 (Rozas & Librado, 2009) was used to phase S7 sequences and to identify the unique haplotypes and alleles of Cyt *b* and S7. The number of variable, parsimony-informative sites within the ingroup was determined with the same software. Arlequin 3.5 (Excoffier & Lischer, 2010) was used to infer population history from diversity indices [haplotype (*h*) and nucleotide (π) diversity] and their standard error (SE) for each river system. Values of $h < 0.5$ and $\pi < 0.5\%$ were considered to represent low diversity (Grant & Bowen, 1998; Lin *et al.*, 2010).

Palaeo-river Reconstruction

Palaeo-rivers of the southward-flowing river systems were assessed to find possible links that could explain the evolutionary history of *Labeo* populations. The reconstructions of the south-western river systems by Swartz *et al.* (2007, 2009) and Chakona *et al.* (2013a) were used.

Population Differentiation and Genetic Structure

Models of nucleotide substitution that best fitted the Cyt *b* and S7 data sets were selected from 1624 models with the Akaike information criterion (AIC; Akaike, 1974) in jModelTest 2.1.7 (Guindon & Gascuel, 2003; Darriba *et al.*, 2012). An analysis of molecular variance (AMOVA) was used to estimate population differentiation (Excoffier *et al.*, 1992). AMOVA calculates where the variation is petitioned in sequences using their allele frequencies and the mutations between the alleles. The optimal model selected with jModelTest was used to estimate genetic distances among alleles. Significance of the results was determined by 1000 permutations. These analyses were computed in Arlequin 3.5 (Excoffier & Lischer, 2010) on *a priori* defined structures. The structure that maximised the variation among the defined groups (ϕ_{CT}) was favoured. Populations were grouped into the following four structures to test the geographic partitioning variance:

1) Orange River vs southern river systems (Gourits, Gamtoos, Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon);

2) Orange vs south-western (Gourits and Gamtoos) vs south-eastern (Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon) river systems;

3) Populations were grouped into contemporary river systems;

4) Populations were grouped into palaeo-river systems, with only the Buffalo and Nahoon river systems assumed to have had a common confluence when the sea level regressed to about 130 m below the current sea level.

Using Arlequin 3.5 (Excoffier & Lischer, 2010) pairwise ϕ_{ST} values among populations were calculated to assess differentiation using the optimal model and significance determined as described above. The mean ϕ_{ST} value of 0.222 was considered as a cut-off for high values for freshwater fishes (Ward, 2000). The programme TCS 1.21 (Clement *et al.*, 2000), which determines parsimony networks based on 95% confidence of connections among alleles (Templeton *et al.*, 1992), was used to determine genealogical relationships among the sequences. The network method was preferred because it takes into account the persistent ancestral nodes, multifurcations and reticulations in contrast to strictly bifurcating phylogenetic trees (Posada & Crandall, 2001). The model selected was used to determine genetic distances among alleles and was used to construct a maximum likelihood (ML) tree with MEGA 6 (Tamura *et al.*, 2013). *Labeo seeberi* was used as an outgroup as it is a sister species to *L. umbratus*, *L. capensis* and *L. rubromaculatus* (Chapter 2).

Historical Demography

Three approaches were employed using Arlequin 3.5 (Excoffier & Lischer, 2010) to assess whether populations have undergone past expansion. Two of the approaches, Tajima's (1989) *D* and Fu's (1997) F_S neutrality tests were used to test the selective neutrality of random samples of DNA sequences. The two analyses were assumed to assess past population expansion. If the two tests are significantly negative, the null hypothesis of no expansion can be rejected. The third approach, mismatch distribution,

is a graphic method of visualising population expansion based on the distribution of the number of pairwise differences between haplotypes. From these distributions, parameters of demographic population expansion can be estimated. Unimodal distributions are indications of population expansion, whereas multimodal distributions are an indication of a stationary population. Two parametric bootstrap statistics (Rogers & Harpending, 1992) were calculated: 1) the sum of square deviations (SSD) between the observed and the expected mismatch curve, and 2) the Harpending's raggedness index (Rag) of the observed distribution, assuming a model population expansion. Significantly small values of the two statistics is an indication of demographic expansion.

Timing of Diversification

Time of divergence was calculated for populations that showed differentiation using the formula: divergence time (t) = net divergence estimate (net divergence × μ) (Elmer *et al.*, 2007), where μ is the substitution rate (0.76% or 2.2% site^{-1} my^{-1}). The two substitution rates are explained in Chapter 2.

RESULTS

Mitochondrial Cytochrome *b*

Sequence variation and diversity

Of the 729 bp fragment analysed for 172 Cyt *b* sequences, 20 sites were variable. Of these, 17 sites were parsimony informative within the ingroup. The TrN+I (Tamura & Nei, 1993) model was selected as the optimal model with jModelTest. The Cyt *b* sequences showed the following statistics: base frequencies, A = 0.303, C = 0.289, G = 0.139, T = 0.27; proportion of invariable sites (I) = 0.765; and rate matrix: R(a) [A-C] = 1.0000, R(b) [A-G] = 34.7040, R(c) [A-T] = 1.0000, R(d) [C-G] = 1.0000, R(e) [C-T] = 14.8347, R(f) [G-T] = 1.0000.

Two divergent *L. umbratus* lineages, designated the Orange lineage (Orange River) and the southern lineage, were resolved in the ML phylogram (Fig. 3.6). Within the southern lineage, minor divergence between alleles was observed, with a minor south-eastern lineage restricted to the Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon rivers resolved. The remaining alleles were restricted to the two south-western river systems (Gourits and Gamtoos) (Figs. 3.6 and 3.7, Table 3.2).

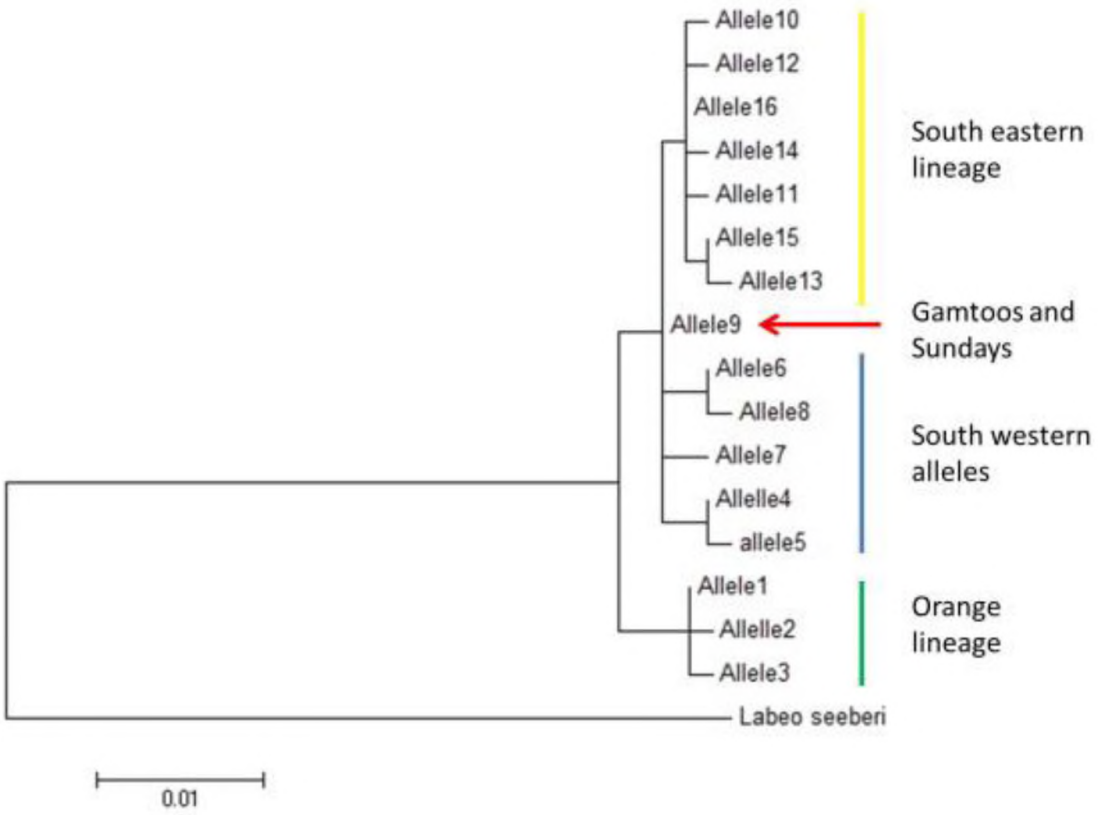


FIG. 3.6. Maximum likelihood phylogram derived from *Labeo umbratus* cytochrome *b* sequence data showing allele association with lineages. *Labeo seeberi* was used as the outgroup. Allele 9 was shared between the south-eastern and south-western lineages.

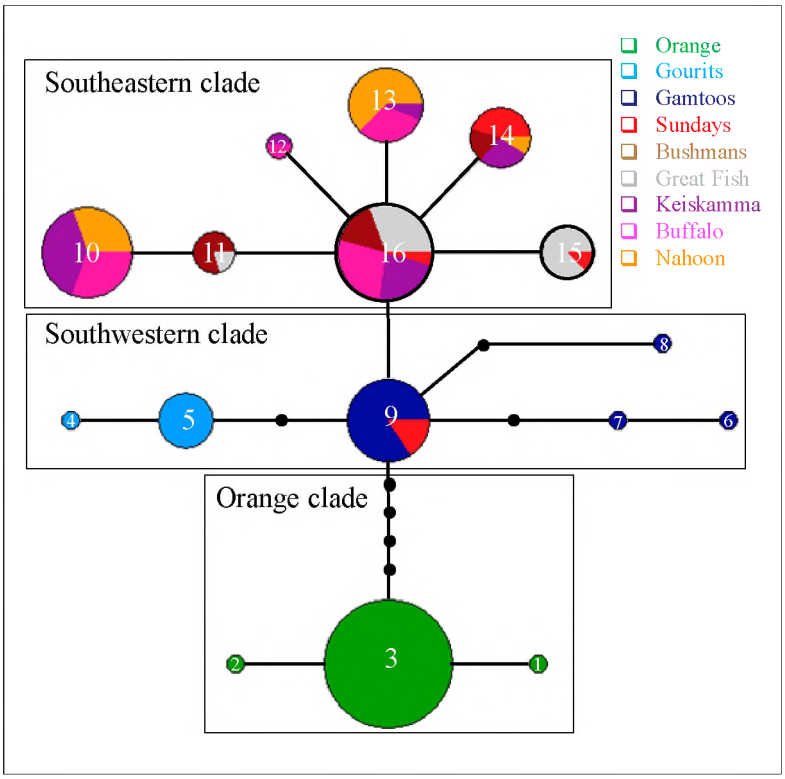


FIG. 3.7. *Labeo umbratus* mitochondrial cytochrome *b* TCS parsimony network showing the relationship between haplotypes (1 to 16) and their associations with river systems and division into three clades. The frequency of the alleles is indicated by the size of the circle.

Differentiation among populations of *L. umbratus* across its distribution range was observed, with maximum sequence divergence of 1.2% between haplotypes in different populations (Table 3.3). Very little differentiation within the Orange River system was observed (Fig. 3.7). Lower differentiation within populations of *L. umbratus* in the southward-flowing river systems was observed. The differentiation was mostly between the south-western and south-eastern populations (maximum 0.80% sequence divergence). The Orange, Gourits and Gamtoos river systems showed low haplotype (0.081–0.298) and nucleotide diversity (0.01–0.1%). The remainder of the southward-flowing drainage systems showed high haplotype diversity (0.568–0.726) and low nucleotide diversity (0.1–0.2%) (Table 3.3, Fig. 3.7).

Some haplotypes were unique to localities, whereas others were private and the majority were shared between populations of *L. umbratus* in isolated river systems (Fig. 3.7, Table 3.2). Haplotypes 1, 2 and 3 were present only in the *L. umbratus* population from the Orange River system. Haplotype 1 was present only within the Vaal population of *L. umbratus* (Table 3.2). Haplotype 2 was detected only in the population of *L. umbratus* from the Gariep impoundment (Table 3.2). Haplotype 3 was shared by all populations of *L. umbratus* from the Orange River system (Table 3.2).

The remaining haplotypes were detected only in the populations *L. umbratus* from the southward-flowing river systems (Fig. 3.7, Table 3.2). Haplotypes 4 and 5 were present only in the population of *L. umbratus* from the Gourits River system. Haplotypes 6, 7 and 8 were only present in the population of *L. umbratus* from the Gamtoos River system. Haplotype 9 was shared between the populations of *L. umbratus* from the Gamtoos and Sundays river systems. This haplotype was ancestral to the alleles in the southward-flowing river systems and the reason for it being shared could be due to ancestral polymorphism. The remaining haplotypes were shared among the populations comprising the south-eastern lineage of *L. umbratus*. Haplotype 11 was shared by populations of *L. umbratus* from the Bushmans and Great Fish river systems. Haplotypes 10 and 13 were shared by populations of *L. umbratus* from the Keiskamma, Buffalo and Nahoon rivers. Haplotype 12 was shared by populations of *L. umbratus* from the Keiskamma and Buffalo river systems. Haplotype 14 was shared among population of *L. umbratus* from the Sundays, Bushmans, Keiskamma and Nahoon river systems. Haplotype 15 was shared between populations of *L. umbratus* from the

Sundays and Great Fish river systems. Haplotype 16 was shared by almost all populations of *L. umbratus* from the south-eastern river systems except the Nahoon in the far east. The frequency of these alleles in each population of *L. umbratus* is shown in Table 3.2.

Genetic structure

The pairwise ϕ_{ST} values were high and significant ($p < 0.05$) among most populations of *L. umbratus* (0.266–0.986) (Table 3.3). This finding was an indication of the differentiation between most populations of *L. umbratus* across the species' distribution. However, the pairwise ϕ_{ST} values were low between the 1) Bushmans and Sundays (0.188), Bushmans and Keiskamma (0.08) and Bushmans and Nahoon (0.096) river systems, 2) Keiskamma and Buffalo (0.006) and Keiskamma and Nahoon (0.109) river systems, and 3) Nahoon and Buffalo (0.013) river systems. These results were an indication of the lack of differentiation between the populations of *L. umbratus* in these river systems (Table 3.3).

Palaeo-river Reconstruction

Reconstruction of palaeo-rivers revealed that only the Buffalo and Nahoon rivers among the large southward-flowing river systems, in which *L. umbratus* was collected, shared a common confluence when the sea level was 130 m lower than the present-day level (Fig. 3.8a and b).

TABLE 3.2. *Labeo umbratus* cytochrome *b* allele frequencies and their distribution among sampled localities. Alleles highlighted in bold are unique to a single river system.

Allele	Vaal	Gariep Dam	Brak	Gourits	Gourits2	Gamtoos	Gamtoos2	Sundays	Bushmans	Great Fish	Keiskamma	Keiskamma2	Buffalo	Buffalo2	Nahoon
	10	49	10	5	5	7	11	10	10	16	10	10	10	10	18
1	**1**	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	**1**	-	-	-	-	-	-	-	-	-	-	-	-	-
3	**9**	**28**	**10**	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	**1**	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	**4**	**5**	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	**1**	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	**1**	-	-	-	-	-	-	-	-
8	-	-	-	-	-	**1**	-	-	-	-	-	-	-	-	-
9	-	-	-	-	-	6	10	3	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-	-	-	4	5	2	5	7
11	-	-	-	-	-	-	-	-	4	1	-	-	-	-	-
12	-	-	-	-	-	-	-	-	-	-	-	1	1	-	-
13	-	-	-	-	-	-	-	-	-	-	1	-	4	1	10
14	-	-	-	-	-	-	-	5	2	-	1	2	-	-	1
15	-	-	-	-	-	-	-	1	-	7	-	-	-	-	-
16	-	-	-	-	-	-	-	1	4	8	4	2	3	4	-

TABLE 3.3. *Labeo umbratus* cytochrome *b* pairwise ϕ_{ST} (above diagonal) and range in percentage sequence divergence (below diagonal) from comparisons among river systems using the Tamura and Nei (TrN+I) distance method (Tamura & Nei, 1993), and molecular diversity indices [nucleotide diversity (π; %) and haplotype allele diversity (*h*)] for each population. The south-western river systems are indicated in bold and south-eastern river systems are underlined.

											Molecular diversity index	
	River system	1	2	3	4	5	6	7	8	9	Nucleotide diversity π (%)	Haplotype diversity *h*
1	Orange		0.986**	0.952**	0.957**	0.968**	0.966**	0.930**	0.931**	0.931**	0.01	0.081
2	**Gourits**	1.00–1.20		0.786**	0.791**	0.853**	0.868**	0.756**	0.756**	0.756**	0.02	0.200
3	**Gamtoos**	0.70–1.20	0.30–0.80		0.443**	0.603**	0.624**	0.556**	0.540**	0.576**	0.10	0.298
4	Sundays	0.70–1.10	0.40–0.70	0.00–0.70		0.188*	0.309**	0.271**	0.288**	0.346**	0.20	0.711
5	Bushmans	0.80–1.10	0.40–0.70	0.10–0.70	0.00–0.30		0.279**	0.080	0.096	0.212*	0.10	0.711
6	Great Fish	0.80–1.10	0.40–0.70	0.10–0.70	0.00–0.30	0.000.30		0.310**	0.266**	0.369**	0.10	0.592
7	Keiskamma	0.80–1.20	0.40–0.80	0.10–0.80	0.00–0.40	0.00–0.30	0.00–0.40		0.006	0.109*	0.20	0.568
8	Buffalo	0.80–1.20	0.40–0.80	0.10–0.80	0.00–0.40	0.00–0.40	0.00–0.40	0.00–0.00		0.013	0.20	0.716
9	Nahoon	1.00–1.20	0.60–0.80	0.30–0.80	0.00–0.40	0.00–0.30	0.10–0.40	0.00–0.00	0.00–0.10		0.20	0.726

Significant ϕ_{ST} values are indicated with asterisks: * $P < 0.05$; ** $P < 0.005$

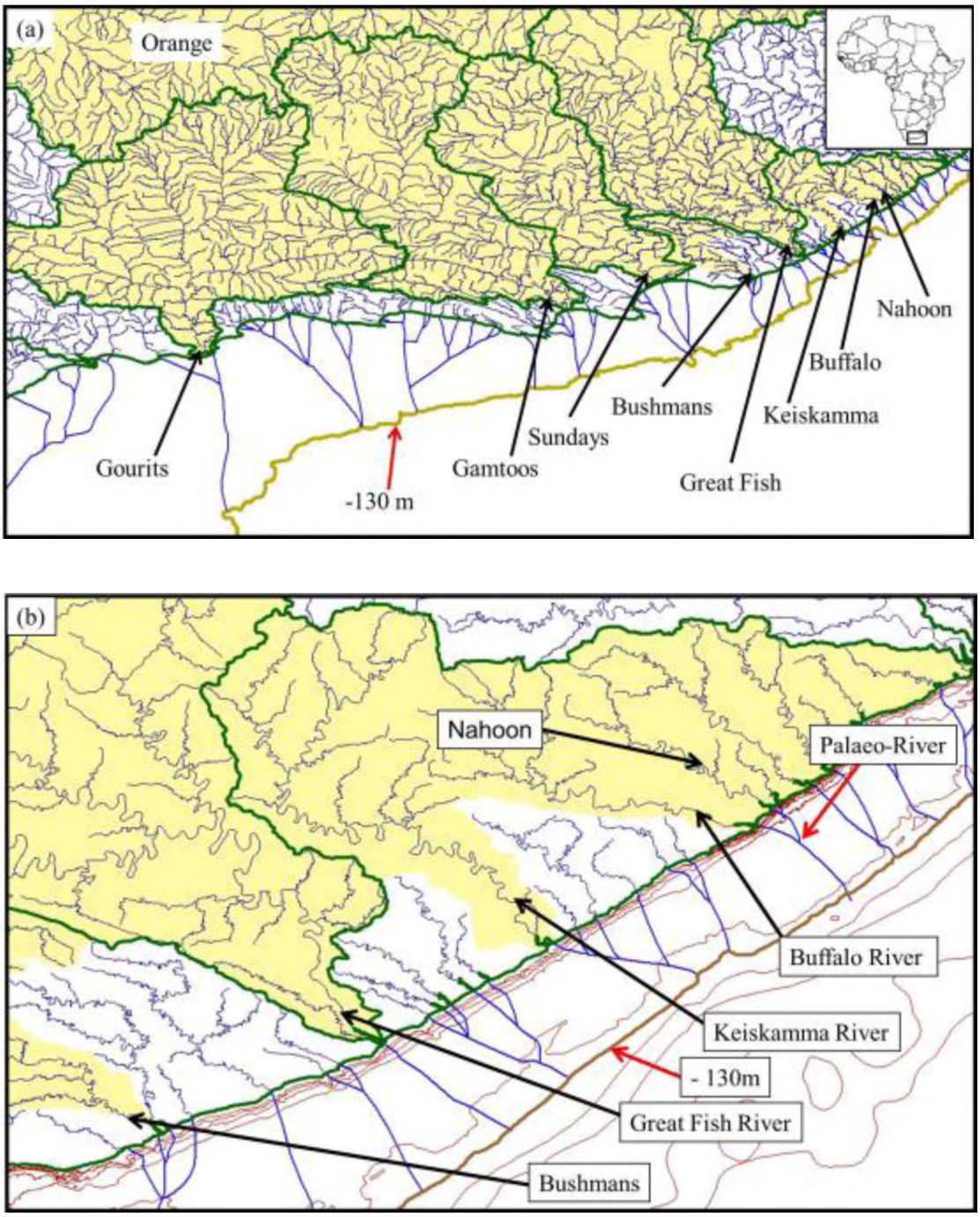


FIG. 3.8. (a) Southward-flowing river system catchments and their confluence during −130 m sea level regression during the Last Glacial Maximum. (b) Magnified view of the confluence and the reconstructed palaeo-river between the Buffalo and Nahoon river systems.

Analysis of molecular variance

The Cyt *b* AMOVA provided additional support for patterns of genetic structure among populations of *L. umbratus*. Of the four arrangements that were defined and tested, the arrangements according to palaeo-river systems and currently isolated river systems showed significant and similar ϕ_{CT} values, which was an indication of the genetic differentiation among the groups (ϕ_{CT} = 0.816 and 0.816, respectively) (Table 3.4). In the Orange vs south-western vs south-eastern systems and the Orange vs southward-flowing systems arrangement, less variation was attributed to the differences among the groups (ϕ_{CT} = 0.785 and 0.773, respectively) compared with the previously mentioned arrangements (Table 3.4).

Historical demography

The hypothesis of population expansion was not rejected only for the Orange River system population of *L. umbratus*. Significantly ($p < 0.05$) negative values for Tajima's *D* (−1.467), Fu's F_s (−3.005) (Table 3.5) and a unimodal mismatch distribution were observed (Fig. 3.9a). The Gamtoos population had a significantly negative Tajima's *D* (−1.559), but a non-significant negative Fu's F_s (−0.823) (Table 3.5) and a multimodal mismatch distribution (Fig. 3.9c). The Gourits River population had a non-significant Tajima's *D* (−1.112) and Fu's F_s (−0.339) (Table 3.5) and a unimodal mismatch distribution (Fig. 3.9b). The remainder of the southward-flowing populations showed no signs of expansion, but this may be due to low sample size. The SSD and Rag statistics for mismatch distribution were non-significant, thus did not support population expansion (Table 3.5).

TABLE 3.4. AMOVA of *Labeo umbratus* cytochrome *b* sequence data with four *a priori* hierarchical arrangements among and within populations. The variance, ϕ-statistic and *P*-values are presented. The analysis employed the Tamura and Nei (TrN) model of substitution (Tamura & Nei, 1993).

Hierarchical arrangement	Variance (%)	ϕ-statistic	*P*-value
Orange vs southward-flowing systems			
Among systems catchments	77.33	ϕ_{CT} = 0.773	0.002
Among localities within systems	11.38	ϕ_{SC} = 0.502	<0.001
Within all localities	11.28	ϕ_{ST} = 0.887	<0.001
Orange vs south-western vs south-eastern-flowing systems			
Among systems catchments	78.45	ϕ_{CT} = 0.785	<0.001
Among localities within systems	7.19	ϕ_{SC} = 0.334	<0.001
Within all localities	14.36	ϕ_{ST} = 0.856	<0.001
Currently isolated river systems			
Among contemporary systems	81.55	ϕ_{CT} = 0.816	<0.001
Among localities within systems	−0.35	ϕ_{SC} = −0.019	0.410
Within all localities	18.80	ϕ_{ST} = 0.812	<0.001
Palaeo-river systems			
Among Palaeo-river systems	81.58	ϕCT = 0.816	< 0.001
Among localities within systems	0.05	ϕSC = 0.003	0.378
Within all localities	18.37	ϕST = 0.816	<0.001

TABLE 3.5. Test statistics for neutrality (Tajima's D test and Fu's F_s test) and mismatch distribution statistics [sum of square deviations (SSD) and Harpending's raggedness index (Rag)] from analysis of *Labeo umbratus* cytochrome *b* sequence data. Values of $P < 0.05$ and statistic values highlighted in bold are significant.

Population	Neutrality tests				Mismatch distribution			
	Tajima's D		Fu's F_s		Sum of square deviations		Harpending's raggedness index	
	D	P-value	F_s	P-value	SSD	P	Rag	P
Orange	**−1.467**	0.032	**−3.005**	<0.001	<0.001	0.252	0.711	0.785
Gourits	−1.112	0.191	−0.339	0.150	0.331	0.088	0.400	0.213
Gamtoos	**−1.559**	0.043	−0.823	0.246	0.062	0.070	0.534	0.498
Sundays	0.549	0.706	−0.459	0.287	0.083	0.073	0.365	0.064
Bushmans	0.830	0.840	0.253	0.469	0.030	0.286	0.218	0.273
Great Fish	0.201	0.679	0.112	0.465	0.039	0.083	0.245	0.066
Keiskamma	0.220	0.631	−0.061	0.486	0.027	0.229	0.106	0.432
Buffalo	0.835	0.829	0.898	0.706	0.011	0.449	0.060	0.773
Nahoon	1.232	0.886	2.402	0.899	0.219	0.114	**0.635**	0.017

TABLE 3.6. *Labeo umbratus* cytochrome *b* estimated divergence times among lineages and populations based on net divergence under assumed 0.76% and 2.2% site^{-1} my^{-1} substitution rates.

Clades	Net divergence (%)	Divergence time (ya)	
		0.76% site^{-1} my^{-1}	2.2% site^{-1} my^{-1}
Orange vs southern	0.8	608 000	1,760,000
South-western vs eastern clades	0.2	152 000	440 000
Gourits vs Gamtoos	0.3	228 000	660 000

my = million years, ya = years ago

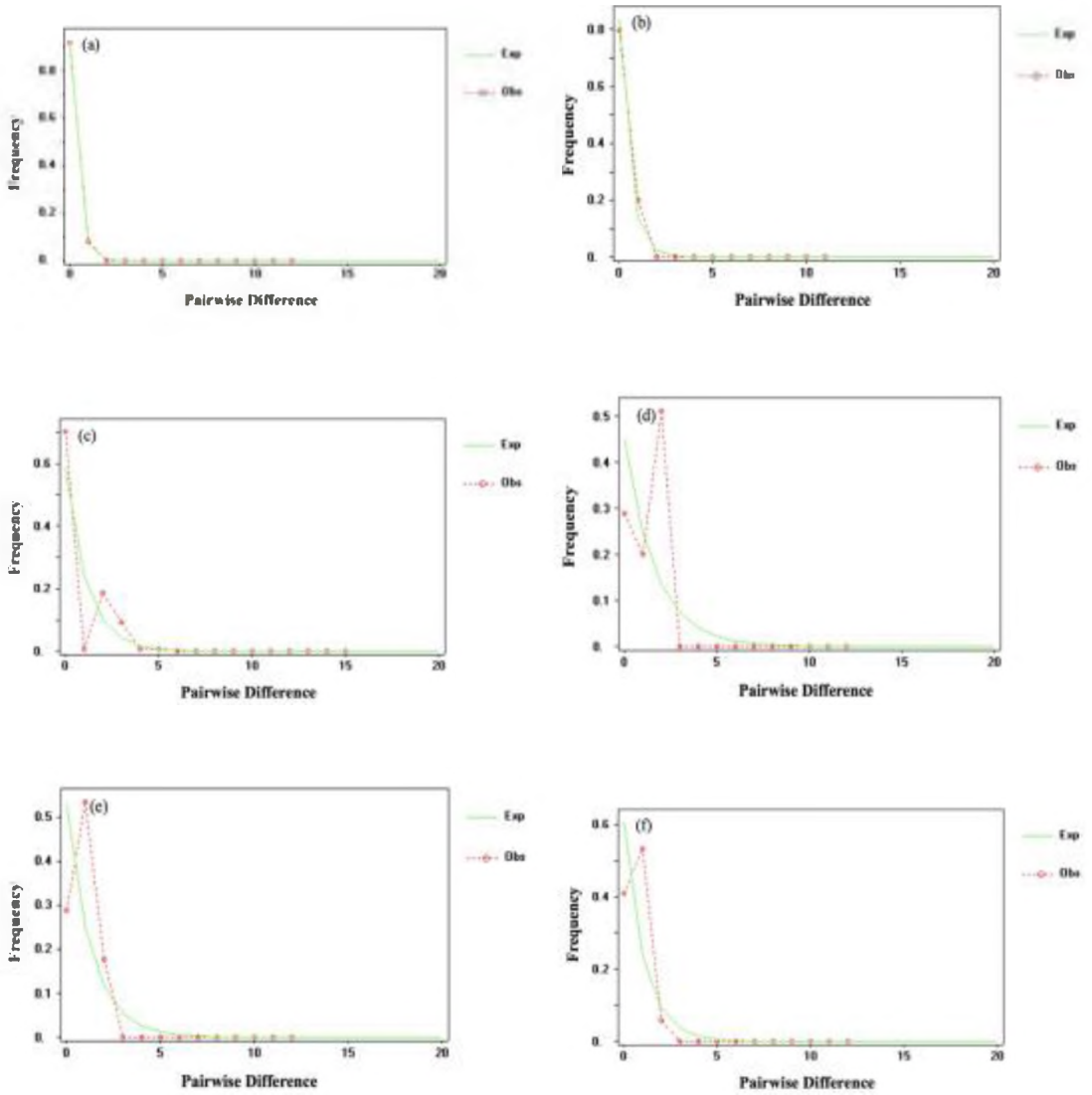


FIG. 3.9. *Labeo umbratus* cytochrome *b* mismatch distribution graphs showing unimodal distributions (population expansion) in the (a) Orange and (b) Gourits river systems, and multimodal distributions (constant population) in the (c) Gamtoos, (d) Sundays, (e) Bushmans, (f) Great Fish, (g) Keiskamma, (h) Buffalo and (i) Nahoon river systems. The red dotted line represents the observed mismatch distribution and the green solid line represents the expected mismatch distribution under a model of population expansion.

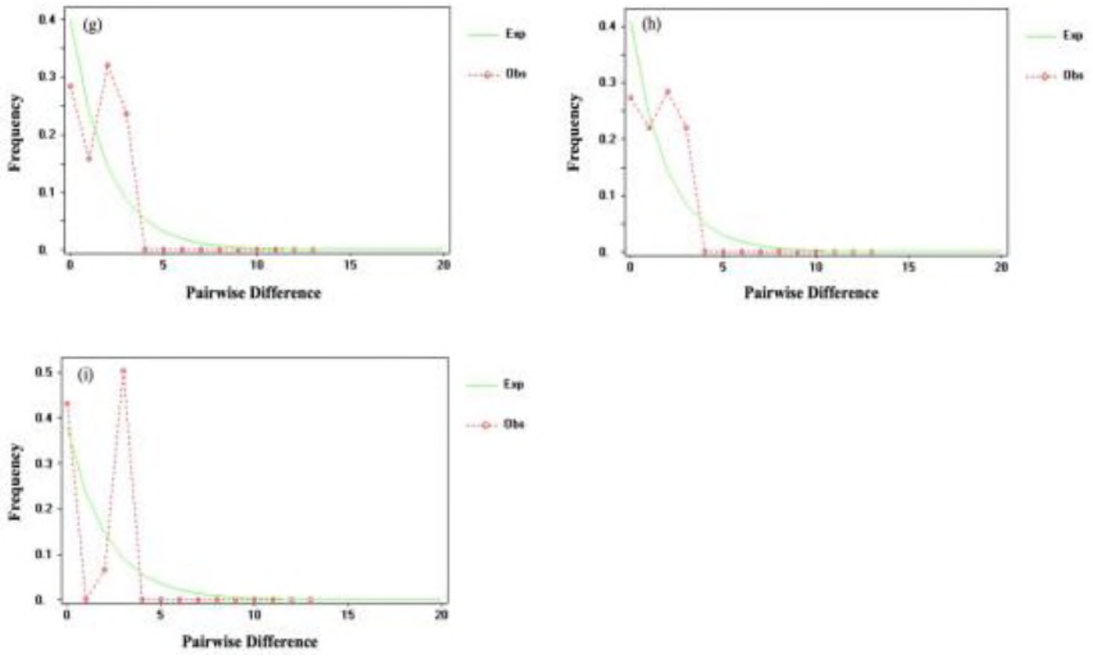


FIG. 3.9 (continued).

Divergence times

According to the net sequence divergence (N_D) among populations, the Orange and southward-flowing river system populations of *L. umbratus* diverged during the late Pleistocene epoch (range c. 608,000–1,760,000 ya) (Table 3.6). Subsequently, divergence between the Gourits and Gamtoos river system populations of *L. umbratus* followed (range c. 228,000–660,000 ya). Finally, divergence between the south-western and south-eastern river system populations of *L. umbratus* was indicated (range c. 152,000–440,000 ya). Haplotype 9 was considered to belong to the Gamtoos population, because it is more common in that river system and is an ancestral haplotype and was present only at a low frequency in the Sundays River system.

Nuclear S7 Ribosomal Protein Gene Intron 1

Sequence variation and diversity

A 609 bp fragment of S7 analysed from 102 individuals yielded 204 sequences (representing alleles from individuals when phased). Only five sites were variable and, of these, four were parsimony informative. Twenty-four sites showed alignment gaps. The HKY (Hasegawa, Kishino and Yano, 1985) model was selected as the optimal model with jModelTest. The S7 data set showed the following statistics: base frequencies A = 0.2957, C = 0.1634, G = 0.1993, T = 0.3415; proportion of invariable sites (I) = 0. Six alleles were detected (Fig. 3.10).

Genetic structure

Sequence divergence among all populations was extremely low (maximum 0.5% between alleles of different populations) (Table 3.7). Pairwise ϕ_{ST} values were significant ($p < 0.05$) and mostly low (0.001–0.391) among all populations of *L. umbratus*, which was indicative of low differentiation between populations, except between the Gourits and Gamtoos (0.999), and between the Great Fish and Nahoon (0.680) populations. These low divergence values may have been influenced by the small sample size and low variation of S7 intron 1.

The Orange and Keiskamma river system populations showed low nucleotide (0.021% and 0.032%, respectively) and allele diversity (0.128 and 0.190, respectively) (Table 3.7). In comparison, the remainder of the southward-flowing river system populations showed high allele (0.233–0.506) and low nucleotide diversity (0.038–0.153%).

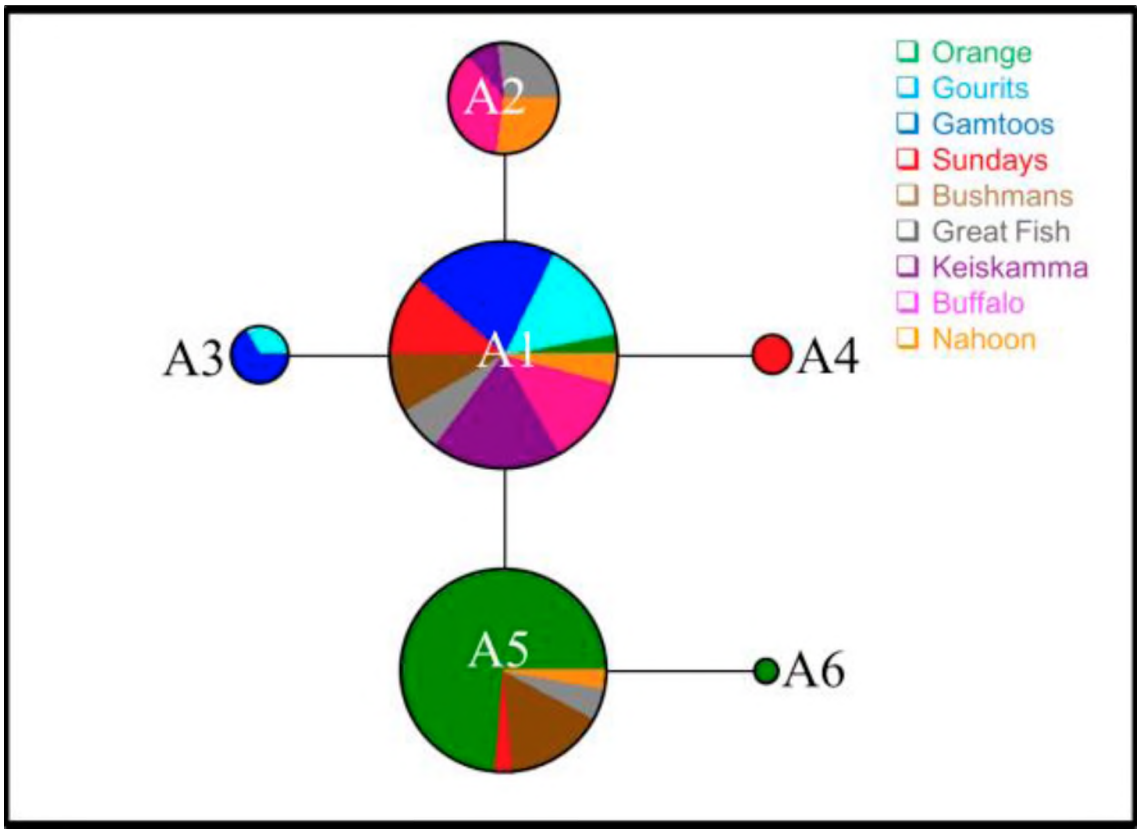


FIG. 3.10. *Labeo umbratus* nuclear S7 ribosomal protein gene intron 1 TCS parsimony network showing alleles (A1 to A6) association with river systems. The frequency of the alleles is indicated by the size of the circle.

Allele 1 was shared across populations of *L. umbratus* from all river systems (Fig. 3.10, Table 3.8). Allele 1 is the most likely representative of ancestral alleles because of its central position in the network. Allele 2 was shared only by populations of *L. umbratus* from the Great Fish, Keiskamma, Buffalo and Nahoon river systems. These systems are at the far east of the southward-flowing river systems. Allele 3 was shared only between populations of *L. umbratus* from the Gourits and Gamtoos rivers, which are in the far west of the southward-flowing river systems. Allele 4 was detected only in the Sundays River system. Allele 5 was carried mostly by individuals from the Orange River system but also shared with Sundays, Bushmans, Great Fish and Nahoon river systems. Allele 6 was detected in only one individual from the Orange River system.

TABLE 3.7. *Labeo umbratus* S7 ribosomal protein gene intron 1 pairwise ϕ_{ST} (above diagonal) and percentage sequence divergence (below diagonal) from comparisons among river systems using the Hasegawa–Kishino–Yano (HKY) distance method, and molecular diversity indices [nucleotide diversity (π; %) and allele diversity (h)] for each population. The south-western river systems are indicated in bold and the south-eastern river systems are underlined.

	River system	1	2	3	4	5	6	7	8	9	Molecular diversity indices	
											Nucleotide diversity π (%)	Allele diversity h
1	Orange		0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.021	0.128
2	**Gourits**	0.0–0.5		0.999	0.032*	0.001**	0.003**	0.180	0.003**	0.001**	0.038	0.233
3	**Gamtoos**	0.0–0.5	0.0–0.2		0.003**	0.001**	0.001**	0.045*	0.001**	0.001**	0.048	0.290
4	Sundays	0.0–0.5	0.0–0.3	0.0–0.3		0.003**	0.008**	0.006**	0.001**	0.002**	0.092	0.508
5	Bushmans	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3		0.008**	0.001**	0.001**	0.001**	0.083	0.505
6	Great Fish	0.0–0.5	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3		0.120	0.210	0.680	0.153	0.700
7	Keiskamma	0.0–0.5	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3		0.065	0.005*	0.032	0.190
8	Buffalo	0.0–0.5	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.2		0.391	0.087	0.505
9	Nahoon	0.0–0.5	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3	0.0–0.3		0.145	0.667

Significant ϕ_{ST} values are indicated with asterisks: * $P < 0.05$; ** $P < 0.005$.

TABLE 3.8. *Labeo umbratus* S7 ribosomal protein gene intron 1 allele frequencies and their distribution among sampled populations.

Alleles	Gariep Dam	Brak	Vaal	Gourits	Gamtoos	Sundays	Bushmans	Great Fish	Keiskamma	Buffalo	Nahoon
	34	12	14	16	24	16	20	16	20	20	12
1	-	-	3	14	20	11	8	6	18	12	4
2	-	-	-	-	-	-	-	6	2	8	6
3	-	-	-	2	4	-	-	-	-	-	-
4	-	-	-	-	-	3	-	-	-	-	-
5	33	12	11	-	-	2	12	4	-	-	2
6	1	-	-	-	-	-	-	-	-	-	-

Analysis of molecular variance

The S7 AMOVA showed that populations of *L. umbratus* were poorly structured in contrast to Cyt *b*. This was because the S7 nuclear gene evolves more slowly (Moore, 1995). Of the four arrangements that were defined and tested, the AMOVA results supported the Orange vs southward-flowing river systems arrangement (Table 3.9). This arrangement showed the highest differentiation among groups ($\phi_{CT} = 0.56$), followed closely by the palaeo-river and currently isolated systems structures ($\phi_{CT} = 0.55$ and $\phi_{CT} = 0.54$, respectively). The Orange vs south-western and south-eastern river systems arrangement showed the lowest among-group differentiation ($\phi_{CT} = 0.48$) (Table 3.9).

Historical demography

The hypothesis of population expansion was not rejected only for the population of *L. umbratus* from Orange River system. This population had a significant Fu's F_s (−2.066) and a non-significant ($P > 0.05$) negative Tajima's D (−1.191) (Table 3.10), and a unimodal mismatch curve (Fig. 3.11a), which all supported population expansion. The Gourits, Sundays and Keiskamma river systems populations also had a non-significant negative Tajima's D and non-significant Fu's F_s. All other populations were shown to have a unimodal mismatch curve (Fig. 3.11b–f). The mismatch statistics were non-significant, thus did not support a hypothesis of population expansion (Table 3.10). However, the low sample size may have reduced the power of the analysis to detect population expansion in the southward-flowing river systems.

TABLE 3.9. AMOVA of *Labeo umbratus* S7 ribosomal protein gene intron 1 sequence data with four *a priori* hierarchical arrangements among and within populations. The variance, ϕ-statistic and *P*-values are presented. The Hasegawa–Kishino–Yano (HKY) model of substitution was employed.

Hierarchical arrangement	Variance (%)	ϕ-statistic	*P* value
Orange vs southward-flowing systems			
Among systems catchments	55.49	ϕ_{CT} = 0.56	0.006
Among localities within systems	12.41	ϕ_{SC} = 0.28	<0.001
Within all localities	32.10	ϕ_{ST} = 0.68	<0.001
Orange vs south-western vs south-eastern-flowing systems			
Among systems catchments	48.36	ϕ_{CT} = 0.48	<0.001
Among localities within systems	12.86	ϕ_{SC} = 0.24	<0.001
Within all localities	38.78	ϕ_{ST} = 0.61	<0.001
Currently isolated river systems			
Among contemporary systems	54.74	ϕ_{CT} = 0.54	0.011
Among localities within systems	0.89	ϕ_{SC} = 0.02	0.019
Within all localities	44.37	ϕ_{ST} = 0.56	<0.001
Palaeo-river systems			
Among palaeo-river systems	54.89	ϕ_{CT} = 0.55	0.002
Among localities within systems	1.06	ϕ_{SC} = 0.02	0.142
Within all localities	44.05	ϕ_{ST} = 0.56	<0.001

TABLE 3.10. Test statistics for neutrality (Tajima's D test and Fu's F_S test) and mismatch distribution [sum of square deviation (SSD) and Harpending's raggedness index (Rag)] from analysis of *Labeo umbratus* S7 ribosomal protein gene intron 1 sequence data. Values of $P < 0.05$ and statistic values highlighted in bold are significant.

Population	Neutrality test				Mismatch distribution Demographic expansion			
	Tajima's D test		Fu's F_S test		Sum of square deviation		Harpending's raggedness index	
	D	P-value	F_S	P-value	SSD	P	Rag	P
Orange	**−1.191**	0.052	**−2.066**	0.010	<0.003	0.350	0.571	0.688
Gourits	−0.448	0.277	0.083	0.286	0.295	0.125	0.339	0.199
Gamtoos	0.139	0.753	0.578	0.404	0.246	0.144	0.261	0.263
Sundays	−0.189	0.398	0.176	0.317	0.017	0.234	0.170	0.345
Bushmans	1.430	0.952	1.409	0.682	0.231	0.070	0.255	0.089
Great Fish	1.262	0.888	13.584	1.000	0.171	0.096	0.190	0.345
Keiskamma	−0.592	0.250	10.256	0.998	0.052	0.180	0.728	0.649
Buffalo	1.430	0.949	19.339	1.000	0.511	0.000	0.755	0.915
Nahoon	0.822	0.808	11.513	1.000	0.215	0.092	0.258	0.268

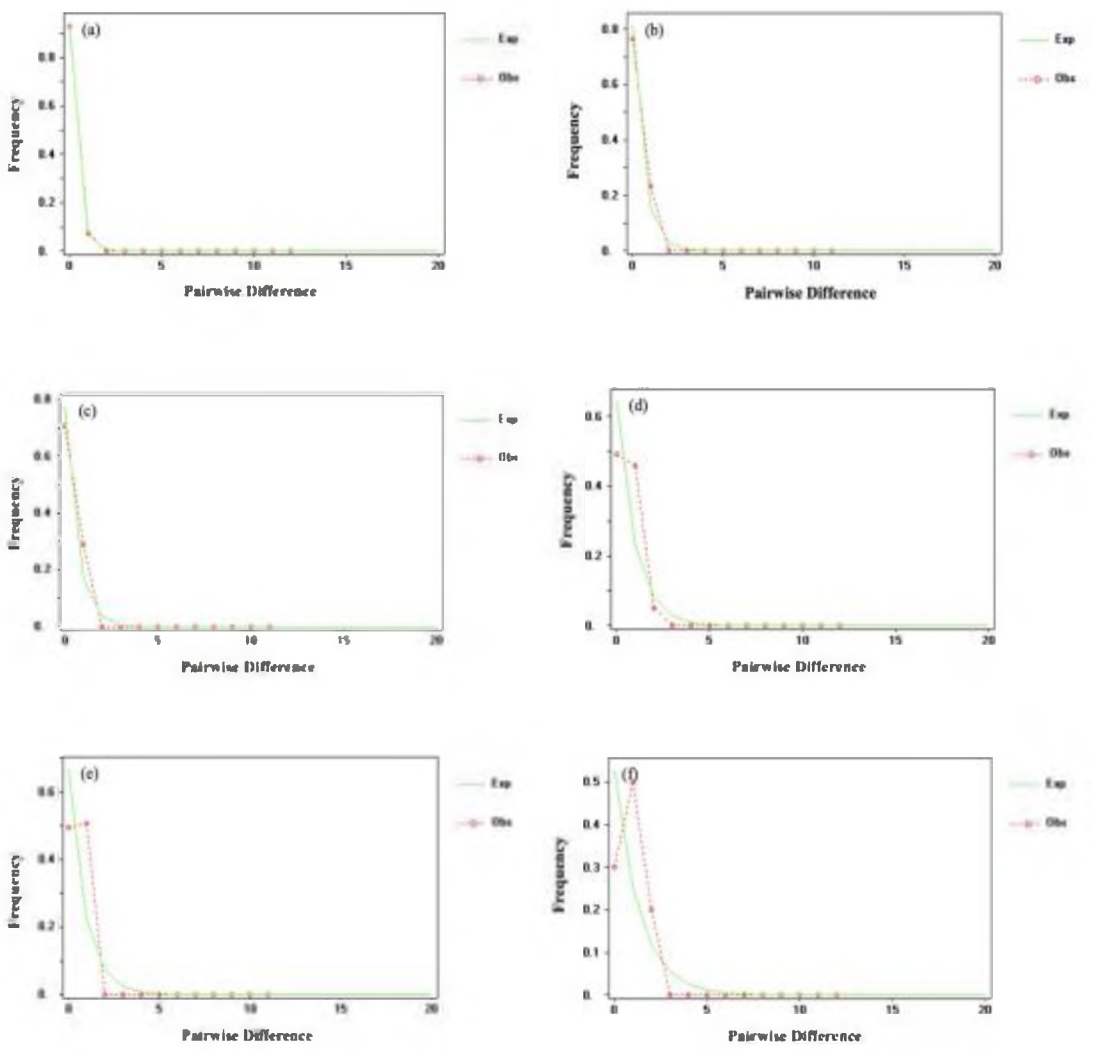


FIG. 3.11. *Labeo umbratus* S7 ribosomal protein gene intron 1 mismatch distribution graphs, showing unimodal (population expansion) in the (a) Orange, (b) Gourits, (c) Gamtoos, (d) Sundays, (e) Bushmans, (f) Great Fish, (g) Keiskamma, (h) Buffalo and (i) Nahoon river systems. The red dotted line represents the observed mismatch curve and the green solid line represents the expected mismatch curve.

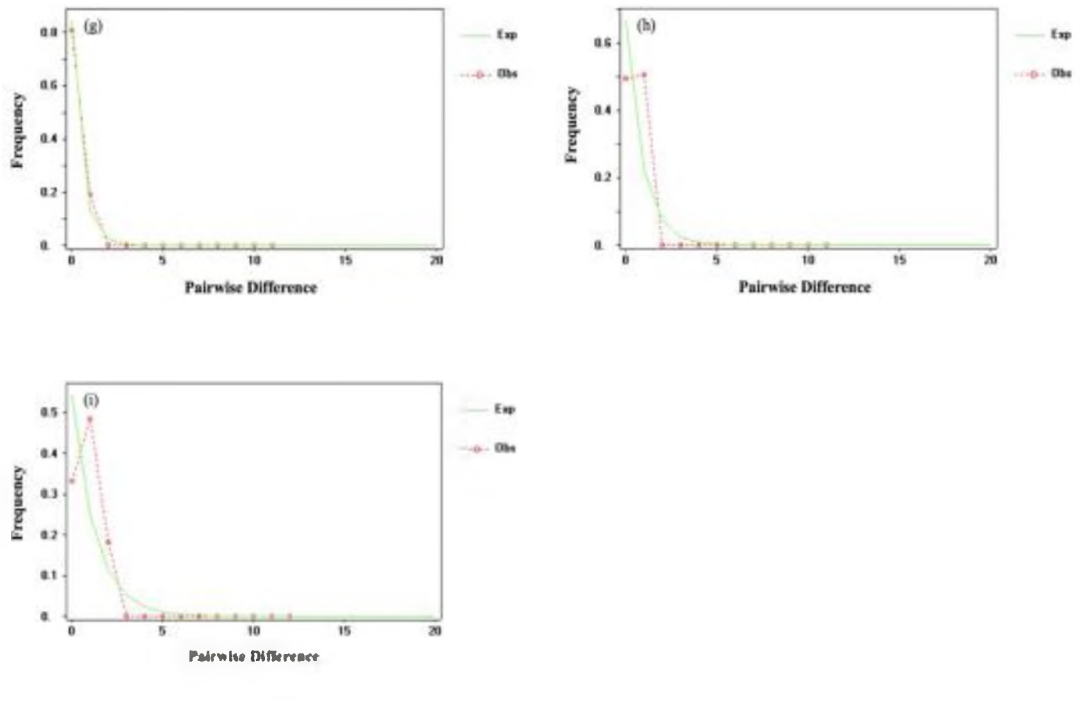


FIG. 3.11 (continued).

DISCUSSION

The results presented in this chapter demonstrated that considerable genetic structuring was detectable among the populations of *L. umbratus* across its distribution. This was mostly reflected in the mtDNA Cyt *b* data set. Two major clades were revealed, namely the Orange River and southward-flowing river systems (consisting of south-western and south-eastern subclades).

Orange and Southern lineages

The differentiation of the Orange and southward-flowing river system populations of *L. umbratus* was not surprising as the two populations are separated by the >1000 m high Great Escarpment (Wellington, 1955). Thus the two populations have been isolated for

a sufficiently long period to become genetically differentiated. The divergence between the two populations of *L. umbratus* was estimated to be in the Pleistocene epoch (608,000–1,760,000 ya) (Table 3.6), although the Great Escarpment is much older (180 mya) (Truswell, 1977). The population of *L. umbratus* may have found a means (e.g., at the lowest portion of the Great Escarpment) of breaching this barrier to disperse into the southward-flowing river systems. This suggestion conforms with the southward *Labeo* species migration hypothesis by Jubb & Farquharson (1965).

The mtDNA Cyt *b* sequence data revealed high percentage divergence (0.8–1.2%) (Table 3.3) and significantly high pairwise ϕ_{ST} values (Table 3.3) between the *L. umbratus* populations from the Orange River system and the southward-flowing river systems. The two populations were distinguished by five nucleotide differences in the Cyt *b* sequence and no haplotypes were shared (Fig. 3.7). These values are very low when compared with species-level differences for other cyprinids in southern Africa (Swartz *et al.*, 2007, 2009). In addition, the differentiation based on the nDNA S7 gene was not as strong as that for Cyt *b* (Fig. 3.8). Analysis of the S7 data showed that the population of *L. umbratus* from the Orange River shared two common, widespread, potentially ancestral alleles with the populations from the southward-flowing river systems (Fig. 3.8, Table 3.8). A similar scenario, whereby nDNA data do not strongly support patterns suggested by mtDNA data, was reported by Koblmüller *et al.* (2008) in a study of the Lake Tanganyika-endemic cichlid tribe Tropheini using two mtDNA gene regions (ND2 and control region) and AFLP markers. Koblmüller *et al.* (2008) attributed the differences in results between the two types of markers to ancient incomplete lineage sorting. Furthermore, mtDNA and nDNA genes evolve differently

(Moore, 1995; Morrison *et al.*, 2006). Given that mtDNA evolves faster than nDNA, mtDNA is more effective in detecting genetic divergence and, because nDNA evolves more slowly, the ancestral alleles persist in a population longer and are reflected as shared alleles.

Investigation of the Orange River population of *L. umbratus* indicated that it had undergone expansion in the recent past (indicated by significantly negative Fu's F_S and Tajima's D values, and a high unimodal mismatch distribution) (Table 3.5, Fig. 3.7a). The population probably experienced a bottleneck or founder event as indicated by the low haplotype and nucleotide diversities (Lin *et al.*, 2010). During the Plio-Pleistocene epochs, southern Africa experienced wet and dry conditions (Partridge, 1993; Cowling *et al.*, 2009) and the Orange River *L. umbratus* may have been restricted to refugia because of dry conditions, therefore enforcing a bottleneck or founder event on the surviving population. Subsequently, during wetter conditions (floods), the population may have expanded, occupying suitable habitat for feeding, shelter and colonisation, e.g. such as the Sak and Brak rivers (tributaries of the Orange River system), which are in close proximity to the Gamtoos River system. It was perhaps at this time that the ancestral southward-flowing river system population of *L. umbratus* dispersed to the southward-flowing river systems as the two populations diverged around this time (608,000–1,760,000 ya). Jubb (1964) also suggested that the Sak River may have played a role in linking the Orange River with the southward-flowing river systems. Jubb (1964) based his decision on the geological history of the region, with its numerous pans and low-graded rivers, such as the Sak River and the distribution of cyprinids. However, it is unclear in the present study how the southern river system invasion may have occurred.

Differentiation within the Southern Lineage

The mtDNA Cyt *b* data showed structuring among the populations from southward-flowing river systems. This is an indication of the duration of isolation among these river systems. This may be due to deeper incisions that are currently observed on the southern coastal systems (Cowling *et al.*, 2009). The populations showed high allelic diversity and low nucleotide diversity. According to guidelines from Grant & Bowen (1998) based on allelic and nucleotide diversity, southward-flowing river populations of *L. umbratus* experienced a bottleneck followed by rapid population growth and accumulation of mutations. The structuring shown by mtDNA (Table 3.3) was not evident in nDNA, as populations of *L. umbratus* were shown to share alleles, but a degree of differentiation was observed between the Gourits and Gamtoos, and the Great Fish and Nahoon rivers (Table 3.7). This may be an indication of an ancient connection between these systems as nDNA is a slow-evolving gene, or the effect of ancestral polymorphism/alleles and incomplete lineage sorting. Evidence of connections between currently isolated southward-flowing systems has been observed previously. The Sundays and Swartkops rivers were indicated to have been connected in the past because of the close relationship between these populations of *Pseudobarbus afer*, as do the Gamtoos, Kabeljous and Swart river systems, and modelling of the paleo-rivers supported the existence of a past link (Swartz *et al.*, 2007). As indicated by Craw *et al.* (2008), the connection between isolated river systems (such as the southward-flowing river systems) can be attributed to three processes.

Firstly, palaeo-river connections during periods of lower sea levels may have played a role in the currently isolated river system. When the sea level regressed to −130 m below the current sea level, neighbouring systems that share the common confluence could have been connected. Among the southern-flowing river systems in this study, only the Buffalo and Nahoon appear to share the same palaeo-river system (Fig. 3.6b). The hypothesis of the palaeo-river as an explanation of the distribution of diversity and patterns of differentiation was also supported by AMOVA partitions for both mtDNA and nDNA markers (Tables 3.4 and 3.10). The connection between the Buffalo and Nahoon rivers may have assisted populations of *L. umbratus* to move between the river systems. Other species indicated to have used this form of dispersal between isolated river systems are *Pseudobarbus afer* (Swartz *et al.*, 2007) and *Galaxias* sp. 'nebula' (Chakona *et al.*, 2013a) in the Cape Floristic Region. These findings indicate the important role that this mechanism played in fish dispersal among the southern African southward-flowing systems.

Secondly, exchange of water between adjacent river systems during flooding may also have connected currently isolated river systems. Southern Africa experienced wetter conditions during the Holocene (8,000–6,000 ya) (Partridge *et al.*, 1999). These conditions may have promoted flooding across the floodplains or low-lying areas connecting adjacent southward-flowing river systems and thus enabled dispersal of *L. umbratus* to other isolated river systems. This form of connection may have prevailed longer in the south-eastern river systems as they are indicated to show a close relationship (Fig. 3.7). Unmack (2013) suggested that only species within the vicinity of the connection between two river systems can utilise such a connection and the

individuals must also have a drive to move. *Labeo umbratus* has such a drive as it migrates during floods (Mulder, 1973; Jackson & Coetzee, 1982; Gaigher, 1984, Cambray, 1990) in search of suitable habitat for spawning, food, shelter and colonisation. Chakona *et al.* (2013a) inferred that *Galaxias* species might have used this form of dispersal as well, but *Galaxias* were aided by the fact that they could breathe air.

Lastly, during river capture, river tributaries are captured by adjacent drainages, thus allowing migrants to move from the captured river to the other river (Burridge *et al.*, 2007). Burridge *et al.* (2008), in a phylogeographic study of *Galaxias* in southern New Zealand, showed that species may have used this form of dispersal. However, this form of dispersal is not well documented in the southward-flowing river systems and is less likely to have aided *L. umbratus* to disperse as this species is not found in tributaries located at higher altitude in the mountains (Jubb, 1964; Skelton, 2001) where river capture is possible.

Gourits and Gamtoos populations of *L. umbratus*

The Gourits river system population of *L. umbratus* was indicated by mtDNA to be the only population in the southward-flowing river system that did not share alleles with the other populations in southward-flowing river systems. The mtDNA data set showed the population of *L. umbratus* from this system to have two private alleles, which is an indication of a long period of isolation from the other systems. This result is supported by divergence time estimates, which estimate the Gourits population of *L. umbratus* to have diverged from other southward-flowing river systems around 228,000-660,000 ya (Table 3.6). This event preceded the divergence between the south-western and south-eastern

lineages (range c. 152,000–440,000 ya). Analysis of nDNA, on the contrary, showed that the Gourits River population of *L. umbratus* shared one allele with a population from the Gamtoos river system and another allele with all populations from the southward-flowing river systems. The differentiation between the Gourits population of *L. umbratus* and other populations from southern river systems, as reflected in mtDNA, may be because of the currently deeper incisions of the river valleys (Cowling *et al.*, 2009) and the distance between the Gourits and Gamtoos river systems (Fig. 3.5) on the floodplain. The population of *L. umbratus* in the Gourits River system is indicated to have undergone a recent population bottleneck on account of the low allelic and nucleotide diversity (Grant & Bowen, 1998). The populations of *L. umbratus* from the Gourits and Gamtoos river systems might have been connected in the past and exchanged migrants. The connection between the latter two river systems was disrupted and thereafter became isolated. Support for this conclusion is that the two systems form a lineage and now contain unique haplotypes. The identification of unique lineages of *Pseudobarbus asper* and *P. tenuis* (Barnard, 1938) within the Gourits River system supports this hypothesis (Swartz *et al.*, 2009).

The population of *L. umbratus* from the Gamtoos River system did not share haplotypes with any of the populations from southward-flowing river systems except the Sundays River system population (Fig. 3.6). The population from the two river systems shared a haplotype that seemed to be ancestral (shared ancestral polymorphism) (Sousa & Hey, 2013) to the populations from the southward-flowing river systems. The population of *L. umbratus* from the Gamtoos River also contained three singleton private haplotypes, which is an indication of a prolonged period of isolation between the Gamtoos River

population and the two adjacent populations (from the Gourits and Sundays river systems). Swartz *et al.* (2007) also detected unique lineages of *Pseudobarbus* between the two river systems. Similar to the Gourits River above, the nDNA data set showed that the population of *L. umbratus* from the Gamtoos river system shared one allele with the Gourits population and another allele with all populations from the southward-flowing river systems. This finding was an indication of the limited ability of nDNA to detect intraspecific variation compared with that of mtDNA (this study and Moore, 1995).

Alien populations of *L. umbratus*

Only three mtDNA haplotypes (H10, H12 and H13) were detected in populations of *L. umbratus* from the Keiskamma, Buffalo and Nahoon rivers, but were absent in the other south-eastern river systems (Fig. 3.7). The sharing of alleles among populations of *L. umbratus* from the Keiskamma, Buffalo and Nahoon river systems may be because they are found in low-lying areas and share the same primary drainage basin, and also because of the palaeo-river that connected the Buffalo and Nahoon river systems (Fig. 3.8). The ϕ_{ST} values indicate that the Keiskamma, Buffalo and Nahoon populations have evolved together for a sufficiently long period to show differentiation from the other south-eastern populations of *L. umbratus*.

Thus, if there was any translocation to the Keiskamma, Buffalo and Nahoon river systems, it would not be from any of the other southward-flowing river systems (Gourits, Gamtoos, Sundays, Bushmans and Great Fish) that are differentiated but from the systems with which they share haplotypes. Jubb (1964) indicated that *L. umbratus* was already known to occur in the Keiskamma River, thus the populations of *L. umbratus*

from the Buffalo and Nahoon rivers could have been introduced from the latter river system. Alternatively, AMOVA of the mtDNA data set (Table 3.4) indicated that the population of *L. umbratus* from the Buffalo River was most likely connected with the Nahoon population via a palaeo-river during the period of lower sea levels 18,000 ya (see Fig. 3.8). Therefore, this finding indicates that the populations of *L. umbratus* from the Buffalo and Nahoon river system might not have been introduced, as hypothesised by Jubb (1964).

Conservation implications

According to Waples (1991), distinct populations of species need to be protected in order to preserve their evolutionary significance. Ryder (1986) suggested that such populations should be designated evolutionary significant units (ESUs). The term ESU recognises a population that has been historically isolated and has the potential of being distinct (Moritz, 1994). To be considered as an ESU, a population must meet two criteria. The population has to be reciprocally monophyletic for mtDNA alleles and show significant divergence of allele frequencies at nuclear loci (Moritz, 1994). If a population does not meet these criteria, it can still be considered for conservation as a management unit (MU). Management units differ from ESUs in that they may not be reciprocally monophyletic for mtDNA but still be divergent in allele frequency.

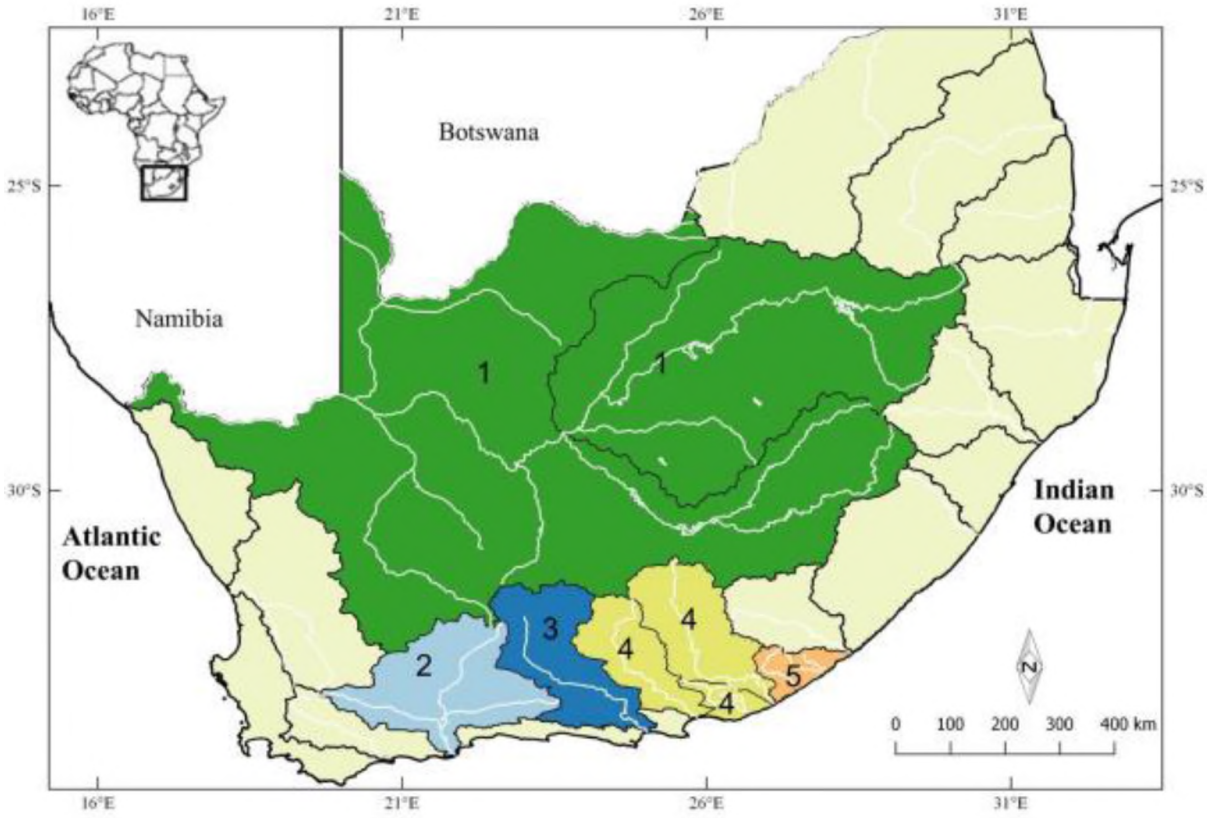


FIG. 3.12. Distribution of evolutionary significant units (nos. 1 and 2-5) and management units (nos. 2, 3, 4 and 5) of *Labeo umbratus* from the southward-flowing river systems.

The southern lineage of *L. umbratus* (Fig. 3.12, nos. 2–5) was shown to be reciprocally monophyletic (Figs. 3.6 and 3.7) and to have been isolated for a sufficiently long period (c. 608,000–1,760,000) (Table 3.6) to have differentiated from the Orange River lineage (Fig. 3.12 no. 1), but did not show significant divergence of alleles frequencies at nuclear loci due to ancestral polymorphism. Therefore, the two populations could be identified as ESUs. The southern lineage can be further divided into two divergent sublineages (Fig. 3.12, no2+3 and 4+5) that potentially may be identified as ESUs, but are not reciprocally monophyletic because the two lineages share one ancestral haplotype (H9; ancestral polymorphism) (Fig. 3.7). The two sublineages may be further divided into two MUs (Gourits and Gamtoos)(Fig. 3.12, no 2 and 3 respectively) and

(Sundays+Bushmans+Great Fish and Keiskamma+Buffalo+Nahoon) (Fig. 3.12, no 4 and 5 respectively) based on their haplotype frequencies (Table 3.2). The genetic integrity of populations in the Sundays and Great Fish river systems is under threat due to ongoing hybridisation in these systems (see Chapter 4), thus, translocation from these systems should be prohibited.

Conclusion

This study contributed to ongoing research into unravelling the history of the southward-flowing river systems in South Africa in indicating the roles played by the Great Escarpment (geological changes), which acts as a physical barrier between the Orange and southward-flowing river systems, the palaeo-river connection of the Buffalo and Nahoon river systems (sea-level changes) and floods (climatic change) in the low-lying areas between river systems. The results show that phylogeographic studies may aid detection of unique populations and clarify the possible processes that may have assisted the dispersal of species to areas where they were perceived not to have occurred due to inadequate sampling. In addition, a unique southern lineage was identified that may be threatened by hybridisation (the topic of Chapter 4).

CHAPTER FOUR: HYBRIDISATION OF TWO SOUTH AFRICAN ENDEMIC FISHES TRIGGERED BY ORANGE-FISH INTER-BASIN WATER TRANSFER SCHEMES

INTRODUCTION

Labeo umbratus (Smith 1841) occurs naturally in the Orange River and in the southern-flowing river systems (Gourits, Gamtoos, Sundays, Bushmans, Great Fish and Keiskamma, Buffalo and Nahoon). The Orange River population is genetically distinct from the populations from southern-flowing rivers systems (see Chapters 2 and 3), which were considered to represent four discrete genetic Management Units (Gourits, Gamtoos, Sundays+Bushmans+Great Fish and Keiskamma+Buffalo+Nahoon) (see Chapter 3). In the Orange River, *Labeo umbratus* occurs together with *Labeo capensis* (Smith 1841). The two species are closely related (see Chapter 2), but differ in several morphological characters (Reid, 1985). *Labeo capensis* (Fig. 4.1a) has larger scales (lateral line scales = 42–50), a larger dorsal fin, a smaller head and a terminal mouth. *Labeo umbratus* (Fig. 4.1b) has smaller scales (lateral line scales = 53–68), a smaller dorsal fin, a bigger head and a subterminal mouth (Reid, 1985; Skelton, 2001). Where they co-occur naturally, the two species occupy different ecological niches (Jubb, 1964; Reid, 1985; Chapter 2) with *L. capensis* preferring fast-flowing waters and *L. umbratus* preferring standing or gently flowing water; the ecological separation may reflect the differences in mouth form and position (Gaigher & Bloemhof, 1975). However, the two species utilise similar breeding grounds and spawn concurrently under similar environmental conditions (Tomasson *et al.*, 1984; Tweddle & Davies, 1997).

FIG. 4.1. (a) *Labeo capensis* from Kanoneiland (middle Orange River) and (b) *L. umbratus* from Brak River (tributary of the Orange River). Photographs by N. Mazungula, SAIAB.

The similarity in spawning conditions could render this species pair susceptible to hybridisation, and other *Labeo* species have been shown to hybridise under natural and aquaculture conditions (Shah *et al.*, 2011; Allu *et al.*, 2014). The distributions of the Orange River *L. umbratus* and *L. capensis* have been altered by the development of inter-basin water transfers schemes (IBWTs) between the Orange and the Great Fish river systems (IBWT 1 in Fig. 4.2) and between the Great Fish and the Sundays river systems (IBWTs 2 and 3 in Fig. 4.2), which facilitated the movement of *L. capensis* and the Orange River lineage of *L. umbratus* from the Orange River to the Great Fish and the Sundays river systems. In this chapter the threat of hybridisation posed by these

transfers to the *L. umbratus* in the Sundays+Bushmans+Great Fish MU was investigated.

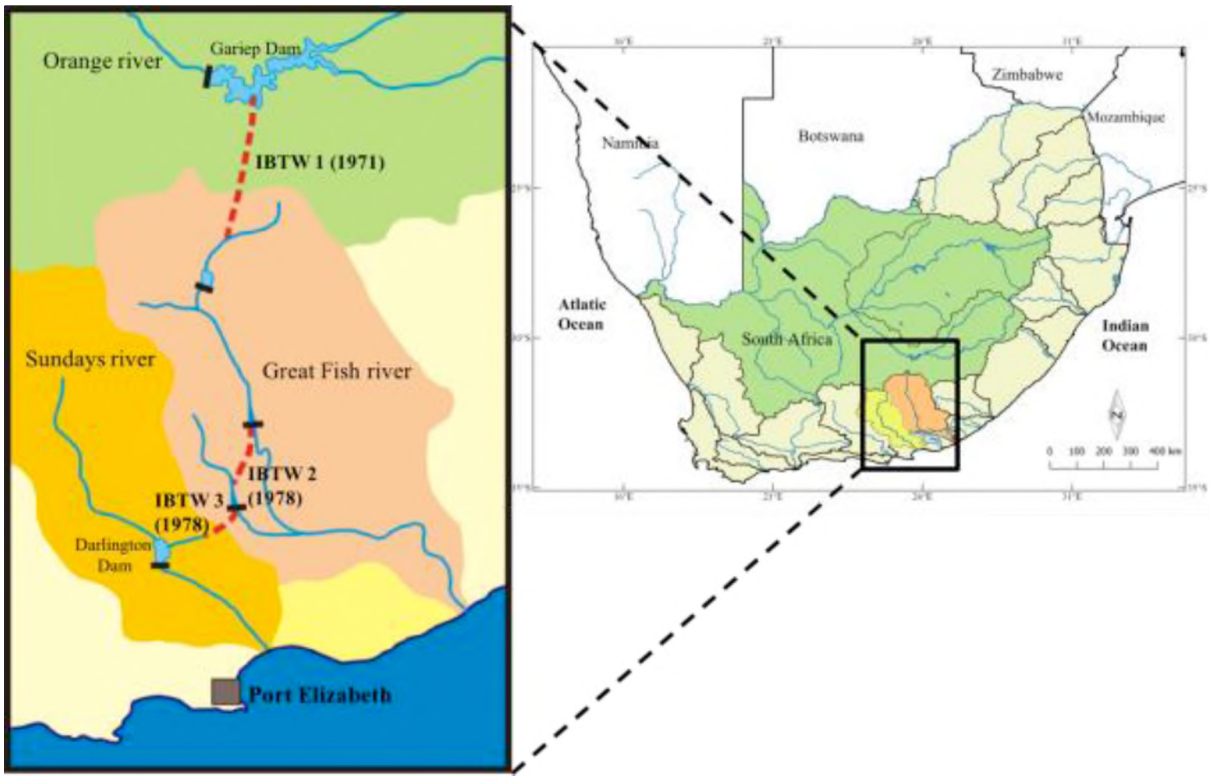


FIG. 4.2. Map of South Africa showing the Orange River Basin (green) in relation to the Great Fish River (orange) and Sundays River (yellow) basins. Inter-basin water transfer schemes (IBWT) and the river systems involved are shown in the left hand panel. IBTW 1 = Orange-Fish tunnel, IBTW 2 and 3 = Cookhouse tunnels. The dates in brackets represent the time the IBWTs became operational.

Inter-basin Water Transfer Schemes and their Impact on River Systems

Inter-basin water transfer schemes are used globally as mechanisms for solving water supply problems in arid and semi-arid environments, as well as in areas of water demand for human population growth and needs (Davies *et al.*, 1993; Snaddon *et al.*, 1999; Shao & Wang, 2003; Gupta & van der Zaag, 2008). Problematically, IBWTs breach historical boundaries between watersheds and facilitate the unintentional

movement of biota across geological barriers (Grant *et al.*, 2012). Examples include the introduction of the sea lamprey *Petromyzon marinus* Linnaeus 1758 into Lake Erie via the Welland Canal (built for shipping) from Lake Ontario in North America (Sullivan *et al.*, 2003), and the introduction of *Galaxias* species across catchments in Australia (Lintermans, 2004). As is the case with most introductions of non-native species (see Ellender & Weyl, 2014), IBWT-mediated fish introductions can have severe consequences for the native organisms, including increased competition with and predation on native biota by non-native fish, the spread of diseases, and homogenisation of fish fauna (Almeida *et al.*, 2013; Kadye & Booth, 2013). In addition, IBWTs may also facilitate hybridisation because they break geographic divides between allopatric species.

Inter-basin Water Transfer Schemes in South Africa

South Africa experiences water shortages because it is situated in the drought belt of the Southern Hemisphere and receives low rainfall (an annual average of 445 mm) (Van Robbroeck, 1979). A number of dams and IBWTs have been constructed to address this problem (Snaddon *et al.*, 1999). The first dam built was the Rand Water Board Barrage (1903), followed by the Vaal and Bloemhof dams successively in the Vaal River (Van Robbroeck, 1979). In the late 1960s, it was evident that the existing system of dams would not meet water demand, particularly in dry regions where natural precipitation did not meet water requirements for human consumption. As a result, several IBWTs were built to transfer water from the larger rivers to drier areas (Gupta & Van der Zaag, 2008). The longest IBWTs (Orange-Fish) were completed in 1975 (Cambray & Jubb, 1977; Van Robbroeck, 1979). These IBWTs connect the west-flowing Orange river

system with the east-flowing Great Fish River and Sundays River drainages (Cambray & Jubb, 1977), first through the Orange-Fish tunnel (from the Orange River to the Great Fish River, completed in 1975) and then the Cookhouse tunnel (from the Little Fish River to the Sundays River, completed in 1978). These IBWTs facilitated the transfer of fishes and insects native to the west-flowing Orange River to the east-flowing Great Fish and Sundays rivers (Cambray & Jubb, 1977; O'Keefe & De Moor, 2006; Woodford *et al.*, 2013). Five fishes [*Austroglanis sclateri* (Boulenger 1901), *Clarias gariepinus* (Burchell 1822), *Labeobarbus aeneus* (Burchell 1822), *Labeo capensis* and *Labeo umbratus*) used this IBWT as an introduction pathway (Laurenson & Hocutt, 1985; Weyl *et al.*, 2009; Kadye & Booth, 2013; Woodford *et al.*, 2013). This chapter focuses on the genetic impact of the introduction of *L. capensis* and the Orange lineage of *L. umbratus* via Orange-Fish IBWTs on the southern lineage of *L. umbratus*.

Hybridisation

Hybridisation is fairly common in freshwater fish species (Freyhof *et al.*, 2005; Bolnick, 2009). According to reviews by Argue and Dunham (1999) on fish hybrid fertility, introgression and backcrossing, and Scribner *et al.* (2001) on freshwater fish hybridisation, the highest prevalence of hybridising species pairs is in the family Cyprinidae. Hybridisation can occur between different species (interspecific hybridisation) or between populations of the same species (intraspecific hybridisation). Natural interspecific hybridisation is common in fish, has influenced the evolution of many animals and plants, and is a potential source of genetic variation and evolutionary novelty (Allendorf *et al.*, 2001; Arnold & Martin, 2010; Abbott *et al.*, 2013). Inter-specific hybridisation can also be facilitated by anthropogenic events and could have

severe consequences (e.g. extinction), or threaten or endanger native species (Allendorf *et al.*, 2001). According to Allendorf *et al.* (2001), it is debatable whether there should be concern with intra-specific hybridisation, as populations of the same species share a similar genetic makeup which could increase the fitness of populations by introducing new genetic variation, and could have a positive effect on the adaptive potential of a population. Allendorf *et al.* (2001), however, also argued that because some populations adapt to their surrounding environment, local adaptation of such populations could be lost if intra-specific hybridisation occurs. An example of such adaptation is that of bull trout [*Salvelinus confluentus* (Suckley 1859)], which has mixed migratory behaviours that have been shown to be linked to periodic fire disturbances (Rieman & Clayton, 1997). The bull trout migrate out of wildfire areas during fires and migrate back afterwards (Rieman *et al.*, 1997). Loss of this adaptation could result in decline or extinction of this species.

Species that hybridise tend to be closely related (Allendorf *et al.*, 2001). They can hybridise naturally due to range overlap (e.g., *Barbus canis* × *Barbus longiceps*; Fishelson *et al.*, 1996), limited spawning habitat (e.g., *Phoxinus eos* × *Phoxinus neogaeus*; Goddard & Dawley, 1990), or because of anthropogenic processes such as aquaculture (e.g., *Hypophthalmichthys molitrix* × *Aristichthys nobilis*; Mia *et al.*, 2005), species translocation or introduction (e.g., *Oreochromis mossambicus* × *Oreochromis niloticus*; Firmat *et al.*, 2013), and loss or alteration of habitat (*Gila cypha* × *Gila elegans*; Douglas *et al.*, 1998). Anthropogenic events are considered to be the main drivers of recent hybridisation (Allendorf *et al.*, 2001; Muhlfeld *et al.*, 2009). Allendorf *et al.* (2001) recognised three final stage "hybrid types" of human-mediated

hybridisation: hybridisation without introgression, which happens because F_1 hybrids are sterile; widespread introgression (hybrids interbreeding with each other or with parental species), which results in a hybrid swarm where individuals of the parental species persist (e.g., somewhere in the river tributaries); and complete admixture, which results in a hybrid swarm where few, if any, pure populations still exist.

Reports of *Labeo* Hybridisation

Labeo capensis and *L. umbratus* have been reported to hybridise in the Hardap impoundment in Namibia (Gaigher & Bloemhof, 1975; Van Vuuren *et al.*, 1989, 1990). Several methodological issues with these studies, however, cast some doubt on the claimed ability to distinguish between the two species and their hybrids based on morphology and allozymes in these studies. Van Vuuren *et al.* (1989) concluded that hybrids could not be identified based on morphology alone. This finding was supported by a review of the different methods of detection and genetic interpretations of natural hybridisation and introgression in fishes (Campton, 1987). It is not clear, however, how Van Vuuren *et al.* (1989) identified pure versus hybrid individuals *a priori*, because they did not sample localities that had only pure populations and based their assessment solely on specimens from the Hardap impoundment. Gaigher & Bloemhof (1975) also mentioned that hybrids between the two species were not confined to Hardap impoundment, because there were reports of *L. capensis* × *L. umbratus* hybrids from the Caledon River (citing an Orange Free State Nature Conservation report from 1972/73) and from Lake Gariep (citing a personal communication by Dr Kas Hamman). These claims could not be verified, however, because no voucher specimens were cited. Voucher specimens act as proof of identity and are therefore an important tool for other

researchers to verify results (Funk, 2005). The shortcomings in these early works indicate that a more comprehensive study on the potential for hybridisation between these two species is needed.

The introduction of *L. capensis* and Orange River *L. umbratus* to the Great Fish and Sundays river systems via the Orange-Fish and Cookhouse tunnels has also raised concerns about the genetic integrity and introgressive hybridisation of unique genetic lineages of *L. umbratus* (see Chapter 3) into the southern-flowing river systems (Cambray & Jubb, 1977; Laurenson & Hocutt, 1985; Laurenson *et al.*, 1989). Given the reported hybridisation of Orange River *L. capensis* and *L. umbratus* in the Hardap impoundment (Gaigher & Bloemhof, 1975; Van Vuuren *et al.*, 1989, 1990), it is reasonable to hypothesise that the species may interbreed in the Great Fish and Sundays river systems. To date, however, no study has presented definitive evidence for hybridisation and/or introgression between the two species.

Aims and Objectives

The current chapter assessed whether *L. capensis* and the Orange River and southern lineages of *L. umbratus* have hybridised in the Sundays and Orange river systems. Three hypotheses were evaluated: (1) that genetically separable groups are also separable morphologically; (2) that hybrids carry introgressed genes of the more abundant species; and (3) that F_1 hybrids are morphologically intermediate to the parental species. Samples were sequenced for mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) genes. The molecular results were compared with morphological characters analysed using linear morphometrics and meristics. First, the different

methods were tested for their ability to identify individuals as either Orange *L. capensis*, Orange *L. umbratus* or southern systems *L. umbratus* in putative pure populations in order to distinguish reference pure populations. Second, individuals from putative hybrid populations were analysed to classify them as being of hybrid origin or potentially pure.

MATERIALS AND METHODS

Sampling

A total of 218 genetic samples and associated voucher specimens collected previously (Ramoejane, 2010) from river systems associated with the Orange-Fish IBWTs and two neighbouring river systems (Bushmans and Keiskamma) were used in this study (Fig. 4.3, Table 4.1). Populations were classified *a priori* as putatively pure or putatively hybridised based on the presence or absence of confirmed prior records of hybridisation.

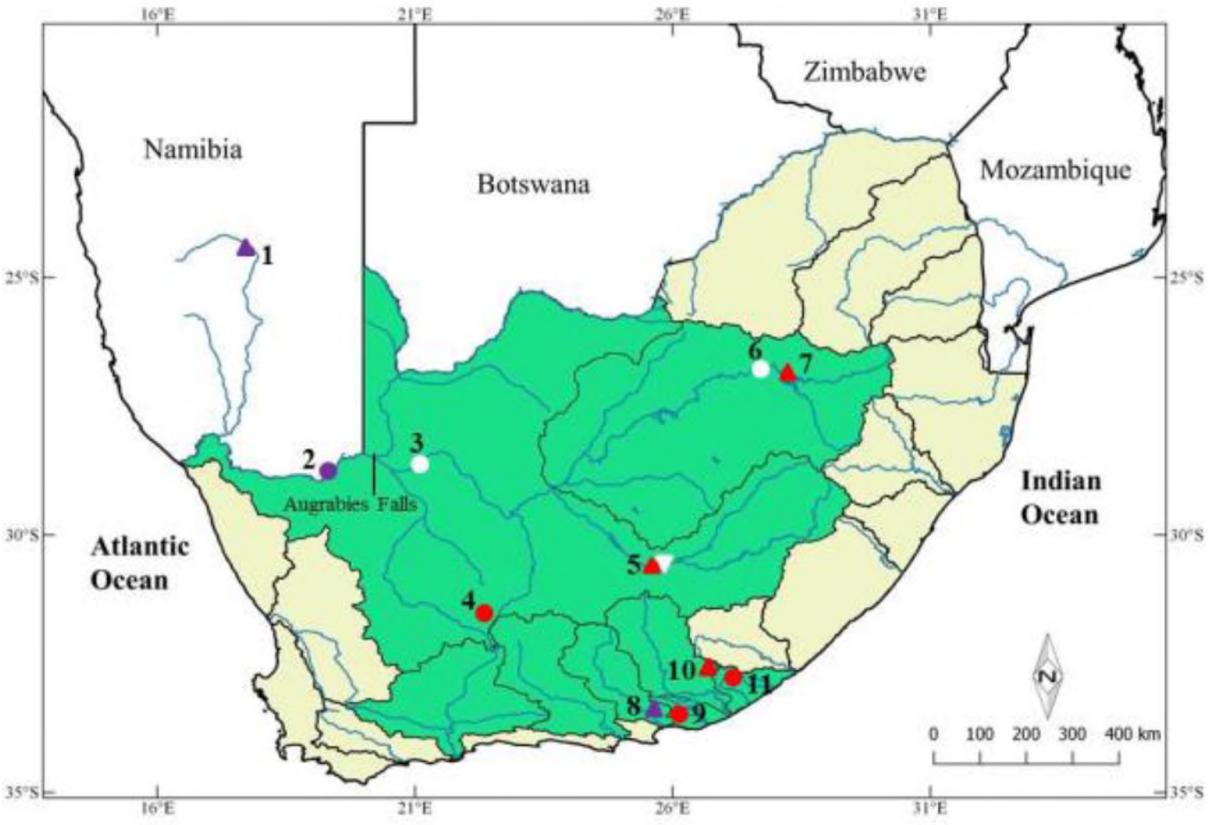


FIG. 4.3. Map of Southern Africa showing the provenance of samples collected in impoundments (triangles) [1-Hardap (n = 45), 5–Gariep (n = 48), 7–Vaal (n = 10), 8–Darlington (n = 25), 9–Slagboom (n = 8) and 10–Kat River (n = 16)] and rivers (circles) [2–Onseepkans (n = 10) on the Orange, 3-Kanoneiland (n = 10) on the Orange, 4–Brak (n = 10), 6–Vaal (n = 10), 9–Bushmans (n = 10), 11–Keiskamma (n = 12)]. White symbols indicate potentially pure *L. capensis*, red symbols indicate *L. umbratus* and purple symbols indicate potential hybrid zones.

TABLE 4.1. List of collection localities showing the number of fish specimens (of both study species and hybrid populations) analysed and the locality coordinates.

Locality	River system	Species	Number of specimens analysed	Coordinates	
				Latitude S	Longitude E
Orange River population (putative pure)					
Kanoneiland	Middle Orange	*L. capensis*	10	28° 38′ 05.7″	21° 05′ 20.3″
Lake Gariep	Upper Orange	*L. capensis*	26	30° 38′ 38.2″	25° 33′ 50.9″
Vaal River	Orange-Vaal	*L. capensis*	10	26° 45′ 57.6″	27° 40′ 56.9″
Orange River population (putative pure)					
Brak	Middle Orange	*L. umbratus*	10	31° 32′ 26.4″	22° 20′ 35.0″
Lake Gariep	Upper Orange	*L. umbratus*	26	30° 42′ 84.5″	25° 43′ 47.3″
Vaal impoundment	Orange-Vaal	*L. umbratus*	10	26° 51′ 58.9″	28° 10′ 14.3″
Southern systems population (putative pure)					
Kat River impoundment	Great Fish	*L. umbratus*	16	32° 33′ 46.5″	26° 46′ 43.0″
Slagboom impoundment	Sundays	*L. umbratus*	8	33° 22′ 31.1″	25° 40′ 45.4″
Amakhala Game Reserve	Bushmans	*L. umbratus*	10	33° 31′ 02.5″	26° 07′ 29.2″
Keiskamma	Keiskamma	*L. umbratus*	12	32° 41′ 12.7″	27° 09′ 09.08″
Potential hybrid populations					
Onseepkans	Lower Orange	Potential hybrids	10	28° 44′ 14.5″	19° 18′ 14.4″
Hardap impoundment	Lower Orange	Potential hybrids	45	24° 28′ 11.3″	17° 47′ 51.9″
Darlington impoundment	Sundays	Potential hybrids	25	33° 10′ 82.2″	25° 07′ 93.0″

Putative pure populations occurred in rivers or impoundments known not to be affected by IBWTs, such as the Kat River impoundment (52 m high Dam) in the Great Fish River system and the Slagboom impoundment in the Sundays River system, which are upstream of the entry point of the tunnels and were built before construction of IBWTs began (1969 and 1955 respectively). Conversely, putative hybrid populations inhabited the Hardap and Darlington impoundments. The Hardap impoundment forms part of the lower Orange River (from below Augrabies Falls to the Orange River estuary) and therefore the lower Orange River is also a potential hybrid zone.

DNA Extraction, Amplification and Sequencing

DNA extraction followed the same protocol described in Chapter 2. The cytochrome *b* (Cyt *b*) (see Chapter 2) and nuclear S7 ribosomal intron 1 (S7) (see Chapter 3) regions were amplified in polymerase chain reactions (PCRs) following the procedures stipulated in the respective chapters. Only 160 nuclear S7 samples were analysed out of the total of 218 samples because of difficulties with amplification.

Sequence Variation and Allele Distribution

DnaSP 5.10 (Rozas & Librado, 2009) was used to phase S7 sequences and to identify the unique haplotypes and alleles of Cyt *b* and S7. Invariable and variable sites were identified with the same software. The software TCS 1.21 (Clement *et al.*, 2000), which determines parsimony networks based on 95% confidence of connections among alleles (Templeton *et al.*, 1992), was used to determine genealogical relationships among the sequences.

Phylogenetic Differentiation and Hybrid Identification

Prior to phylogenetic reconstruction, models of nucleotide substitution that best fit the Cyt *b* and S7 data sets were selected from 1624 models with the Akaike information criterion (AIC) (Akaike, 1974) in jModelTest 2.1.7 (Guindon & Gascuel, 2003; Darriba *et al.*, 2012). Using the same software, base frequencies, Ti:Tv ratio or substitution rate matrix, proportion of invariable sites (*I*) and the α value of the gamma distribution (rate variation among sites) were estimated. These parameters were used to determine genetic distances among alleles using maximum likelihood with MEGA 6 (Tamura *et al.*, 2013). Gaps in the S7 gene were treated as missing data.

Mitochondrial Cyt *b* and nuclear S7 maximum likelihood phylogenetic trees were constructed in order to explore similarities among the populations and their association with individuals from putative hybrid zones. Populations were considered different if they formed reciprocally monophyletic clades. In the S7 data, individuals from pure populations would have fixed differences at each polymorphic locus, whereas potential F_1 hybrid individuals would be heterozygous at each locus where the species exhibit fixed differences, as they would inherit alleles from both parental species. The F_2 generation or backcrosses would potentially show different patterns of heterozygosity in different individuals.

Hybrid Identification from Genetic Data

The information collected from each gene was compared for each individual to identify whether they were possible hybrids. As mtDNA data were available for a greater number of individuals compared with the nDNA data, only individuals that were

analysed for nDNA were used in the comparison. Identifications of individuals on the basis of the nDNA data were compared with identifications based on the mtDNA data. F_1 individuals were identified as described in the preceding section, whereas individuals identified as pure from the nDNA data and possessing the mtDNA of the other species were considered to be possible F_2/backcross hybrids.

Discrimination of Pure and Hybrid Individuals from Morphological Characters

All of the 218 voucher specimens were measured to the nearest 0.1 mm for 33 morphological characters (Table 4.2, Fig. 4.4) using dial callipers. The measurements were size-standardised using Burnaby's Allometric method implemented in PAST 2.11 (Hammer *et al.*, 2001), which removes allometric size-dependent shape variation from the log-transformed data by projecting the data set into space orthogonal to the first principal component of the pooled covariance matrix (following Reist, 1985; Elliott *et al.*, 1995). Four meristics characters were counted (Table 4.3).

The size-standardised morphometric data and meristic data were analysed separately using a covariance matrix and correlation matrix, respectively, by principal component analysis (PCA) with PAST 2.11 (Hammer *et al.*, 2001) to reduce dimensionality and visualise the pattern of variation among individuals. Canonical variates analysis (CVA) based on multivariate analysis of variance (MANOVA) implemented in PAST 2.11 were used (for morphometrics and meristics analyses) to test the diagnosability and statistical distinctiveness of the three groups of putative pure individuals representing Orange River *L. capensis*, Orange River *L. umbratus* and *L. umbratus* populations in the southern systems.

TABLE 4.2. Characters used in the morphometric analysis and their abbreviations (codes).

No.	Characters	Codes
1	Standard length	SL
2	Dorsal fin length	DF
3	Dorsal fin base length	DFB
4	Caudal peduncle depth	CDP
5	Posterior dorsal fin to dorsal caudal fin	PDF-DCF
6	Pectoral fin length	P1F
7	Pelvic fin length	P2F
8	Pectoral fin to pelvic fin	P1F-P2F
9	Anal fin	AF
10	Anal fin base length	AFB
11	Pelvic fin to anterior anal fin	P2F-AAF
12	Anterior dorsal fin to pectoral fin	ADF-P1F
13	Posterior dorsal fin to pectoral fin	PDF-P1F
14	Anterior dorsal fin to pelvic fin	ADF-P2F
15	Posterior dorsal fin to pelvic fin	PDF-P2F
16	Anterior dorsal fin to anterior anal fin	ADF-AAF
17	Posterior dorsal fin to anterior anal fin	PDF-AAF
18	Posterior dorsal fin to posterior anal fin	PDF-PAF
19	Pectoral fin to ventral caudal fin	P1F-VCB
20	Anterior anal fin to dorsal caudal fin	AAF-DCF
21	Caudal peduncle length	CPL
22	Posterior anal fin to dorsal caudal fin	PAF-DCF
23	Operculum to pre-operculum	O-PO
24	Operculum to eye	O-E
25	Head length	HL
26	Eye to snout	E-SN
27	Eye diameter	ED
28	Snout to pectoral fin	SN-P1F
29	Snout to pelvic fin	SN-P2F
30	Snout to posterior dorsal fin	SN-PDF
31	Snout to anterior dorsal fin	SN-ADF
32	Eye to posterior nares	E-N
33	Left to right nares	L-RN

Hybrid Identification from Morphological Characters

The position of individuals from the putative hybrid zone in PCA morphospace was compared with the 95% confidence ellipses of the putative pure groups to determine with which group hybrids were mostly associated. F_1 hybrids were predicted to be intermediate to the pure groups (due to inheritance of morphological features from both parental species) and advanced-generation hybrids (F_2 hybrids or backcrosses) were placed closest to the parental species that contributed most to the breeding history. Identifications based on genetic and morphological data were compared to assess the utility of the two data types for identification of individuals with hybrid origin.

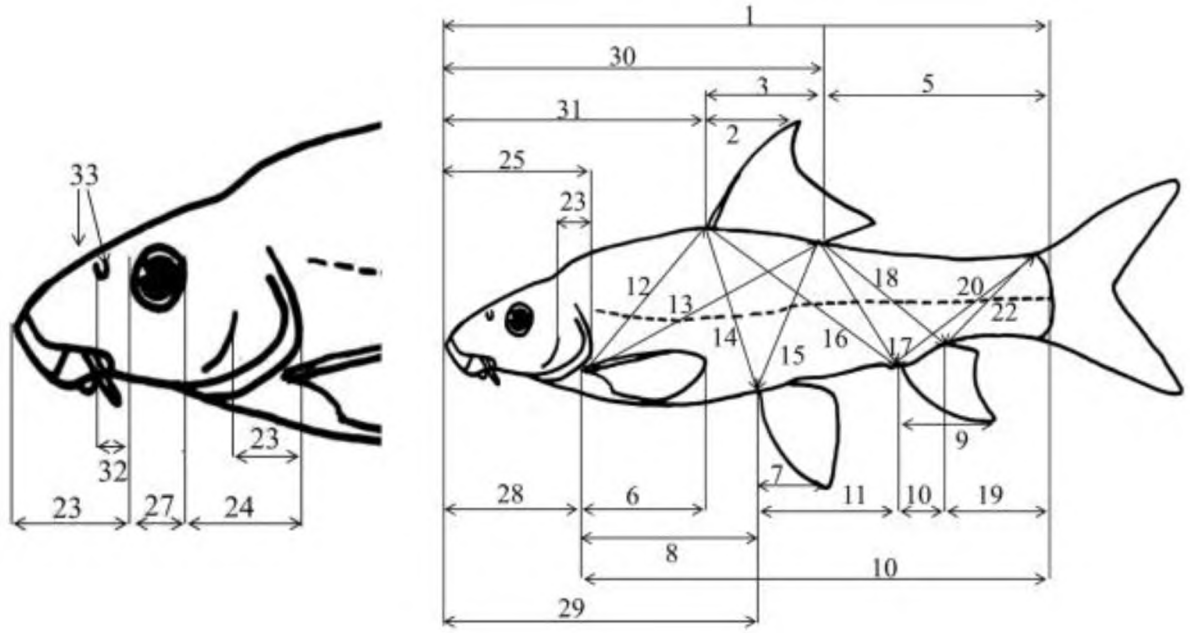


FIG. 4.4. The landmarks of the morphometric measurements of the *Labeo* fish. 1= standard length, 2= dorsal fin length, 3= dorsal fin base length, 4= caudal peduncle depth, 5= posterior dorsal fin to dorsal caudal fin, 6= pectoral fin length, 7= pelvic fin length, 8= pectoral fin to pelvic fin, 9= anal fin, 10= anal fin base length, 11= pelvic fin to anterior anal fin, 12= anterior dorsal fin to pectoral fin, 13= posterior dorsal fin to pectoral fin, 14= anterior dorsal fin to pelvic fin, 15= posterior dorsal fin to pelvic fin, 16= anterior dorsal fin to anterior anal fin, 17= posterior dorsal fin to anterior anal fin, 18= posterior dorsal fin to posterior anal fin, 19= pectoral fin to ventral caudal fin, 20= anterior anal fin to dorsal caudal fin, 21= caudal peduncle length, 22= posterior anal fin to dorsal caudal fin, 23= operculum to pre-operculum, 24= operculum to eye, 25= head length, 26= eye to snout, 27= eye diameter, 28= snout to pectoral fin, 29= snout to pelvic fin, 30= snout to posterior dorsal fin, 31= snout to anterior dorsal fin, 32= eye to nares, 33= left to right nares.

TABLE 4.3. Characters used in the meristic analysis and their abbreviations (codes).

Characters	Codes
Lateral line scale count	LL
Lateral line to the origin of the dorsal fin scale count	LL-DF
Lateral line to pelvic fin scale count	LL-P2F
Caudal peduncle scale count	CP

RESULTS

Phylogenetic Differentiation

An initial 160 S7 sequences were phased to produce 320 copies, which were represented by 18 unique nuclear alleles. The model that best fit the variation among these alleles was HKY+I (Hasegawa *et al.*, 1985). Of the 608 base pairs used, 553 sites were invariable and 16 were variable. Of the 16 variable sites, 11 were parsimony informative and five were autapomorphic. Two monophyletic groups were apparent in the phylogenetic analysis phylogram (Fig. 4.5). As alleles in the second group were present in homozygous genotypes and were from localities where mostly *L. umbratus* is found, the individuals were classified as *L. umbratus*. The first group was also homozygous and was associated with *L. capensis*. Alleles from heterozygous individuals were present in both species groups (represented by A5 and A6 occurring in combination with A20 and A23 in the Hardap, and A5 and A25 in the Darlington impoundments), which was consistent with hybridisation between the two species.

The 214 mtDNA sequences yielded 25 unique haplotypes. The model that best fits the variation among these allele fragments was TrN+I (Tamura & Nei, 1993). Of the 730 base pairs used, 704 sites were invariable and 26 were variable. Of the 26 variable sites, 14 were parsimony informative and 12 were autapomorphic. Two *L. umbratus* clades were resolved and the remainder of the alleles represented *L. capensis* (Fig. 4.6). All of the individuals identified as *L. umbratus* and hybrids by nDNA possessed mtDNA of *L. capensis*, and thus were identified as hybrids.

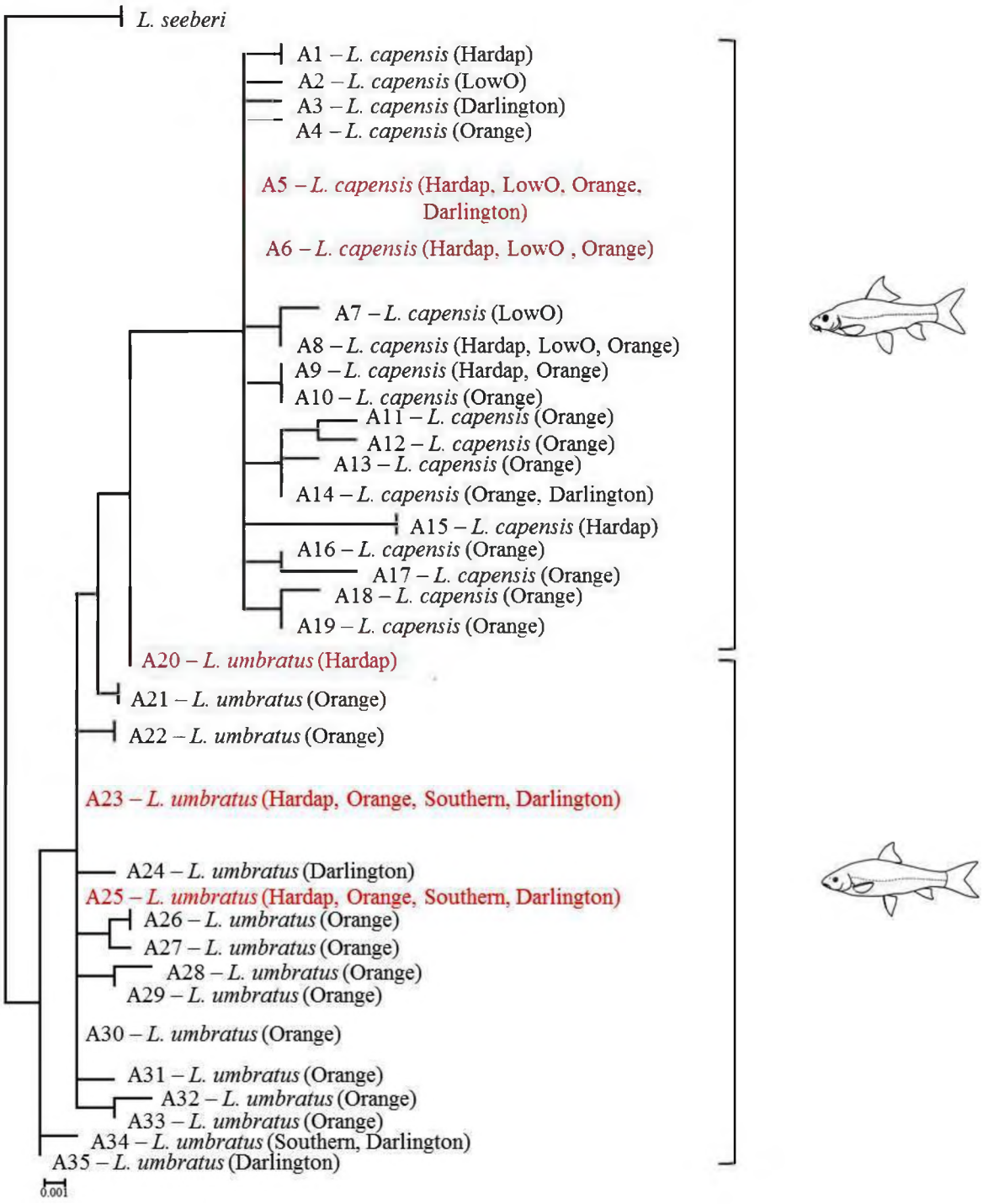


FIG. 4.5. Maximum likelihood phylogram of nuclear S7 gene constructed using the HKY+I model, showing the relationships between *L. capensis* (A1–19) and *L. umbratus* (A20–35). Alleles A5 and A6 occurred in combination with A20 and A23 in the Hardap and A5 and A25 in the Darlington impoundments (highlighted in red), indicating the individuals were hybrids because they carried one allele each from the two putative parental species. *Labeo seeberi* was used as an outgroup. LowO = Lower Orange.

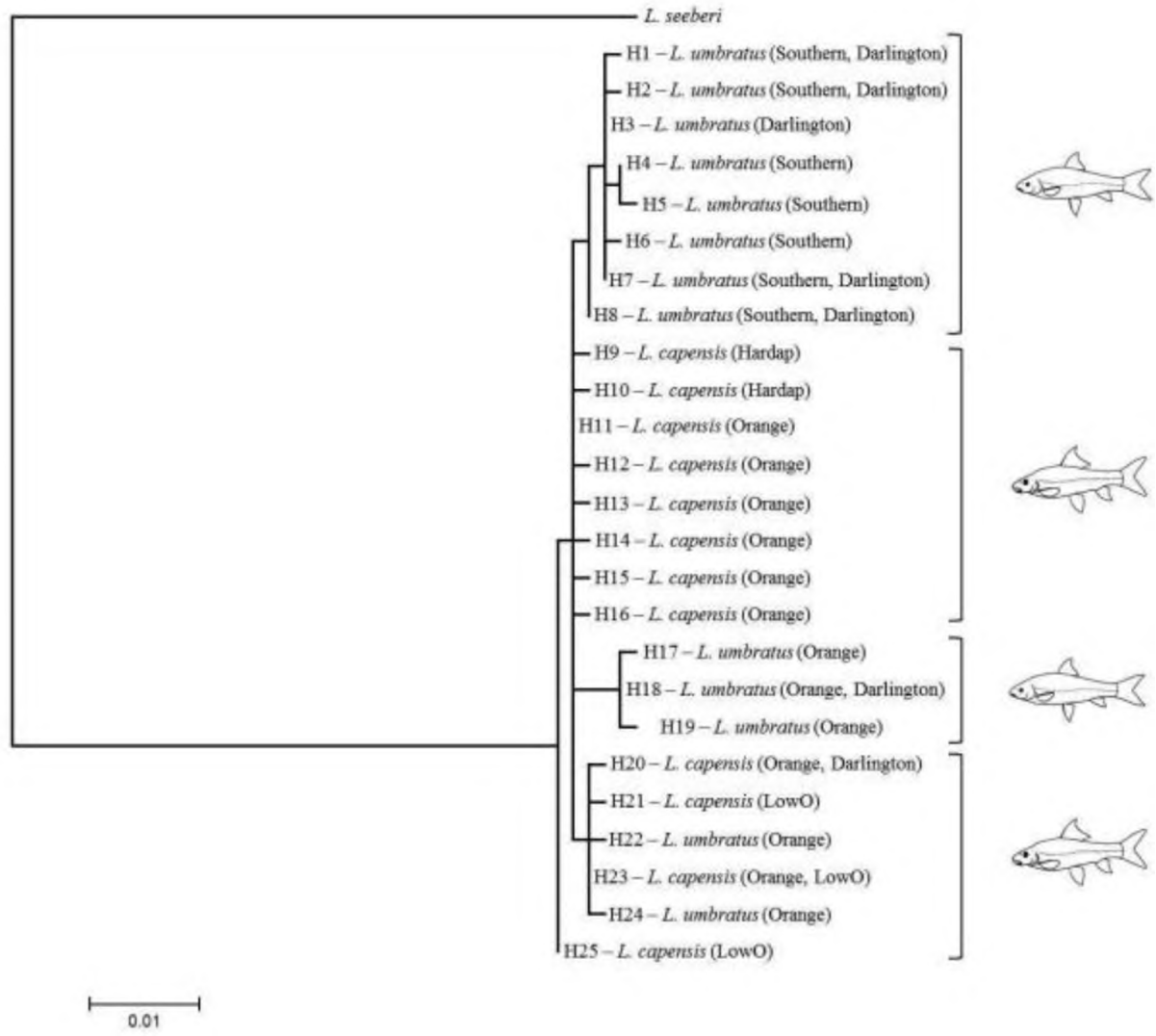


FIG. 4.6. Maximum likelihood phylogram of mitochondrial Cyt *b* gene constructed using the TrN+I model, showing the close relationship of *L. capensis* alleles (H9–16 and H20–25) to the two distinct *L. umbratus* clades (Southern H1–8 and Orange H17–19). *Labeo seeberi* was used as an outgroup. LowO = Lower Orange.

Detection and Geographic Distribution of Mitochondrial DNA Haplotypes

A total of 25 unique mtDNA sequences (haplotypes) was found. Three groups (lineages) of similar haplotypes were distinguished, namely Orange River *L. capensis* and *L. umbratus* representing lineages A (alleles 4–17) and B (alleles 1–3), respectively, and populations from southern-flowing river systems [lineage C (alleles

18–25)] (Fig. 4.7). Lineage (C) was associated with *L. umbratus* from southern-flowing rivers, which is the only indigenous species of *Labeo* in these river systems. Only Orange lineage (A), associated with *L. capensis*, was present in the middle and lower Orange River, and Orange lineage (B) was associated with *L. umbratus*. Surprisingly, *L. capensis* alleles were indicated to be more closely related to the two *L. umbratus* clades than the two *L. umbratus* clades were related to each other, but this could be the result of incomplete lineage sorting (Fig. 4.7).

Most haplotypes detected were confined to the Orange River basin, which reflected the natural occurrence of two species and probably also the greater number of sampling locations. Only eight haplotypes (H18–H25) were detected in southern-flowing river systems (including hybrid zones), of which most were associated with *L. umbratus* that occurred naturally in these river systems. Some of the haplotypes from the Orange River lineage (A) (H4 and H12) and lineage (B) (H2) were also detected in the Darlington impoundment, which is part of the southern-flowing river system (Fig. 4.3). The abundant haplotypes detected in the Darlington impoundment belonged to lineage (C). In the Hardap impoundment, only Orange River lineage (A) haplotypes, associated with *L. capensis*, were detected (H4, H16 and H17). Lineage A showed the highest haplotype diversity, whereas the Orange River lineage (B) showed the lowest. The Orange River lineage (B), associated with *L. capensis* (Fig. 4.7; highlighted in red), differed from the Orange lineage (A) associated with *L. umbratus* (dark green) by five mutations (0.6% divergence), and differed from lineage C, associated with *L. umbratus* (light green) from southern-flowing river systems, by only one mutation (Fig. 4.7).

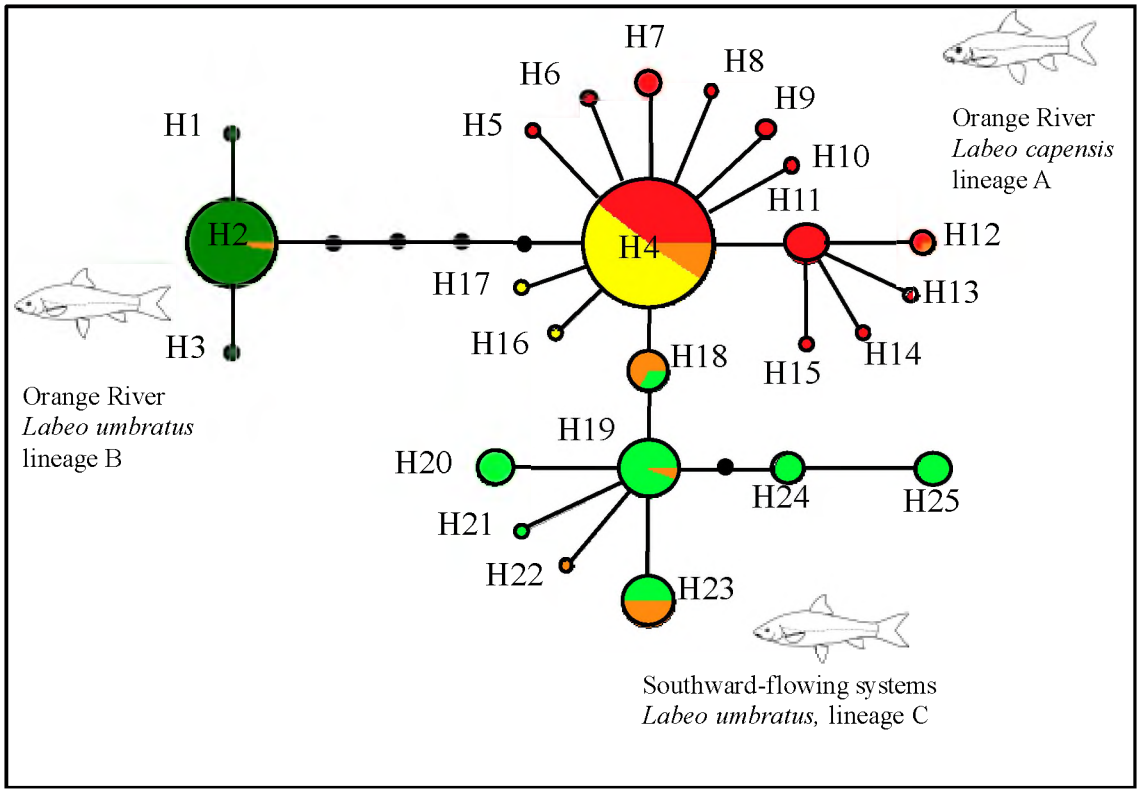


FIG. 4.7. Mitochondrial cytochrome *b* TCS parsimony network showing the geographic distribution of *Labeo umbratus* and *Labeo capensis* haplotypes. Putative pure populations colour codes: dark green = Orange River system *L. umbratus*, red = Orange River system *L. capensis*, light green = southward-flowing systems *L. umbratus*. Potential hybrid populations colour codes: yellow = Hardap, orange = Darlington. Black dots represent missing alleles.

Hybrid Identification from Genetic Data

Both Cyt *b* mtDNA and S7 nDNA data distinguished the pure species strains. No hybrids were identified in the lower Orange River (alleles were in the homozygous state for *L. capensis*) and Lake Gariep (alleles were homozygous for either pure *L. capensis* or *L. umbratus*), where hybridisation has been reported previously. Hybrids were only identified in the Hardap and Darlington impoundments. Certain individuals from the Hardap (MR08F029, 044, 060, 061 and 074) and Darlington (DIFS 07-133) impoundments were identified as F_1 hybrids on the basis of heterozygosity of S7 at each locus where the parental species exhibit fixed differences. Six individuals from Hardap that were homozygous for *L. umbratus* alleles (MR08F071 represented by A20; MR08F030, 050, 059 and 062 represented by A23; and MR08F056 represented by A25) in the nDNA phylogeny (Fig. 4.5) were classified as *L. capensis* on the basis of mtDNA data; these results were indicative of introgression, therefore the individuals were classified as possible backcross hybrids.

TABLE 4.4. Principal component analysis loadings for 33 morphometric characters on four principal components (PC1 to PC4), eigenvalues and percentage variance explained for three putative pure populations of Orange River *L. capensis*, *L. umbratus* and southern *L. umbratus*. Numbers highlighted in bold (values above 0.25) indicate loadings that are above the cut-off rule (V vectors) (Chatfield & Collins 1980).

Character	PC 1	PC 2	PC 3	PC 4
Eigenvalue	0.032	0.014	0.006	0.003
Variance explained (%)	50	21.914	9.484	4.186
SL	−0.025	−0.007	0.095	0.091
DF	**0.257**	0.179	−0.082	−0.007
DFB	**0.414**	0.130	−0.034	0.022
CPD	0.058	**−0.297**	0.177	−0.078
PDF-DCF	−0.218	0.065	0.208	−0.093
P1F	0.129	0.107	−0.082	0.077
P2F	0.221	0.210	−0.104	0.041
P1F-P2F	0.101	−0.115	0.027	−0.051
AF	**0.382**	0.176	−0.135	−0.079
AFB	**0.291**	0.141	−0.043	−0.142
P2F-AAF	−0.157	−0.060	0.215	0.035
ADF-P1F	0.036	−0.117	0.071	−0.054
PDF-P1F	0.136	−0.070	0.031	−0.013
ADF-P2F	0.092	−0.161	0.010	−0.105
PDF-P2F	0.056	−0.192	0.036	−0.082
ADF-AAF	0.023	−0.045	0.133	−0.018
PDF-AAF	−0.152	−0.150	0.219	−0.064
PDF-PAF	−0.155	−0.063	0.220	−0.018
P1F-VCF	−0.001	−0.026	0.098	−0.073
AAF-DCF	0.046	0.099	0.110	−0.189
CPL	−0.113	0.176	0.138	**−0.315**
PAF-DCF	−0.052	0.083	0.155	−0.168
O-PO	**−0.351**	**0.539**	−0.139	−0.021
O-E	**−0.303**	**0.252**	−0.204	0.089
HL	−0.094	0.062	−0.108	**0.280**
E-SN	0.008	**−0.255**	−0.176	**0.427**
ED	0.050	0.016	0.152	**0.522**
SN-P1F	−0.085	0.047	−0.004	**0.329**
SN-P2F	0.023	−0.039	0.021	0.108
SN-PDF	0.085	−0.027	0.030	0.084
SN-ADF	−0.007	−0.072	0.062	0.089
E-N	−0.192	**−0.379**	**−0.677**	**−0.259**
L-RN	−0.068	0.015	−0.218	0.024

SL= Standard length, DF= Dorsal fin length, DFB= Dorsal fin base length, CPD= Caudal peduncle depth, PDF-DCF= Posterior dorsal fin to dorsal caudal fin, P1F= Pectoral fin length, P2F= Pelvic fin length, P1F-P2F= Pectoral fin to pelvic fin, AF= Anal fin, AFB= Anal fin base length, P2F-AAF= Pelvic fin to anterior anal fin, ADF-P1F= Anterior dorsal fin to pectoral fin, PDF-P1F= Posterior dorsal fin to pectoral fin, ADF-P2F= Anterior dorsal fin to pelvic fin, PDF-P2F= Posterior dorsal fin to pelvic fin, ADF-AAF= Anterior dorsal fin to anterior anal fin, PDF-AAF= Posterior dorsal fin to anterior anal fin, PDF-PAF= Posterior dorsal fin to posterior anal fin, P1F-VCF= Pectoral fin to ventral caudal fin, AAF-DCF= Anterior anal fin to dorsal caudal fin, CPL= Caudal peduncle length, PAF-DCF= Posterior anal fin to dorsal caudal fin, O-PO= Operculum to pre-operculum, O-E= Operculum to eye, HL= Head length, E-SN= Eye to snout, ED= Eye diameter, SN-P1F= Snout to pectoral fin, SN-P2F= Snout to pelvic fin, SN-PDF= Snout to posterior dorsal fin, SN-ADF= Snout to anterior dorsal fin, E-N= Eye to nares, L-RN= Left to right nares.

Morphological Differentiation of Pure and Hybrid Individuals

Four principal components (PC) described 85.6% of the morphometric variation among individuals in the size-standardised pure populations data set (Table 4.4). The first component explained 50% of the variation remaining after allometric correction, and was positively loaded by high values for Dorsal-fin length (DF) (0.257), Dorsal-fin base width (DFB) (0.414), Anal-in length (AF) (0.382) and Anal-fin base width (AFB) (0.291), and negatively for distance from operculum to pre-operculum (O-OP) (−0.351) and operculum to the eye (O-E) (−0.303). The first component fully separated Orange River *L. capensis* and the two *L. umbratus* populations (Fig. 4.8a), and indicated that Orange River *L. capensis* differed from the two *L. umbratus* populations by possessing longer dorsal and anal fins and shorter opercles.

The second principal component explained 22% of the total variance and was positively loaded by distances between the operculum and the pre-operculum (O-PO) (0.539) and the operculum and eye (O-E) (0.252), negatively by caudal peduncle depth (CPD) (-0.297), distance from the eye to the snout (E-SN) (-0.255) and distance from the eye to the nostril (E-N) (-0.379). The second component largely separated the two *L. umbratus* populations, but with slight overlap (Fig. 4.8a). This result indicated that the Orange River *L. umbratus* have, on average, larger heads, shorter snouts and slimmer caudal peduncles compared with *L. umbratus* populations in southern-flowing river systems.

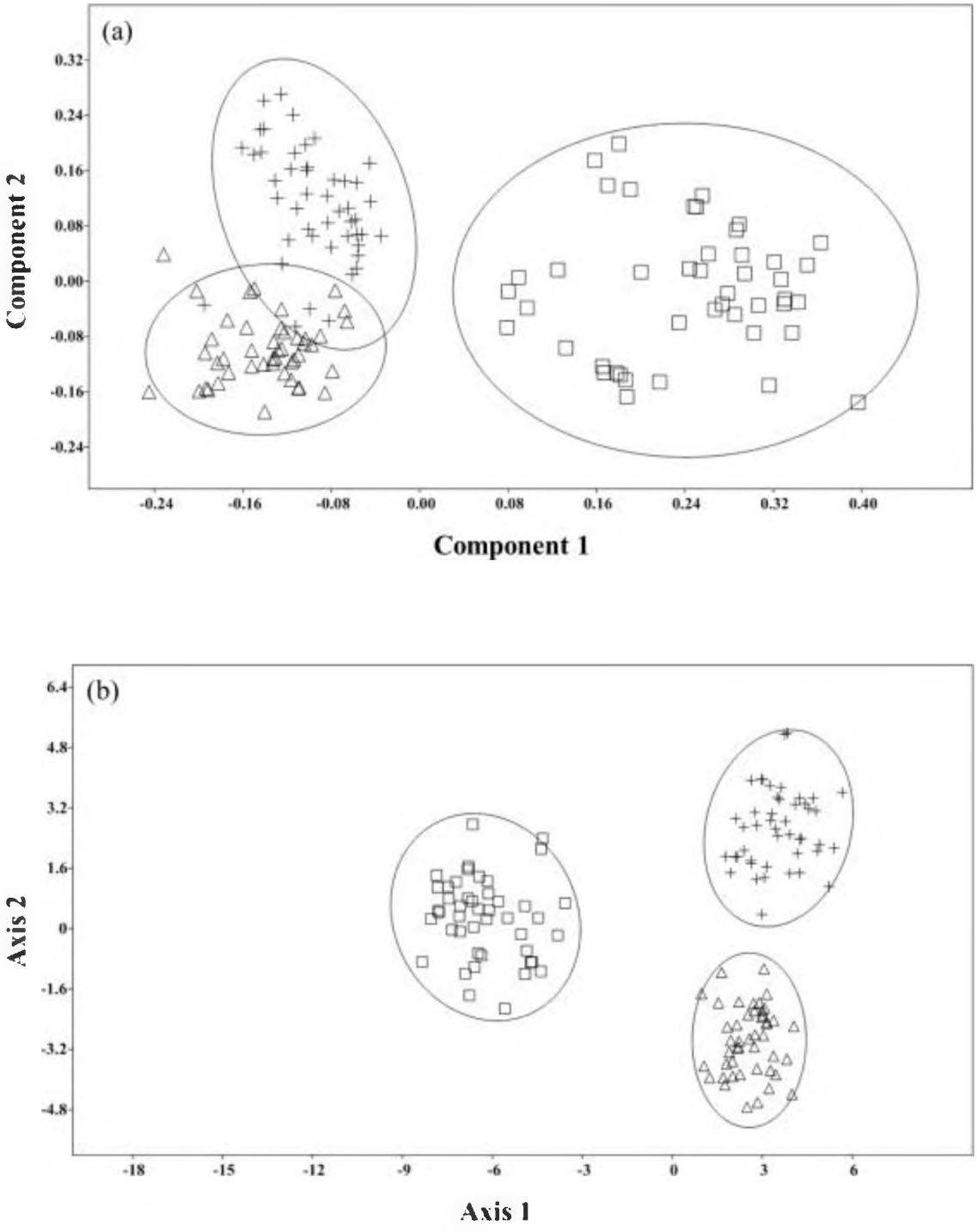


FIG. 4.8. (a) Principal component analysis scatterplot of the first and second principal components for specimens of Orange River *L. capensis*, and Orange River and southern *L. umbratus* derived from 33 morphometric characters. Orange River *L. capensis* (□) is shown to be distinct from the other two populations, while Orange River *L. umbratus* (+) and southern river systems *L. umbratus* (Δ) are shown to overlap slightly. The ellipses represent 95% confidence intervals. (b) Canonical variate analysis scatterplot of the specimens on the first and second discriminant functions.

TABLE 4.5. MANOVA/CVA loadings for 33 morphometric characters on the first two canonical axes (Axis 1 and 2), eigenvalues and percentage variance explained for three putative pure populations (Orange River *L. capensis*, Orange River *L. umbratus* and southern *L. umbratus*.

Character	**Axis 1**	**Axis 2**
Eigenvalue	19.68	5.487
Variance explained (%)	78.2	21.8
SL	11.408	27.386
DF	−4.706	7.053
DFB	−3.357	−1.18
CPD	4.415	5.59
PDF-DCF	15.497	16.466
P1F	−6.149	−19.186
P2F	17.121	24.941
P1F-P2F	−1.85	−16.466
AF	−14.438	6.370
AFB	0.901	13.003
P2F-AAF	5.451	−24.656
ADF-P1F	13.752	19.668
PDF-P1F	−3.818	−13.652
ADF-P2F	−28.754	−11.093
PDF-P2F	25.562	14.463
ADF-AAF	−4.270	−6.220
PDF-AAF	−15.658	−23.689
PDF-PAF	23.956	−7.889
P1F-VCF	15.464	79.149
AAF-DCF	−8.677	−28.202
CPL	−0.789	−7.406
PAF-DCF	−3.714	−2.465
O-PO	−2.136	−11.32
O-E	6.324	10.265
HL	−0.066	0.523
E-SN	−12.421	−19.506
ED	−7.838	−1.807
SN-P1F	55.503	65.397
SN-P2F	−33.275	−25.716
SN-PDF	−52.488	−34.083
SN-ADF	−0.763	−42.685
E-N	−0.458	−1.723
L-RN	1.898	14.555

SL= Standard length, DF= Dorsal fin length, DFB= Dorsal fin base length, CPD= Caudal peduncle depth, PDF-DCF= Posterior dorsal fin to dorsal caudal fin, P1F= Pectoral fin length, P2F= Pelvic fin length, P1F-P2F= Pectoral fin to pelvic fin, AF= Anal fin, AFB= Anal fin base length, P2F-AAF= Pelvic fin to anterior anal fin, ADF-P1F= Anterior dorsal fin to pectoral fin, PDF-P1F= Posterior dorsal fin to pectoral fin, ADF-P2F= Anterior dorsal fin to pelvic fin, PDF-P2F= Posterior dorsal fin to pelvic fin, ADF-AAF= Anterior dorsal fin to anterior anal fin, PDF-AAF= Posterior dorsal fin to anterior anal fin, PDF-PAF= Posterior dorsal fin to posterior anal fin, P1F-VCF= Pectoral fin to ventral caudal fin, AAF-DCF= Anterior anal fin to dorsal caudal fin, CPL= Caudal peduncle length, PAF-DCF= Posterior anal fin to dorsal caudal fin, O-PO= Operculum to pre-operculum, O-E= Operculum to eye, HL= Head length, E-SN= Eye to snout, ED= Eye diameter, SN-P1F= Snout to pectoral fin, SN-P2F= Snout to pelvic fin, SN-PDF= Snout to posterior dorsal fin, SN-ADF= Snout to anterior dorsal fin, E-N= Eye to nares, L-RN= Left to right nares.

The MANOVA/CVA showed a significant overall difference between groups (Wilk's $\lambda = 0.007452$; $F = 32.71$, $p < 0.05$), supported by *post-hoc* analysis of pairwise Hostelling's T squared (Bonferroni corrected) (Orange *L. capensis* vs *L. umbratus*, $p <$ 0,05; Orange *L. capensis* vs southern *L. umbratus*, $p < 0,05$; Orange *L. umbratus* vs southern *L. umbratus*, $p < 0,05$). Thus, the morphological equivalence of the three pure groups was rejected. The CVA scatterplot showed 100% discrimination of the three *a priori* groups (Fig. 4.8b). Axis 1 explained 78.2% of the total variation and Axis 2 accounted for 21.8% (Table 4.5). All variables showed a significant correlation with the two discriminant functions.

In the PCA of meristic data, PC1 described the vast majority of variation (91.6%) among individuals in the pure populations data set (Table 4.6). The axes were positively loaded by high values for all four variables. This result confirmed that *L. umbratus* has more and smaller scales than *L. capensis*. The PCA scatterplot derived from the meristic data clearly differentiated *L. capensis* and *L. umbratus* individuals, but the two *L. umbratus* populations were not distinct (Fig. 4.9a). Analysis of data for the two *L. umbratus* populations alone revealed similar results (data not presented). MANOVA showed a significant overall difference between groups (Wilk's $\lambda = 0.003766$; $F =$ 146.4, $p < 0.05$). This finding was supported by *post-hoc* analysis of pairwise Hotelling's T squared (Bonferroni corrected) using all components (Orange *L. capensis* vs *L. umbratus*, $p < 0.05$; Orange *L. capensis* vs southern *L. umbratus*, $p < 0.05$; Orange *L. umbratus* vs southern *L. umbratus*, $p < 0.05$). CVA revealed 100% classification among *L. capensis* and the two *L. umbratus* groups, and 86.86% classification of individuals between the two *L. umbratus* groups (Fig. 4.9b). The hypothesis that the

three populations are not distinct was rejected only when comparing *L. capensis* and the two *L. umbratus* groups. Axis 1 explained 98.2% of the total variation and Axis 2 accounted for 1.439% (Table 4.6). All variables showed a significant correlation with the two discriminant functions.

TABLE 4.6. Principal components analysis loadings for the first two principal components (PC1 and 2), and MANOVA/CVA loadings for the first two canonical axes (Axis 1 and 2), for four meristic variables for three putative pure populations (Orange River *L. capensis*, Orange River *L. umbratus* and southern *L. umbratus*. Numbers highlighted in bold (values above 0.25) indicate PCA loadings that are above the cut-off rule (V vectors) (Chatfield & Collins 1980). Eigenvalues and percentage variance explained for PCA and MANOVA/CVA are also listed.

	PC 1	**PC 2**	**Axis 1**	**Axis 1**
Eigenvalues	0.022	0.001	15.16	0.329
Variance explained (%)	91.605	3.934	97.88	2.125
LL	**0.469**	−0.018	34.574	38.62
LL-DF	**0.541**	0.083	19.641	−31.569
LL-P2V	**0.452**	**0.792**	−2.443	−18.268
CP	**0.532**	**−0.605**	8.827	12.597

LL= Lateral line scale count, LL-DF= Lateral line to the origin of the dorsal fin scale count, LL-P2F= Lateral line to pelvic fin scale count, CP= Caudal peduncle scale count

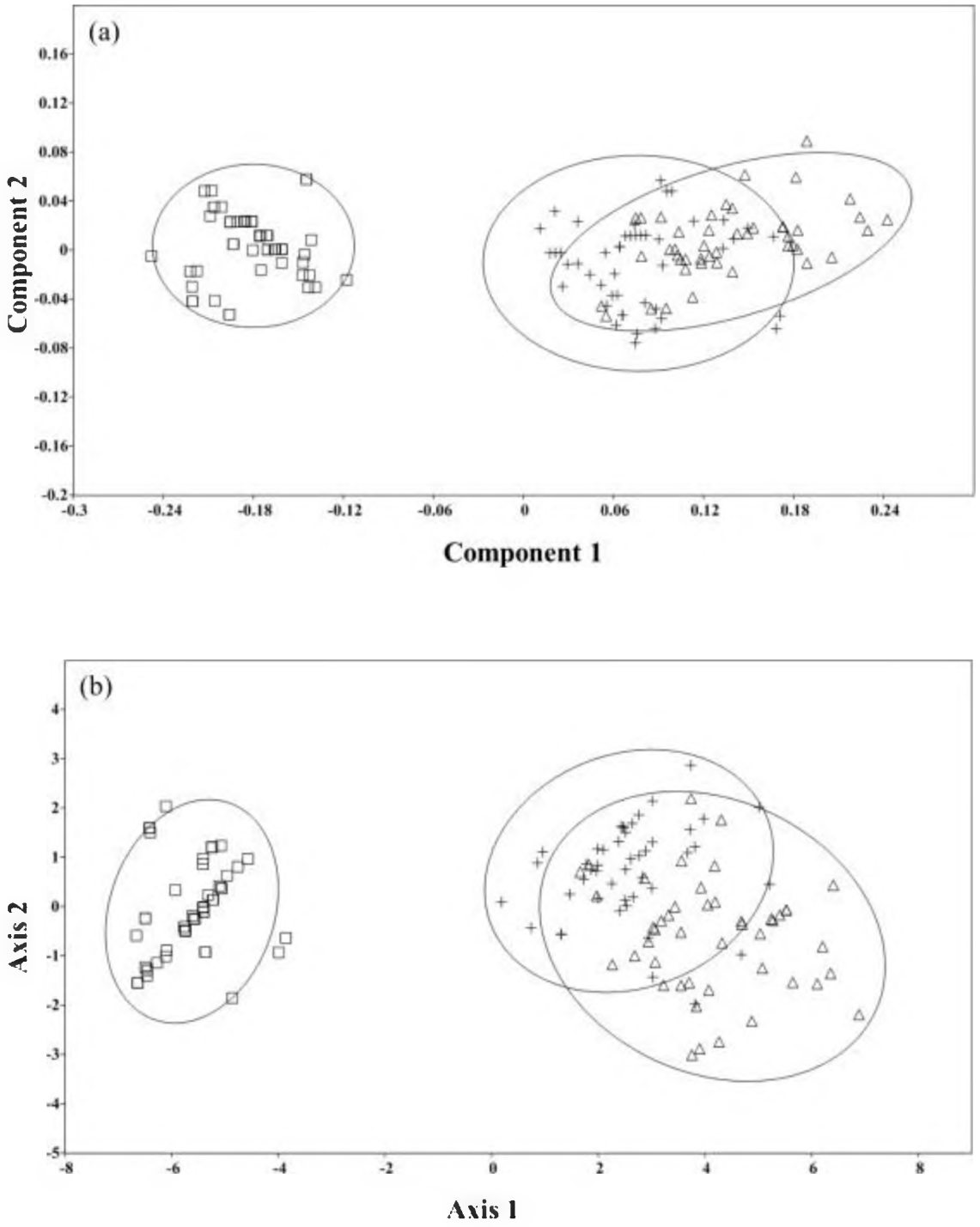


FIG. 4.9. Principal components analysis scatterplot of the first and second principal components for specimens of Orange River *L. capensis*, and Orange River and southern *L. umbratus* derived from four meristic variables. Orange *L. capensis* (□) is shown to be distinct from the other two populations, while Orange River *L. umbratus* (+) and southern systems *L. umbratus* (Δ) show considerable overlap. The ellipses represent 95% confidence intervals. (b) Canonical variate analysis scatterplot showing discrimination of the specimens on the first and second discriminant functions.

Hybrid Identification from Morphological Data

Individuals from the putative introgressed Hardap population were included in the morphometric and meristic data sets together with members of the three pure populations, and the data were subjected to PCA. Some Hardap individuals were grouped with *L. capensis* population, a few with the *L. umbratus* group, and the remainder were placed intermediate between the two groups. Two individuals classified as F_1 hybrids from genetic data were placed between the two species groups outside the respective 95% confidence ellipses in the PCA scatterplot derived from morphometric data (Fig. 4.10a). The other three individuals genetically classified as F_1 hybrids were grouped with the parental species, two with *L. umbratus* and one with *L. capensis*. Eight individuals genetically classified as F_2 hybrids were grouped with the two parental species, three with *L. umbratus* and five with *L. capensis*. In the PCA of the meristic data, only one individual genetically classified as a F_1 hybrid was grouped with *L. umbratus*; the other four individuals were placed between the two parental species (Fig. 4.10b). Of the nine individuals genetically classified as F_2 hybrids, three were grouped with *L. umbratus*, two with *L. capensis*, three were intermediate between the two groups but two were placed close to *L. capensis*, and one was placed close to *L. capensis* but outside the 95% confidence ellipses. Most of the putative hybrids had scale counts intermediate between those of the two species as indicated by the meristic data (Fig. 4.10b, Table 4.7).

TABLE 4.7. Scale count ranges for pure Orange River *L. capensis*, Orange River *L. umbratus* and southern systems *L. umbratus*, and putative hybrids from the Hardap and Darlington impoundments.

Character	Orange River *L. capensis*	Orange River *L. umbratus*	Southern systems *L. umbratus*	Hardap hybrids		Darlington hybrids F_1
				F_1	F_2/backcrosses	
LL	40–45	52–64	54–67	43–51	42–53	45
LL-DF	8–10	11–14	12–15	10–11	8–12	10
LL-P2V	6–8	8–10	8–12	7–8	6–9	9
CP	20–26	28–41	29–39	28–34	27–35	33

LL= Lateral line scale count, LL-DF= Lateral line to the origin of the dorsal fin scale count, LL-P2F= Lateral line to pelvic fin scale count, CP= Caudal peduncle scale count

When data for individuals of the Darlington population were analysed with the same parental populations, most Darlington specimens were grouped with the *L. umbratus* group, a few with the *L. capensis* group, and only one individual, genetically classified as a F_1 hybrid, was morphologically intermediate but closer to the *L. umbratus* group (Fig. 4.10c and d). These results indicated that pure *L. capensis* and *L. umbratus* were represented in the Darlington impoundment.

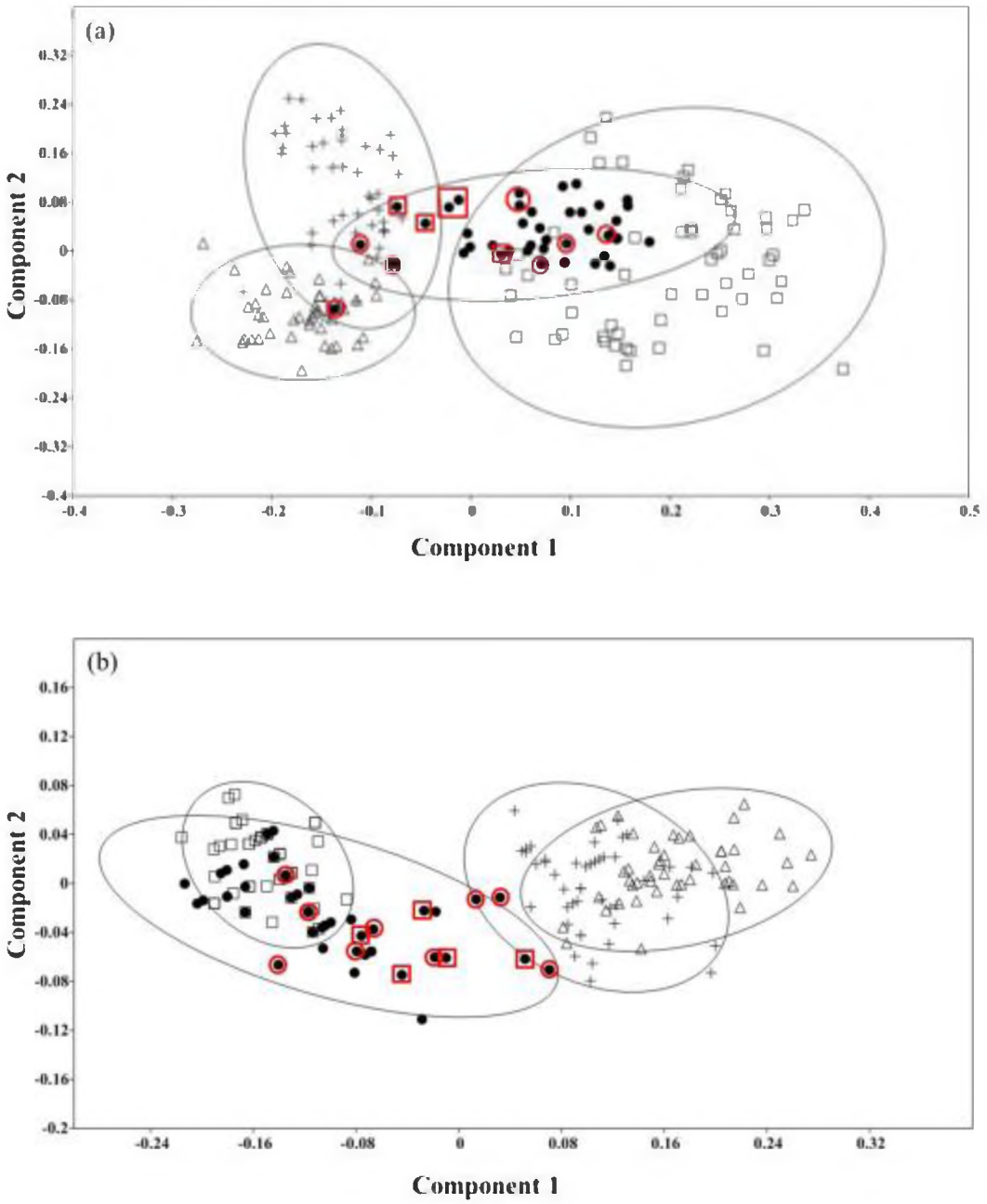
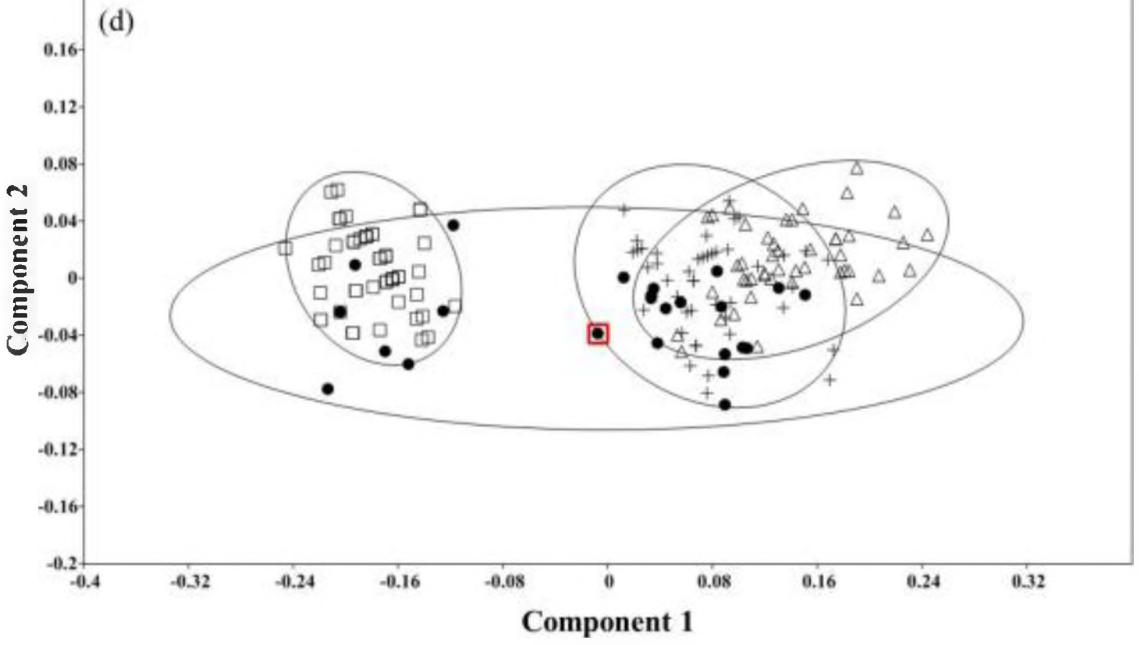


FIG. 4.10. Principal component analysis scatterplots derived from morphometric and meristic data for the three reference pure populations (Orange River *L. capensis* (□), Orange River *L. umbratus* (+) and southern systems *L. umbratus* (Δ)) and individuals from putative hybrid zones (•). (a) Hardap morphometrics, (b) Hardap meristics, (c) Darlington morphometrics and (d) Darlington meristics. Putative first-generation (F_1) hybrids are indicated by red squares and second-generation hybrids (F_2/backcrosses) are indicated by red circles, as identified from genetic data.


(d)
0.16
0.12
0.08
0.04
0
-0.04
-0.08
-0.12
-0.16
-0.2
Component 2
-0.4
-0.32
-0.24
-0.16
-0.08
0
0.08
0.16
0.24
0.32
Component 1

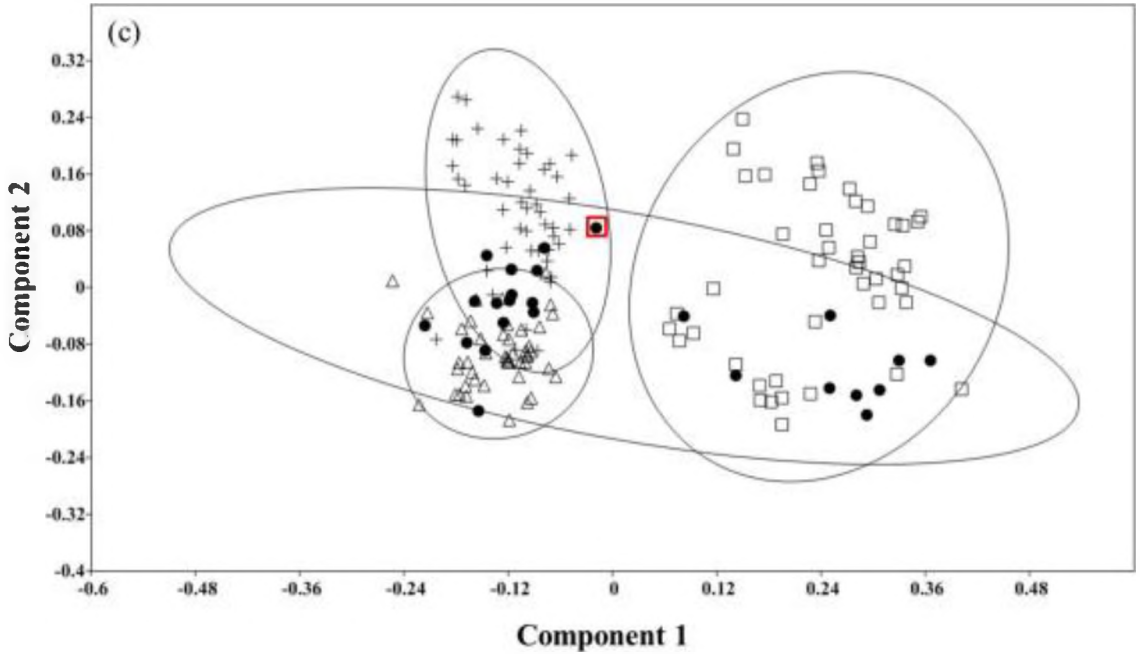


FIG. 4.10. (continued.)

DISCUSSION

This study is the first comprehensive assessment of the genetic and morphological characteristics of pure and putative hybrid *L. capensis* and *L. umbratus* using relatively large sample sizes (218 individuals) across their distributional range. Nine morphometric characters (dorsal-fin length, dorsal-fin base length, caudal peduncle depth, anal-fin length, anal-fin base length, operculum to pre-operculum, operculum to the eye distance, eye to snout, and eye to nostril), four meristic characters (number of scales along the lateral line, between the lateral line and the origin of the dorsal fin, between the lateral line and origin of the pelvic fin, and around the caudal peduncle), four fixed nuclear DNA mutations (first intron of the S7 ribosomal protein coding gene) and five mtDNA mutations (cytochrome *b*) distinguished the three *Labeo* populations (*L. capensis*, Orange River *L. umbratus* and southern *L. umbratus*). According to Van Vuuren *et al.* (1989, 1990), who studied populations of *L. umbratus* and *L. capensis* from the Hardap impoundment, the two species can be distinguished by the number of dorsal-fin spines. However, the present results indicate that this is not the case because *L. capensis* has three dorsal-fin spines and *L. umbratus* has either three or four dorsal-fin spines.

Groups that were hypothesised *a priori* to represent distinct and non-introgressed populations (middle Orange, Brak, Vaal, Bushmans, Keiskamma and Gariep, and Kat rivers and the Slagboom impoundment) were discriminated by genetic data (Figs. 4.5 and 4.6) and morphology (Figs. 4.8 and 4.9). Individuals of *L. capensis* from the lower Orange River, in which hybridisation was suspected to have occurred, were classified on the basis of genetic and morphological data as pure *L. capensis*. The combination of genetic and morphological methods enabled identification of possible F_1 hybrids and

backcrosses in the Hardap (Fig. 4.10a and b) and Darlington (Fig. 4.10c and d) impoundments. Unlike the study by Van Vuuren *et al.* (1989), it was also possible to identify putative F_1 hybrids from morphological characters (Fig. 4.10, Table 4.7). Putative F_1 hybrids could be identified using nDNA sequence data as they carry an equal mixture of genes from both parental genomes, but morphological characters were less informative as most of the putative F_1 hybrids were grouped with or closest to one of the putative parental species, probably because of dominance or epigenetic effects (genetic effects not encoded in the DNA sequence of an organism) (Chen, 2007). Backcrosses and F_2 hybrids are more difficult to identify using only the two data types as such hybrids show phenotypes of one of the parental species with greater frequency (Campton, 1987; Pacheco *et al.*, 2002). Hybridisation in the Hardap and Darlington impoundments was indicated to be at different stages, which may reflect different processes. This is because the two impoundments were indicated to show different patterns of hybrid types (individual identification with meristics, morphometrics, nDNA and mtDNA) between the two impoundments (Table 4.8). Introgression in the Hardap impoundment seems to be distinctly older and more extensive than in the Darlington impoundment, because in the former impoundment putative hybrids were more frequently encountered during field surveys and most of the putative hybrids sampled were indicated to be F_2 hybrids or backcrosses.

Hybridisation in the Hardap Impoundment

The persistence of only *L. capensis* mtDNA alleles in the Hardap impoundment (Fig. 4.7) indicated that all of the specimens identified as *L. umbratus* (from nDNA sequences and morphological characters) have a history of recent or more ancient hybridisation (see also Bernatchez *et al.*, 1995; Wilson & Bernatchez, 1998; Freyhof *et*

al., 2005) and that *L. capensis* females were mostly involved in hybridisation (Wirtz, 1999). Hybridisation in the Hardap impoundment was indicated to have led to the complete fixation of *L. capensis* mtDNA in hybrid individuals and that either pure *L. umbratus* no longer persists or few pure individuals remain and were not sampled. Several studies have reported similar mtDNA replacement (Aubert & Solignac, 1990; Wilson & Bernatchez, 1998; Freyhof *et al.*, 2005).

For example, Freyhof *et al.* (2005), who examined evidence for introgression of mtDNA in Dalmatian cyprinids found in Lake Busko (Rièina River), reported that one population of *Scardinius dergle* Heckel & Kner 1857 grouped with the *Scardinius* genus on the basis of morphological and genetic (nDNA) data but had *Squalius tenellus* Heckel 1843 mtDNA. These authors speculated that introgression must have occurred following construction of a dam in 1962, probably trapping the two parental species in the pits below the dam. This is similar to the findings of the current study, except that in the case of *L. capensis* × *L. umbratus* hybrids, hybridisation is indicated to have led to mtDNA replacement between congeneric species.

The presence of a combination of possible hybrid 'types' in the Hardap impoundment (Table 4.8) is suggestive of ongoing introgressive hybridisation (Miller, 1963). This finding suggests that hybrids of *L. capensis* and *L. umbratus* are fertile, which implies there is a risk of widespread introgression and potential for complete admixture (Hitt *et al.*, 2003; Boyer *et al.*, 2008). According to the criteria of Allendorf *et al.* (2001), hybridisation in the Hardap impoundment can be categorised as Type 5 hybridisation (i.e., widespread introgression).

TABLE 4.8. Summary of the categories of pure (*L. capensis* = CAP, *L. umbratus* = UMB) and hybrid specimens that were identified using a combination of data types (morphology and genetic) in the putative hybrids areas (Hardap and Darlington impoundments).

	Morphology		**Genetics**			
Types	Morphometrics	Meristics	Nuclear DNA	Mitochondrial DNA	Locality	*N*
Pure categories						
CAP	CAP	CAP	CAP	CAP	Both dams	12
UMB	UMB	UMB	UMB	UMB	Darlington	14
Possible F_1 hybrid categories						
1	Hybrid	Hybrid	Hybrid	CAP	Hardap	5
2	Hybrid	Hybrid	Hybrid	UMB	Darlington	1
Possible backcross hybrid categories						
3	UMB	Hybrid	CAP	CAP	Hardap	1
4	UMB	Hybrid	Hybrid	CAP	Hardap	2
5	UMB	Hybrid	UMB	CAP	Hardap	2
6	Hybrid	CAP	CAP	CAP	Hardap	7
7	Hybrid	CAP	Hybrid	CAP	Hardap	1
8	Hybrid	CAP	UMB	CAP	Hardap	1
9	Hybrid	Hybrid	UMB	CAP	Hardap	2

Labeo capensis phenotypes were more abundant than *L. umbratus* in the Hardap impoundment (Table 4.9). At this locality only 14 of the 81 specimens collected in a two-day gill net survey showed *L. umbratus* phenotypes. Gaigher & Bloemhof (1975) suggested that *L. capensis* dominated because the impoundment is mostly rocky bottomed, which is habitat favoured by *L. capensis*. Hamman (1980) noted that the two species shared the same breeding sites in Lake Gariep, and Tweddle & Davies (1997) noted that, like *L. umbratus* (Tomasson *et al.*, 1984), *L. capensis* undertook lateral migration into newly inundated littoral habitats to spawn in the Katse impoundment.

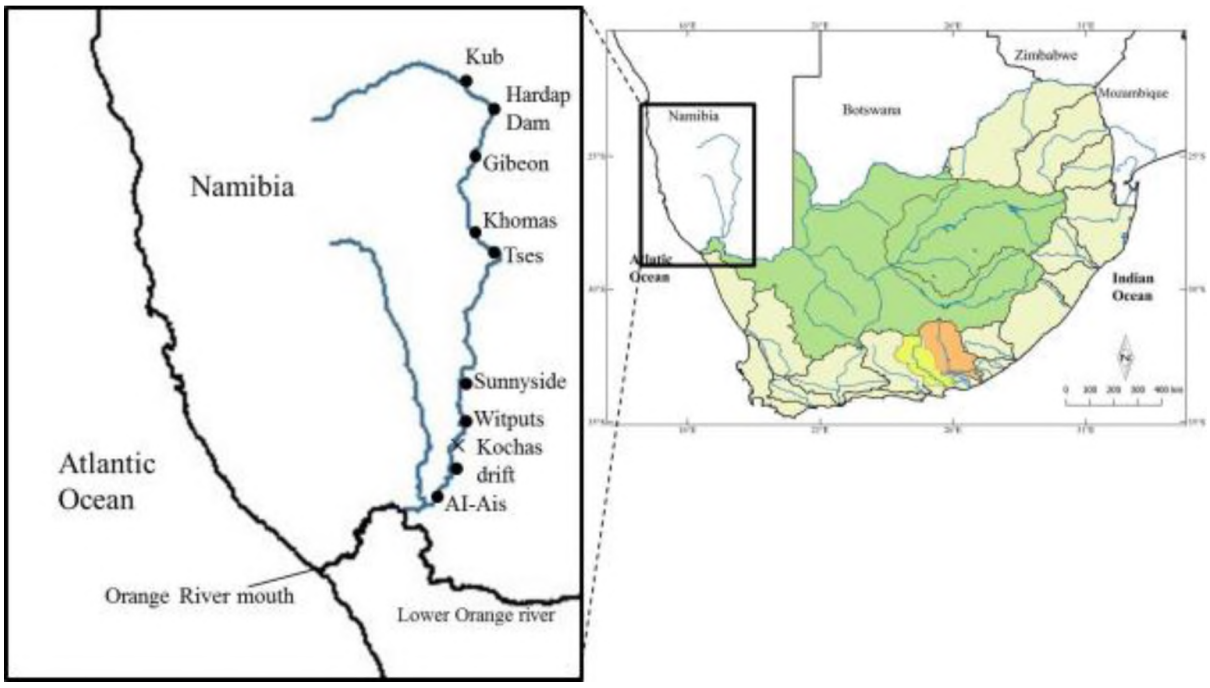


FIG. 4.11. Expanded view of the sampling localities in the Fish River, Namibia. The × symbol between Witputs and Kochas drift represents a waterfall.

It is therefore possible that *L. capensis* out competes *L. umbratus* in dams (Winker *et al.*, 2012). When *L. capensis* is the dominant species, *L. umbratus* might struggle to find conspecific mates leading to extirpation of genetically pure *L. umbratus* from the Hardap impoundment with only hybrids expressing some *L. umbratus* phenotypes now remaining in the population.

According to Hay (1991), who studied the distribution of fish in the Fish River tributary of the Orange River, putative hybrids and *L. umbratus* were only sampled from the upper reaches of the river at Kub above the Hardap Dam to below the dam at Sunnyside (Fig. 4.11). No putative hybrids and *L. umbratus* were found below the waterfall to the lower Orange River. The present study also identified only pure *L. capensis* in the lower Orange River. This finding may be a result of inadequate sampling or indicates that *L.*

umbratus and hybrids with *L. capensis* have not spread to this part of the river system. The results, however, do not imply that the genetic integrity of *L. capensis* downstream of the Hardap impoundment and in the lower Orange is not at risk.

Populations of *L. capensis* and *L. umbratus* above the Augrabies Falls were indicated not to hybridise naturally and thus are secure from direct invasion of hybrids because the 60-metre-high Augrabies Falls acts as a barrier to the potential spread of hybrids upstream. The populations are not, however, secure from translocations (e.g., bait bucket transfers of juveniles) that might or might not introduce the risk of hybridisation. *Labeo umbratus* has never been recorded from the lower Orange River (Skelton, 1986; Hay, 1991; Van Zyl, 1991), possibly because the river below the Augrabies Falls has reduced habitat diversity (Skelton, 1986) or the falls are a barrier. The absence of *L. umbratus* could explain why hybridisation has not been detected in this area. If *L. umbratus* cannot adapt to the lower Orange River, it is possible that there could be selection against the establishment of hybrids with *L. umbratus* characteristics. The absence of *L. umbratus* from the lower Orange River may suggest that *L. umbratus* was introduced into the Hardap impoundment. There are records of the translocation of *L. umbratus* elsewhere in South Africa (e.g., De Moor & Bruton, 1988) and the possibility that this species was introduced into the Hardap impoundment cannot be excluded. Such an introduction of relatively few fish could have sparked hybridisation with *L. capensis* due to the scarcity of conspecific mates. Dowling *et al.* (1989) stated that the less abundant species tends to possess more introgressed alleles than the more common species, which is consistent with the pattern observed in the Hardap impoundment.

Hybridisation in the Darlington Impoundment

The introduction of *L. capensis* from the Orange River system into the Great Fish and Sundays river systems has led to interspecific hybridisation with the indigenous *L. umbratus* in the Darlington impoundment. Relatively recent or limited hybridisation was indicated, with only one potential F_1 hybrid identified. This finding suggests that the present frequency of hybridisation in this impoundment may be lower compared with that in the Hardap impoundment. Despite sampling efforts being biased towards location of potential hybrids (looking for individuals with phenotypes that differed from potentially pure species) and potentially pure *L. capensis* specimens in the Darlington impoundment, mtDNA analysis detected *L. capensis* alleles in only 11 individuals (23%) compared with indigenous southern *L. umbratus* alleles occurring in 32 individuals (68%). Survey catch data from impoundments in the Eastern Cape also support the low number of *L. capensis* vs *L. umbratus* individuals (Table 4.9, Fig. 4.12). In addition, Orange River *L. umbratus* alleles were detected in four individuals (8.5%), confirming that genes of both Orange River species had reached the Darlington impoundment in the Sundays River system. The results also indicated that fixation of one parental mtDNA genome had not occurred at this locality, as was observed for the Hardap impoundment.

Putative hybrid in the Darlington impoundment had mtDNA of the indigenous *L. umbratus* lineage, indicating that females of this species can breed with *L. capensis* males. This may be due to the low number (38) of *L. capensis* individuals in the Darlington Dam (Weyl *et al.*, 2009) (Table 4.9), which might have made it difficult for *L. capensis* males to find conspecific females to mate with. In addition, construction of the IBWT into the previously seasonal mainstream regions of the Great Fish and

Sundays rivers has transformed these systems into permanently flowing systems (Laurenson & Hocutt, 1985). This transformation might favour *L. capensis*, as this species prefers fast-flowing waters of the mainstream (Skelton, 1986). The IBWT provides a constant source of Orange River propagules, so it is likely that *L. capensis* will eventually establish in the Great Fish and Sundays river systems (see Woodford *et al.*, 2013) and hybridisation between the two species may become more common.

TABLE 4.9. Catch data from impoundments for *L. capensis* and *L. umbratus*. The locations of the South African impoundments are shown in Figure 4.12.

No.	Impoundment	Nearest town	*L. capensis* mass (kg)	*L. umbratus* mass (kg)	*L. capensis* no. of individuals	*L. umbratus* no. of individuals
1	Gariep	Gariep Town	662.955	21.135		
2	Grassridge	Hofmeyr Town	31.877	4.807	103	10
3	Darlington	Jansenville	32.3	518.92	38	692
4	Glen Melville	Grahamstown		12.218		21
5	Pikoli	Grahamstown		247.747		459
6	Mangazana	Grahamstown		47.444		34
7	Tyefu	Ndlambe		17.829		52
8	Community Dam	Fort Hare		34.5894		177
9	Lombard	Fort Hare		29.918		84
9	Laing	King William's Town		341.792		798
10	Dimbaza	King William's Town		168.589		533
12	Hardap	Mariental (Namibia)			52	14

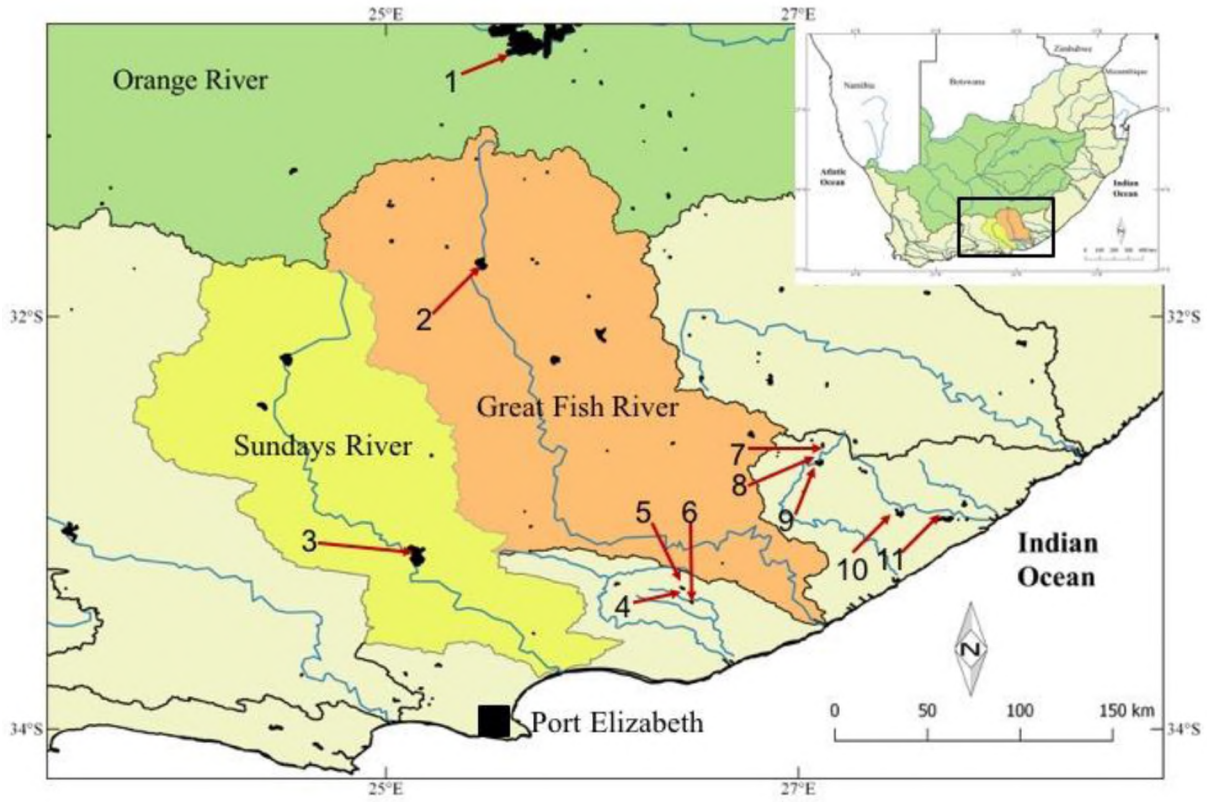


FIG. 4.12. Locations of impoundments in the Eastern Cape listed in Table 4.9. 1= Gariep, 2= Grassridge, 3= Darlington, 4= Glen Melville, 5= Pikoli, 6= Mangazana, 7= Tyefu, 8= Community Dam, 9= Lombard, 10= Laing and 11= Dimbaza.

Intraspecific hybridisation between *L. umbratus* from the Orange River system and individuals from populations of southern-flowing river systems was not detected. However, this was due to the lack of a suitable method for identification of intergrades, as differences in morphology and nDNA sequences between pure populations of these two areas were inadequate for detection of potential hybrids. The only method used that could reliably discriminate the two lineages of *L. umbratus* was mtDNA, which is incapable of detecting interbreeding between different lineages, unless it is combined with evidence from other data sources (Scribner *et al*., 2001). The *Labeo* population in the Darlington impoundment therefore is indicated to be undergoing Type 4

hybridisation (i.e., hybridisation without introgression) or it could be Type 5 hybridisation (i.e., widespread introgression) but insufficient specimens were sampled to detect a higher frequency of hybrids. Additional data is needed to test this hypothesis, especially additional nDNA sequence data.

Loss of local adaptation due to intraspecific hybridisation, as discussed by Allendorf *et al.* (2001), maybe be applicable for the *L. umbratus* population in southern-flowing river systems. In these river systems (including the Great Fish and Sundays river systems), *L. umbratus* had to adapt to the harsh conditions of these rivers, such as extreme seasonal flow regimes (Laurenson & Hocutt, 1985; Roux *et al.*, 2002). Possible interbreeding between *L. umbratus* from the Orange River and *L. umbratus* from the Great Fish and Sundays river systems may negatively impact on local adaptations to the southern habitats, but could also increase genetic variation (Allendorf *et al.*, 2001). The Great Fish and Sundays river systems are components of the *L. umbratus* Sundays+Bushmans+Great Fish management unit that requires protection from the threat of hybridisation. As controlling the flow of propagules via the IBWT is not a viable management option (Woodford *et al.*, 2013), conservation action should focus on preventing the spread of fish from the lower 'invaded' reaches to above dams, which isolate the currently pure *L. umbratus* in some tributaries of the Great Fish and Sundays river systems. Thus, no movement of this species should be permitted within or between catchments.

Conclusion

Evidence presented in this chapter supported the ongoing hybridisation between *L. capensis* and *L. umbratus* in the Hardap impoundment on the Orange River, and between *L. capensis* and *L. umbratus* in the Darlington impoundment in the southern-flowing river systems. Hybrids possessed the mtDNA of the most abundant species and most putative F_1 hybrids were morphologically intermediate between the parental species, whereas few putative hybrids were grouped close to the abundant parental species. The genetic integrity of the *L. umbratus* Sundays+Bushmans+Great Fish management unit is at risk if this hybridisation is not contained by precluding translocation of fish from the Great Fish and Sundays river systems to the other southern-flowing river systems.

CHAPTER FIVE: GENERAL DISCUSSION AND RECOMMENDATIONS

Labeo species are important components of aquatic ecosystems and are exploited for subsistence and recreational fisheries and for aquaculture (Reid, 1985; Skelton, 2001; Booth & Weyl, 2004). In South Africa *Labeo* are among the largest native fishes in the rivers in which they occur and, as a result of their algivorous and detritivorous feeding habits, fulfil an important role in controlling algae and cycling nutrients in aquatic ecosystems (Skelton, 2001). However, many *Labeo* spp. are subject to a variety of threats, which include water pollution, habitat degradation and interactions with non-native species (IUCN, 2012). Although *L. umbratus* is currently classified as Least Concern under the IUCN Red List criteria, this species is under threat of hybridisation with introduced congeneric species. A better understanding of the conservation implications of this threat is therefore required. In this regard, this thesis contributes to an improved understanding of *Labeo umbratus*, by contextualising the phylogenetic relationships of this species among other southern African *Labeo* spp., assessing phylogeographic patterns among *L. umbratus* lineages, and evaluating the potential impact of interspecific hybridisation resulting from the translocation of congeneric species between river systems.

Phylogenetic relationships among southern African *Labeo* species

Prior to the present study, the genetic relationships among the southern African *Labeo* species and the species groups proposed by Reid (1985) on the basis of morphological assessment [i.e. *Labeo forskalii* group (LFG), *Labeo niloticus* group (LNG), *Labeo coubie* group (LCG) and *Labeo umbratus* group (LUG)] were uncertain. As a result, the level of threat from human-induced impacts, such as the introduction of non-native *Labeo* spp. into river systems and the subsequent risk of interspecific hybridisation, could not be assessed. This thesis has, however, has made considerable contributions towards better understanding of the evolutionary relationships of African *Labeo* spp. by clarifying the phylogenetic affinities of the southern African *Labeo* spp. in relation to other African species.

In Chapter 2, for example, five monophyletic lineages, which corresponded to the previously proposed *Labeo niloticus*, *Labeo forskalii*, *Labeo macrostoma*, *Labeo coubie* and *Labeo umbratus* species groups (Reid, 1985), as well as two additional groups – a *Labeo ruddi/Labeo vulgaris* group and a divergent *Labeo batesii* group – were resolved in phylogenetic analyses of mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) sequence data. Resolution of the latter two additional groups also contributes to refinement of the classification of African *Labeo* spp. and provides evidence for a potential increase to the number of species groups from six to eight. Reid (1985) placed *L. congoro* in the LCG *sensu lato* based on morphological similarities, but the present molecular phylogenetic analysis (see Chapter 2) indicates that *L. congoro* shows greatest genetic affinity with the LNG *sensu lato*. The results also suggest that, contrary to placement of *L. ruddi* in the LNG by Reid (1985), *L. ruddi* and *L. vulgaris* together comprise a distinct phylogenetic lineage.

The phylogenetic assessment supported the species groups proposed by Reid (1985) for the remainder of the southern African species. Of particular interest is the *Labeo umbratus* group, for which the genus name *Abrostomus* Smith 1841 is available, based on the description of *Abrostomus umbratus* and *Abrostomus capensis* by Smith (1841).

This species group may warrant segregation at the genus level because (1) members are morphologically distinct from other *Labeo* groups (see Reid, 1985), (2) the members of this group are genetically distinct from other *Labeo* groups (see Chapter 2), and (3) the group is restricted to the southern temperate region of South Africa, where it is geographically separated from other species groups except in the Tugela River system, where *L. rubromaculatus* co-occurs with *L. molybdinus* (the latter species is a member of the *Labeo forskalii* group).

Labeo umbratus is of particular interest among the southern African *Labeo* species because, unlike the other members of the LUG which have restricted distributions (*Labeo capensis* is endemic to the Orange River, *L. rubromaculatus* is endemic to the Tugela River and *L. seeberi* is restricted to the Olifants River (Du Plessis, 1963; Swartz & Impson, 2007), *L. umbratus* has a relatively widespread distribution that includes the Orange River system as well as the southward-flowing Gourits, Gamtoos, Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon rivers (Swartz & Impson, 2007; see Chapter 3). Understanding the phylogeographic relationships and potential genetic isolation of *L. umbratus* lineages is important for the development of a species conservation plan and management plans for certain (if not all) river systems within its distribution.

Phylogeography of *Labeo umbratus*

The Orange River and the southward-flowing river systems inhabited by *L. umbratus* are considered to have experienced past geological, climatic and sea-level changes that might have influenced the distribution and genetic structure of this species (Swartz *et al.* 2007; Cowling *et al.*, 2009). The phylogeographic analysis of populations of *L. umbratus* in the southward-flowing river systems, presented in Chapter 3, revealed evidence for genetic structuring within *L. umbratus* and that the species encompasses two genetically distinct and geographically isolated lineages, one in the westward-flowing Orange River and the second 'southern lineage' in the southward-flowing rivers of South Africa. As these two lineages were reciprocally monophyletic, it was hypothesised that they represented Evolutionarily Significant Units (ESUs). In addition, the "southern lineage" could be divided further into two sublineages, namely southwestern (Gourits and Gamtoos) and southeastern (Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon) lineages. The two sublineages could also be further divided into four Management Units (MUs) [Gourits (2 unique haplotypes), Gamtoos (3 unique haplotypes), Sundays+Bushmans+Great Fish (2 unique haplotypes) and Keiskamma+Buffalo+Nahoon (3 unique haplotypes)], reflecting significant differences in haplotype frequencies, for conservation purposes. Based on the results of these genetic analyses and those of other studies on fish species that are widespread in isolated river systems (Swartz *et al.*, 2007; Chakona *et al.*, 2013a), continuation of such studies is advisable as additional hidden genetic diversity may be uncovered.

The populations of *L. umbratus* that occur in the Buffalo and Nahoon rivers were previously suspected to have been introduced by anglers as bait (Jubb, 1964). Contrary to Jubb's (1964) hypothesis, phylogeographic analyses indicated that the precursors of

these populations most likely dispersed to these river systems naturally via flood-plain connections during flooding events [~6,000–8,000 years ago (ya)] due to mesic conditions (Partridge *et al.*, 1999) and via a paleo-river connection between the Buffalo and Nahoon rivers during the Last Glacial Maximum (18,000 ya) when the sea level was considerably lower than present (–300 m) (Siesser & Dingle, 1981).

Hammer *et al.* (2013) considered geographic speciation to be one of primary evolutionary drivers of speciation in freshwater fishes. These authors stated that this type of speciation occurs when populations become separated from each other via fragmentation of the original range, or after dispersal. Three modes of speciation are potentially responsible for geographic speciation: allopatric speciation (where populations are completely isolated from each other); parapatric speciation (where populations are partially isolated from each other); and peripatric speciation (where populations are isolated at the periphery of their ranges). In the present phylogeographic analysis of *L. umbratus*, the populations from the Orange, Gourits and Gamtoos rivers were indicated to have been completely isolated for a sufficiently long period (152,000–1,760,000 ya) to have undergone genetic differentiation and are currently still isolated from each other, and from the other southeastern river systems (Sundays, Bushmans, Great Fish, Keiskamma, Buffalo and Nahoon). Thus, allopatric separation would be responsible for their evolutionary divergence. The southeastern populations of *L. umbratus* shared alleles. This may reflect ancestral polymorphism or translocation as the southward-flowing river systems are currently isolated and mesic conditions similar to the past climate have not been experienced since the Holocene (6,000–8,000 ya) (Partridge *et al.*, 1999), which would have led to heavy flooding and thus connection of adjacent drainage systems and facilitated gene flow. Ancestral polymorphism is the

process whereby separate populations retain alleles derived from the common ancestor. The results presented in Chapter 3 supported ancestral polymorphism as the most likely process responsible for the shared alleles among southeastern populations of *L. umbratus*. The results also showed that genetic structuring within *L. umbratus* is partitioned according to contemporary river systems. Human-mediated movement of fish would therefore be required to explain the latter results.

The southward-flowing river systems were indicated to be connected through inter-drainage connections in low-lying areas and through one paleo-river connection between the Buffalo and Nahoon rivers, which thus would have facilitated the dispersal of southern populations of *L. umbratus* (see Chapter 3). This finding supports the hypothesised role played by past climatic and geological changes, especially in southern Africa, as indicated by previous studies on small bodied (7-13 cm SL) *Pseudobarbus*, *Galaxias* and *Sandelia* spp. in the southward-flowing rivers of South Africa (Swartz *et al.*, 2007; Chakona *et al.*, 2013a). While these fishes were also indicated to have dispersed between river systems via paleo-river connection and via inter-drainage connections, their distributions stretched to higher-altitude streams where the possibility of river capture provided an alternative hypothesis. Dispersal of *L. umbratus*, however, is only likely to have occurred via inter-drainage and paleo-river connections because the fish are larger (maximum SL 40–50 cm) and occur mostly in the main channels of rivers or larger tributaries. A comparative analysis between *L. umbratus* and other *Labeo* spp., such as *L. cylindricus* and *L. molybdinus*, that occur both in the main channels and the headwaters of currently isolated eastward-flowing river systems (Tugela, Pongola, Incomati, Limpopo, Save, Buzi, Pungwe and Zambezi) is therefore warranted because, under comparable environmental conditions, *Labeo* spp. migrate

similarly and would most likely be subject to similar drivers of dispersal (Bell-Cross & Minshull, 1988; Cambray, 1990; Skelton, 2001).

Hybridisation of *Labeo* species

Prior to the research presented in Chapter 4, hybridisation among African *Labeo* spp. had only been documented using morphological characters and allozymes to discriminate individuals of *L. capensis*, *L. umbratus* and their hybrids in the Hardap impoundment (Gaigher & Bloemhof, 1975; Van Vuuren *et al.*, 1989, 1990). A major constraint of these studies was that they did not use pure reference populations for comparison and they used samples from only one locality (the Hardap impoundment). The research presented in Chapter 4 is therefore the first comprehensive assessment to use morphological and genetic data to discriminate pure *L. capensis* and *L. umbratus* and their putative hybrids with a relatively large sample size across their distributional range. In addition, it is the first study to provide evidence that genes from both *L. capensis* and *L. umbratus* from the Orange River system have entered the distributional range of the 'southern lineage' in the Sundays River system due to translocation of fish via the inter-basin water transfer schemes (Orange-Fish and Cookhouse tunnels), and that *L. capensis* and southern *L. umbratus* have hybridised in the Darlington impoundment (Chapter 4).

As the genetic evidence for interspecific hybridisation in the Darlington impoundment is currently based on one F_1 hybrid (Chapter 4), future sampling should be extended to include also the main channels of the Sundays and Great Fish rivers because hybrids may thrive in certain habitats (Barton & Hewitt, 1985). In addition, future research should use other types of reliable nuclear molecular makers for identification of

hybrids, such as co-dominant PCR markers [microsatellites and single-nucleotide polymorphisms (SNP)] or PCR-restriction fragment length polymorphisms (RFLP) (Hashimoto *et al.*, 2012). Co-dominant markers are superior to PCR-RFLP and multiplex-PCR for identification of hybrids but are time-consuming and expensive to develop (Hashimoto *et al.*, 2012).

Allendorf *et al.* (2001) categorised anthropogenic hybridisation into three types: hybridisation without introgression; widespread introgression; and complete admixture. The first type happens if species hybridise to produce only F_1 hybrids. The second type happens when hybrids interbreed with other hybrids and backcross with the parental species. The third type happens when few or no individuals of the parental populations remain. The analysis of morphological and genetic data presented in Chapter 4 revealed that hybridisation that is occurring in the Hardap impoundment can be characterised as either widespread introgression, as some individuals that carry a parental genome remain, or complete admixture. In contrast, hybridisation in the Darlington impoundment is indicated to be at an early stage (hybridisation without introgression) as only one possible F_1 hybrid was identified, although hybridisation may be more widespread in the impoundment but insufficient specimens were sampled to detect a higher frequency of hybrids.

The results presented in this thesis demonstrate that morphology and genetics are useful tools in combination to identify putative hybrids between *L. capensis* and *L. umbratus*. In some instances, genetic data alone was adequate to identify putative hybrids; for example, F_1 hybrids in the Hardap and Darlington impoundments were identifiable from nuclear S7 ribosomal intron 1 sequence data alone, because the fish were

heterozygous for loci at which the parental species showed fixed genetic differences. In other instances, specimens that were identified as *L. umbratus* from the nDNA data were found to have *L. capensis* mtDNA, which was evidence of introgressive hybridisation, but the stage of anthropogenic hybridisation represented cannot presently be determined. Such a pattern was also documented in the phylogenetic analysis of southern African *Labeo* spp. (Chapter 2). A specimen of *Labeo rubromaculatus* as identified by nDNA was found to have *L. capensis* mtDNA, and specimens of *L. cylindricus* from the Bua River (Malawi) and *L. molybdinus* from Revue River system (Mozambique) as identified by nDNA were found to also have *L. lunatus* mtDNA. As previously mentioned, future studies should use other nuclear markers for better detection of hybrids.

Divergence estimates revealed that the some of the hybridising species (e.g., *L. umbratus* and *L. capensis*) are closely related and belong to the same evolutionary lineage, whereas others (e.g., *L. cylindricus* and *L. lunatus*) are distantly related between lineages or belong to different species groups that have recently diverged (e.g., *Labeo forskalii* group and *Labeo niloticus* group). Thus, hybridisation, or at least the potential for hybridisation, might be relatively common among *Labeo* spp. Future research could explore the extent and consequences of natural and human mediated hybridisation events between other *Labeo* spp., which might better contextualise the conservation risk posed by interspecific hybridisation.

Scribner *et al.* (2001) included fish introductions as one of the contributing factors that can facilitate hybridisation. As *L. umbratus* does not occur naturally below the Augrabies Falls on the Orange River, one possible pathway for its introduction into the

Hardap impoundment was via direct stocking (Skelton, 2001). According to Ellender & Weyl (2014) native fishes were translocated all over South Africa for conservation reasons in the period 1961–1980 and during this period the Hardap impoundment was built. *Labeo umbratus* could have been introduced into the Fish River at the Hardap impoundment (lower Orange River) for similar reasons.

In the Darlington impoundment *L. capensis* were introduced via an inter-basin water transfer scheme (IBTW). This IBWT facilitated the introduction of Orange River fauna firstly between the Orange and Great Fish river systems, and secondly between the Great Fish and Sundays river systems (Cambray & Jubb, 1977). As there is currently no evidence that *L. capensis* and *L. umbratus* hybridise in the middle and upper Orange River system, it is likely that the resultant mixture of the Orange River *L. capensis* with the southern lineage of *L. umbratus* led to the observed interspecific (*L. capensis* × *L. umbratus*) hybridisation. Similarly, the Tugela-Orange IBTWs could cause hybridisation between *L. rubromaculatus* and *L. capensis* in the Tugela River systems. The evidence of hybridisation between the two species in the Tugela River system, however, was consistent with historical hybridisation rather than recent, and thus could not be due to Tugela-Orange IBTW as hypothesised. Future studies should properly assess the hybridisation in Tugela River system and the spread of hybridisation in the Sundays and Great Fish river systems using the other reliable molecular marker methods as mentioned above.

Conservation concerns and recommendations

Hybridisation is a potential cause of species extinction, especially of rare species when they come into contact with a more abundant species (Allendorf *et al.*, 2001). In

southern Africa *L. capensis* and *L. umbratus* have been shown to hybridise in impoundments in the Orange River (Hardap Dam) and in the Sundays River (Darlington Dam) (see Chapter 4). Such hybridisation may pose a threat to the native *Labeo* populations of the respective river systems. The ongoing hybridisation in the Hardap impoundment poses a low risk for contamination of the middle and upper Orange River *L. umbratus* populations as *L. umbratus* is not known to occur in the lower Orange River (Skelton, 1986; Hay, 1991; Van Zyl, 1991) and the 60-m-high Augrabies Falls acts as a barrier between the lower reaches of the river and the remainder of the Orange River system. Conversely, as there is a high likelihood that *L. umbratus* were introduced into Hardap Dam, there is a significant risk to the *L. capensis* populations of the lower Orange River. Thus, there is a need to determine the extent of threat that hybridisation poses in this system.

The research presented in Chapter 3 demonstrates that *L. umbratus* consists of two unique lineages, which I recommend should be managed as distinct ESUs because they are reciprocally monophyletic and have been historically isolated for between 608,000 and 1,760,000 years. Loss of either ESU would lead to a significant loss of overall genetic diversity of the species, and possibly of unique evolutionary potential for future adaptation. The genetic integrity of the southern ESU is under threat due to hybridisation with *L. capensis* (at least in Darlington Dam) and via invasion of the Orange River lineage of *L. umbratus* into the Great Fish and Sundays river systems. Two impoundments, the Kat River Dam on the Great Fish River and Slagboom Dam on the Sundays River systems still harbour pure *L. umbratus* populations representative of the two systems. To protect these populations of *L. umbratus* from complete

introgression, translocation of water or fish from downstream should be avoided, both into and upstream of the Kat River and Slagboom impoundments.

Recommendations for further research into the biogeography of the southern African *Labeo* species

Molecular dating of the Bayesian inference phylogeny estimated that all *Labeo* species groups (LFG, LUG, LNG, LCG, *L. batesii* and *L. ruddi*/*L. vulgaris*) diverged around 2.09–5.58 million years ago (mya), during a period associated with landscape changes, warping and rifting of the East African Rift System and the sickle-shaped Congo basin. The first species group to diverge was LFG, which was followed by divergence between the common ancestor of LUG+LNG+*L. batesii* and LCG+*L. ruddi*/*L. vulgaris* lineages. This result concurs with estimates in previous studies on other fish species indigent to this region: *Mastacembelus* eels (2.5–6.0 mya) (Brown *et al.*, 2010); African tigerfish (*Hydrocynus*), divergence (1.3–5.3 mya) between *H. tanzaniae* and the *H. vittatus* complex (Okavango, Zambezi, Buzi, Save, Incomati and Pongola) (Goodier *et al.*, 2011); *Synodontis* catfishes (1.9–9.6 mya) (Pinton *et al.*, 2013); and divergence of the southern and eastern lineages of *Petrocephalus* and a divergence event at 5.1 mya that separated *Marcusenius livingstoni*, which occurs in parts of the Malawi-Shire and Ruvuma river systems, from all other *Marcusenius* spp. (Maake, 2014). During this period (the Pliocene Epoch) the common ancestors of the *Labeo* spp. groups (except LFG) may have first migrated into the Orange River basin via the southward-flowing Okavango system (Jubb & Farguharson, 1965). The common ancestor of lineages B1 (*L. cylindricus*) and B2 (*L. molybdinus*, *L. ansorgii* and *L. lunatus*) of the LFG was estimated to have diverged from lineage B3 (undescribed *Labeo* sp.) from the Kwanza River system during the Pleistocene (1.7–2.78 mya). This finding could imply that

lineages B1 and B2 formed part of the second invasion into southern Africa, as proposed by Jubb & Farguharson (1965), and might explain the overlap in distribution with other southern African *Labeo* spp. groups. A comprehensive comparative biogeographic analysis is necessary to properly understand the patterns of distributions of *Labeo* spp. in relation to landscape changes in southern and east Africa.

Challenges and recommendations for future research

Several challenges were experienced with the genetic assessment of *Labeo* presented in this thesis. For example, the S7 ribosomal protein intron 1 gene was difficult to amplify as most sequence traces had background noise and thus were difficult to read or were too short (<400 bp) to be used together with other sequences for the analysis. The few specimens for which the S7 gene was amplified successfully were used in Chapters 3 and 4. Additional nDNA sequence data that correspond with available mtDNA data are needed, but due to financial constraints and time limitation, only available data was used. In Chapter 3, additional nDNA data would have enabled better understanding of allelic variation and the phylogeographic history of populations inhabiting southern-flowing river systems. In Chapter 4, additional nDNA data might have enabled detection of additional hybrid individuals because nDNA sequence data are better suited for detection of hybrids due to biparental gene inheritance. In Chapter 2, the paucity of S7 sequences and limited taxonomic coverage resulted in exclusion of this gene from the analysis. It is possible that inclusion of sequence data for this gene may have led to a difference in interpretation, as the S7 gene evolves faster than the Rag1 gene (Bufalino & Mayden, 2010). Additional polymorphic nuclear genes (e.g., those genes mentioned above) that are readily amplified are needed for future studies. Attempts to extract DNA from certain old tissue samples (e.g., *L. rosae*) stored in the

SAIAB collection facility were unsuccessful. This was probably because the DNA in the tissue was degraded. It is advisable to use fresh tissue samples for successful DNA extraction. However, financial constraints hampered sampling trips to collect *L. ruddi* and *L. rosae* from the Limpopo, Incomati and Pongola river systems. In addition to financial constraints, *L. ruddi* could not be located at the type locality or neighbouring localities by researchers from the University of Limpopo and SAIAB sampling in the area.

Conclusion

This thesis has contributed to ongoing resolution of phylogenetic relationships of the African *Labeo* species by providing data on the relationships and evolutionary history of southern African *Labeo* species, with particular focus on the *Labeo umbratus* species group. *Labeo umbratus* was shown to comprise two evolutionary lineages that inhabit the Orange and the southward-flowing river systems, respectively. For the southern lineages, conservation action needs to be directed at the genetic conservation units (ESUs or MUs) that were identified for *L. umbratus*. This is particularly pertinent because of the demonstrated threat of hybridisation between *L. capensis* and *L. umbratus*. It is however important that future research investigates the extent of hybridisation in the extralimmital and native ranges of southern African *Labeo* spp..

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, Boughman, N. J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P, Möst, M., Mullen, S., Nichols, R., Nolte, A. W., Parisod. C., Pfennig, K., Rice, A. M., Ritchie, M. G., Seifert, B., Smadja, C. M., Stelkens, R., Szymura, J. M., Väinölä, R., Wolf, J. B. W. & Zinne, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology* **26(2)**, 229–46. doi: 10.1111/j.1420-9101.2012.02599.x PubMed PMID: WOS:000313747600001.

Akaike, H. (1974) "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* **19 (6)**, 716–723, doi:10.1109/TAC.1974.1100705,MR 0423716.

All Cypriniformes Species Inventory. <https://www.flmnh.ufl.edu/fish/freshwater-research/all-cypriniformes/>. (Accessed April 2016).

Allendorf, F. W., Leary, R. F., Spruell, P. & Wenburg, J. K. (2001). The problem with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution* **16**, 613–622.

Allu, P. K. R., Chakraborty, B., Das, M., Mahapatra, N. R. & Ghosh, S. (2014). PCR-based segregation of one hybrid variety of *Labeo rohita* and *Catla catla* from their wild-types. *Aquaculture International* **22**, 775–782.

Almeida, D., Ribeiro, F., Leunda, P. M., Vilizzi, L & Copp, G. H. (2013). Effectiveness of FISK, an invasiveness screening tool for non-native freshwater fishes, to

perform risk identification assessments in the Iberian Peninsula. *Risk Anal*ysis **33 (8)**, 1404–1413. http://0-dx.doi.org.wam.seals.ac.za/10.1111/risa.12050.

Andersson, L. (1990). The driving force: Species concepts and ecology. *Taxon* **39**,375–382.

Argue, B. J. & Dunham, R. A. (1999). Hybrids fertility, introgression, and backcrossing in fish. *Reviews in Fisheries Science* **7(3-4)**, 137-195. DOI: 10.1080/10641269908951360

Arnold, M. L. & Martin, N. H. (2010). Hybrid fitness across time and habitats. *Trends in Ecology and Evolution* **25(9)**, 530–536.

Arunachalam, M., Nandagopal, S. & Mayden, R. L. (2013). Morphological diagnoses of garra (Cypriniformes: Cyprinidae) from northeastern India with four new species description from Brahmaputra River. *Journal of Fisheries and Aquaculture* **4(3)**, 121–138.

Aubert, J. & Solignac, M. (1990). Experimental evidence for mitochondrial DNA introgression between *Drosophila* species. *Evolution* **44**, 1272–1282.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology, Evolution, and Systematics* **18**, 489–522.

Avise, J. C. & Ball Jr, R. M. (1990). Principles of genealogical concor dance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* **7**:45–67.

Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7**, 371–379.

Avise, J. C. (2000). Phylogeography: the history and formation of species. Harvard University Press, Cambridge, p 447.

Avise, J. C. (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography* **36**, 3–15.

Balon, E. K. (1974). Fishes from the edge of Victoria Falls, Africa: demise of a physical barrier for downstream invasions . *Copeia* **3**, 643–660.

Barlow, G. W. (2002). How behavioural studies contribute to the species problem: a piscine perspective. *Fish and Fisheries* 3, 197–212.

Barnard, K. H. (1943). Revision of the indigenous freshwater of the S.W. Cape region. *Annals of the South African Museum* **36(2)**, 102–262.

Barraclough, T. G. & Nee, S. (2001). Phylogenetics and speciation. *Trends in Ecology & Evolution* **16**, 391–399.

Barton, N. H. & Hewitt, G. M. (1985). "Analysis of hybrid zones". *Annual Review of Ecology and Systematics* **16**, 113–148.

Baum, D. A., and K. L. Shaw. 1995. Genealogical perspectives on the species problem. In *Experimental and molecular approaches to plant biosystematics* (Hoch, P. C. & Stephenson, A. G., eds.). pp. 289–303. Missouri Botanical Garden, St. Louis.

Bayona, J., Bills, R., Cambray, J., Hanssens, M., Kazembe, J. & Marshall, B. (2010). *Labeo congoro*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Bell-Cross, G. (1976). *The fishes of Rhodesia*. National Museums and Monuments of Rhodesia, Salisbury.

Bell-Cross, G. & Minshull, J. L. (1988). The fishes of Zimbabwe. National Museums and Monuments of Zimbabwe, Harare, Zimbabwe.

Bernatchez, L., Glemet, H. & Wilson, C. C. (1995). Introgression and fixation of Arctic char (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 179–185.

Bills, R. & Cambray, J. (2007). *Labeo molybdinus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed July 2015).

Bills, R., Engelbrecht, J., Marshall, B.E. & Cambray, J. (2007). *Labeo rosae*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Bills, R., Marshall, B., Moelants, T. & Vreven, E. (2010a). *Labeo altivelis*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Bills, R., Cambray, J., Getahun, A., Hanssens, M., Marshall, B., Moelants, T. & Ntakimazi, G. (2010b). *Labeo cylindricus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed July 2015).

Bolnick, D. I. (2009). Hybridisation and speciation in centrarchids. In *Centrarchid fishes: diversity, biology, and conservation* (Cooke, S. & Philipp, D. J., eds), pp. 39–41. Blackwell: Oxford.

Booth, A. J. & Weyl, O. L. F. (2004). Investigations into a potential directed fishery for *Labeo cylindricus* on Lake Chicamba, Mozambique. *African Zoology* 39(1), 77-90.

Boulenger, G. A. (1909-1916). Catalogue of the freshwater fishes of Africa in the British Museum (Natural History).

Boyer, M. C., Muhlfeld, C. C. & Allendorf, F. W. (2008). Rainbow trout (Oncorhynchus mykiss) invasion and the spread of hybridization with native

westslope cutthroat trout (Oncorhynchus clarkii lewisi). *Canadian Journal of Fisheries and Aquatic Sciences*. **65**, 658–669.

Brown, W. M., George, M. & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences* USA **76**, 1967–1971.

Brown, K. J., Ruber, L., Bills, R. & Day, J. J. (2010). *Mastacembelid* eels support Lake Tanganyika as an evolutionary hotspot of diversification. *BMC Evolutionary Biology* 10,188.

Bruton, M. N & Van As, J. G. (1986). Faunal invasions of aquatic ecosystems in southern Africa, with suggestions for their management. In: *The Ecology and Management of Biological Invasions in Southern Africa*. (MacDonald, I. A. W., Kruger, F. J. & Ferrar, A. A, eds),. Oxford University Press, Cape Town, South Africa, pp 47–61.

Bufalino, A. P. & Mayden, R. L. (2010). Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. *Molecular Phylogenetics and Evolution* **55**, 274–283.

Burridge, C. P., Craw, C. & Waters, J. M. (2007). An empirical test of freshwater vicariance via river capture. *Molecular Ecology* **16**, 1883–1895.

Burridge, C. P., Craw, C., Jack, D. C., King, T. M. & Waters, J. M. (2008). Does fish ecology predict dispersal across a river drainage divide? *Evolution* **62**, 1484–1499.

Butzer, K. W. & Helgren, D. M. (1972). Late Cenozoic evolution of the Cape Coast between Knysna and Cape St Francis, South Africa. *Quaternary Research*, **2**, 143–169.

Cain, A. J. (1963) 'The Natural Classification'. *Proceedings of the Linnean Society of London* **174**, 115–121.

Camardi, G. (2001). Richard Owen, Morphology and Evolution. *Journal of the History of Biology* **34,** 481–515.

Cambray, J. A. (2007). *Labeo rubromaculatus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed July 2015).

Cambray, J. A. (1990). Adaptive significance of a longitudinal migration by juvenile freshwater fish in the Gamtoos River system, South Africa. Suid Afrikaanse Tydskrif vir. *Natuurnavorsing* **20(4)**, 149–156.

Cambray, J. A. & Jubb, R. A. (1977). Dispersal of fishes via the Orange-Fish tunnel, South Africa. *Journal of the Limnological Society of Southern Africa* **3**, 33–35.

Cambray, J. A., Davies, B. R. & Ashton, P. J. (1986). The Orange-Vaal system. In: The ecology of river systems (Davies, B. R. & Walker, K. F., ed.), pp. 89–122. Dr W. Junk Publishers: Dordrecht, The Netherlands.

Cambray, J. A. (2000). Threatened fishes of the world' series, an update. *Environmental Biology of Fishes* **59**, 353–357.

Campton, D. E. (1987). Natural hybridization and introgression in fishes: methods of detection and genetic interpretations. In *Population Genetics and Fishery Management* (Ryman, N. & Utter, F., eds), pp. 161–192, University of Washington Press.

Carrizo, S. F., Smith, K. G. & Darwall, W. R. T. (2013): Progress towards a global assessment of the status of freshwater fishes (Pisces) for the IUCN Red List: application to conservation programmes in zoos and aquariums. *International Zoo Yearbook* **47**, 46–64.

Chakona, A. & Swartz, E. R. (2013). A new redfin species, *Pseudobarbus skeltoni* (Cyprinidae, Teleostei), from the Cape Floristic Region, South Africa. *Zootaxa* **3686**, 565–577. doi: 10.11646/zootaxa.3686.5.5

Chakona, A., Swartz, E. R., Gouws, G. & Bloomer, P. (2013a). A freshwater fish defies ancient mountain ranges and drainage divides: extrinsic and intrinsic influences on the evolutionary history of a recently identified galaxiid. *Journal of Biogeography* **20(7)**, 1365–2699.

Chakona, A., Swartz, E. R. & Gouws, G. (2013b). Evolutionary drivers of diversification and distribution of a southern temperate stream fish assemblage: Testing the role of historical isolation and spatial range expansion. *PLoS ONE* **8(8)**: e70953. doi:10.1371/journal.pone.0070953.

Chatfield, C. & Collins, A. J. (1980). Introduction to multivariate analysis. Chapman & Hall, London (1980).

Chen, Z, J. (2007). Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids. *Annual Reviews Plant Biology* **58**, 377–406.

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.

Chen, X. L., Yue, P.Q. & Lin, R. D. (1984). Major groups within the family Cyprinidae and their phylogenetic relationships. *Acta Zootaxon. Sin.* **9**, 424–440 (in Chinese with English summary).

Chow, S. & Hazama, K. (1998). Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology* **7**, 1247–1263.

Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657–1659.

Coetzee, L., Du Preez, H. H. & Van Vuren, J. H. J. (2002). Metal concentrations in *Clarias gariepinus* and *Labeo umbratus* from the Olifants and Klein Olifants River, Mpumalanga, South Africa: Zinc, copper, manganese, lead, chromium, nickel, aluminium and iron. *Water SA* **28(4)**, 433–448.

Cowling, R. M., Proches, S. & Partridge, T. C. (2009). Explaining the uniqueness of the Cape flora: Incorporating geomorphic evolution as a factor for explaining its diversification. *Molecular Phylogenetics and Evolution* **51**, 64–74.

Cracraft, J. (1983). Species concepts and speciation analysis. *Current Ornithology* **1**, 159–187.

Crandall, K. A., Bininda-Emonds, O. R. P., Mace, G. M. & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *TREE* **15(7)**, 290–295.

Craw, D., Burridge, C., Norris, R. & Waters, J. (2008). Genetic ages for Quaternary topographic evolution: a new dating tool. *Geology* **36**, 19–22.

Cucherousset, J & Olden, J. D. (2011). Ecological Impacts of Non-native Freshwater Fishes. *Fisheries* **36 (5)**, 215–230, DOI: 10.1080/03632415.2011.574578.

Cypriniformes Tree of Life. http://bio.slu.edu/mayden/cypriniformes/home.html. (Accessed April 2016).

Da Costa, L. (2007). *Labeo ansorgii*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methords* **9**, 772.

Darwall, W. R. T., Smith, K. G., Tweddle, D. & Skelton, P. (2008). The status and distribution of freshwater biodiversity in southern Africa. *IUCN Red List of Threatened Species* **2008**, 1–36.

Darwall, W. R. T., Smith, K. G., Tweddle, D. & Skelton, P. (2009). The status and distribution of freshwater biodiversity in southern Africa. *IUCN Red List of Threatened Species* **2009**, 1–36.

Davies, B. R., O'Keeffe, J. H. & Snaddon, C. D. (1993). *A synthesis of the ecological functioning, conservation and management of South African river ecosystems*. Water Research Commission, Pretoria, South Africa, Report No. WRC TT62/93.

De Moor, I. J. & Bruton, M. N. (1988). Atlas of alien and translocated indigenous aquatic animals of southern Africa. *South African National Scientific Programmes Report* **144**, 78–133.

Department of Water Affairs. (1978). https://www6.dwa.gov.za/iwrp/DSS/UserFiles/IWRP/Documents/Tugela%20Vaal%20Project%20Drainage%20Regions%20300%20and%202000.pdf. (Accessed July 2015).

De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In *Endless forms: Species and speciation* (Howard, D. J. & Berlocher,S. H., eds.). pp. 57–75. New York: Oxford University Press.

De Queiroz, K. (2005). Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences* 102, 6600–6607.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* **56(6)**, 879–886.

De Wit, M. C. J. (1993). Cainozoic evolution of drainage systems in the north-western Cape. University of Cape Town, Cape Town.

Dingle, R. V. & Hendey, Q. B. (1984). Mesozoic and tertiary sediment supply to the Western Cape basin and palaeo drainage systems in south-western Africa. *Marine Geology* **56**, 13–26.

Dobzhansky, T. (1950). Mendelian populations and their evolution. *American Naturalist* **84**, 401–418.

Dobzhansky, T. (1970). *Genetics of the evolutionary process*. New York: Columbia University Press.

Dollar, E. S. J. (1998) Palaeofluvial geomorphology in southern Africa: a review. *Progress in Physical Geography* **22(3)**, 325–349.

Donoghue, M. J. (1985). A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* **88**, 172–181.

Douglas, M. E., Miller, R. R. & Minckley, W. L. (1998). Multivariate Discrimination of Colorado Plateau *Gila* spp.: The "Art of Seeing Well" Revisited. *Transactions of the American Fisheries Society* 127(2) DOI:10.1577/15488659(1998)127<0163:MDOCPG>2.0.CO;2

Dowling, T. E., Smith, G. R. & Brown, W. (1989). Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (Family Cyprinidae): comparison of morphology, allozymes, and mitochondrial DNA. *Society for the Study of Evolution* **43**, 620–634.

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.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A-H., Soto, D., Melanie L. J. Stiassny, M. L. J. & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**, 163–182.

Du Plessis, S. S. (1963). Revision of the genus *Labeo* in the Transvaal, with description of a new species. *Annals of the Transvaal Museum* **24(4)**, 327–337.

Ellender, B. R., Weyl, O. L. F., Winker, H. & Booth, A. J. (2010). Quantifying the annual fish harvest from South Africa's largest freshwater reservoir. *Water South Africa* **36**, 45–51.

Ellender, B. R. & Weyl, O. L. F. (2014). A review of current knowledge, risk and ecological impacts associated with non-native freshwater fish introductions in South Africa. *Aquatic Invasions* **9(2)**, 117–132.

Elliott, N. G., Haskard, K. & Koslow, J. A. (1995). Morphometric analysis of orange roughy (*Hoplostethusatlanticus*) off the continental slope of southern Australia. *Journal of Fish Biology* **46**, 202–220.

Elmer, K. R., Dávila, J. A. & Lougheed, S. C. (2007). Cryptic diversity and deep divergence in an upper Amazonian leaflitter frog, *Eleutherodactylus ockendeni*. *BMC Evolutionary Biology* 7, 247 doi:10.1186/1471-2148-7-247.

Engelbrecht, J., Bills, R. & Marshall, B. E. (2007). *Labeo ruddi*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Eschmeyer, W. N. & Fong, J. D. (2015). Species by Family/Subfamily. Available at http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp. (Accessed May 2015).

Esposti, M., Devries, D. S., Crimi, M., Ghelli, A., Patarnello, T., Meyer, A. (1993). Mitochondrial cytochrome b evolution and structure of the protein. *Biochimica et Biophysica Acta* **1143**, 243–271.

Excoffier, L. & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–567.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131(2)**, 479–91. ISSN 0016-6731. PMC 1205020.PMID 1644282.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* **17**, 368-376.

Felsenstein, J. (2004). Inferring phylogenies. Sunderland (MA): Sinauer Associates.

Firmat, C., Alibert, P., Losseau, M., Baroiller, J. & Schliewen, U. K. (2013). Successive Invasion-Mediated Interspecific Hybridizations and Population Structure in the Endangered Cichlid *Oreochromis mossambicus*. *PLoS ONE* **8(5)**, e63880. doi:10.1371/journal.pone.0063880

Fishelson, L., Goren, M., Van Vuren, J. & Manelis, R. (1996). Some aspects of the reproductive biology of *Barbus* spp., *Capoeta damascina* and their hybrids (Cyprinidae, Teleostei) in Israel. *Hydrobiologia* **317(7)**, 9–88.

Flügel, T. J., Eckardt, F. D. & Cotterill, F. P. D. (2015). The Present Day Drainage Patterns of the Congo River System and their Neogene Evolution. In *Geology and Resource Potential of the Congo Basin*. (Maarten J. de Wit, M. J., Guillocheau, F. & de Wit, M. C. J., eds), pp.315–337. Springer: Berlin, Heidelberg.

Freyhof, J., Lieckfeldt, D., Pitra, C. & Ludwig, A. (2005). Molecules and morphology: Evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution* **37**, 347–354.

Friesen, V. L. (2000). Introns. In *Molecular Methods in Ecology* (Baker, A. J., ed). pp. 274–294. Blackwell Science, Oxford.

Froese, R & Pauly, D. (2000). FishBase 2000: concepts, design and data sources. ICLARM, Los Banos, Laguna, Philippines. p 344.

Fu, Y-X (1997). Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. *Genetics* **147(2)**, 915–925.

Funk, V. A., Hoch, P. C., Prather, L. A. & Wagner, W. L. (2005). The importance of vouchers. *Taxon* **54**, 127–129.

Gaigher, G. I. & Bloemhof, H. J. (1975). Hybrids *Labeo umbratus* x *L. capensis* from the Hardap *Dam*. *Madoqua* **9**, 53–55.

Gaigher, I. G. (1984). The reproduction of *Labeo umbratus* in Wuras Dam, a shallow turbid impoundment. *South African Journal of Zoology* **19**, 105–108.

Galtier, N., Nabholz, B., Glemin, S. & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* **18**, 4541–4550.

Gene, S. H. (2007). *Fish conservation: A guide to understanding and restoring global aquatic biodiversity and fishery resources*. Island press, USA.

Gernhard, T. (2008). The conditioned reconstructed process. *Journal of Theoretical Biology* **253**, 769–778.

Getahun, A. & Twongo, T. K. (2010). *Labeo forskalii*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed August 2015).

Gilchrist, J. D. F. & Thompson, W. W. (1913). The freshwater fishes of South Africa (continued). *Annuals of the South African Museum* **11(6)**, 346–364.

Gill, B. F., Slikas, B. & Sheldon, F. H. (2005). Phylogeny of titmice (Paridae): ii. Species relationships based on sequences of the mitochondrial cytochrome-*b* gene. *The Auk* **122(1)**, 121–143.

Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D. M. & Hood, M. E. (2008). "Speciation in Fungi". *Fungal Genetics and Biology* **45**, 791–802.doi:10.1016/j.fgb.2008.02.001.

Godbout, J. D., Aday, D. D., Rice, J. A., Bangs, M. R. & Quattro, J. M. (2009). Morphological models for identifying Largemouth Bass, Spotted Bass, and Largemouth × Spotted Bass hybrids. *North American Journal of Fisheries Management* **29**, 1425–1437.

Goddard, K. A. & Dawley, R. M. D. (1990). Clonal Inheritance of a Diploid Nuclear Genome by a Hybrid Freshwater Minnow (*Phoxinus eos-neogaeus*, Pisces: Cyprinidae). *Evolution* **44(4)**, 1052–1065.

Goodier, S. A. M., Cotterill, F. P. D., O'Ryan, C., Skelton, P. H. & de Wit, M. J. (2011). Cryptic diversity of African tigerfish (Genus Hydrocynus) reveals palaeogeographic signatures of linked Neogene geotectonic events. *PLoS One* **6(12)**. e28775. doi:10.1371/journal.pone.0028775.

Grant, W. S. & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89**, 415–426.

Grant, E. H. C., Lynch, H. J., Muneepeerakul, R., Arunachalam, M., Rodriguez-Iturbe, I. & Fagan, W. (2012). Inter-basin water transfer, riverine connectivity and spatial

controls on fish biodiversity. *PLoS ONE* **7(3)**: e34170. Doi:10.1371/journal.pone.0034170. (Accessed March 2009).

Guindon, S. & Gascuel, O. (2003). A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* **52**, 696–704.

Guo, X. & Chen, D. (2010). Comparative evolution of the mitochondrial cytochrome *b* gene and nuclear S7 ribosomal protein gene intron 1 in sinipercid fishes and their relatives. *Hydrobiologia* **649**, 139–156.

Gupta. J. & Van der Zaag. P. (2008). Interbasin water transfers and integrated water resources management: Where engineering, science and politics interlock. *Physics and chemistry of earth* **33**, 28–40.

Hall, B. G. (2011). Phylogenetic trees made easy: a how-to manual. 4th edition. Sinauer Associates: Sunderland (MA).

Hall, B. G. (2013). Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution* **30(5)**, 1229–1235.

Hamman, K. C. D. (1980). Post-impoundment trends in the fish populations of the Hendrik Verwoerd Dam, South Africa. *Journal of the Limnological Society of Southern Africa* **6**, 101–108.

Hamman, K. C. D., Thorne, S. C. & Scott, H.A. (1982). Preliminary report on the effects of the January 1981 flood damage on the indigenous fish fauna of certain sections of the Bree and Gourits River systems. Unpublished report, Cape Department of Nature and Environmental Conservation, Cape Town. 7pp.

Hammer, Ø., Harper, D. & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**, 9.

Hammer, M. P., Adams, M. & Hughes J. M. (2013). Evolutionary processes and biodiversity. In *Ecology of Australian Freshwater Fishes*, (Humphries, P & Walter, K., eds), pp. 26–48. CSIRO publishing: Australia.

Hanssens, M., Lalèyè, P. & Moelants, T. (2010). *Labeo parvus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed August 2015).

Harwood, D. M. (1985). Late Neogene climatic fluctuations in the southern high-latitudes: implications of a warm Pliocene and deglaciated Antarctic continent. *South African Journal of Science* **82**, 239–241.

Hasegawa, M., Kishino, H. & Yano, T. (1985). "Dating of human-ape splitting by a molecular clock of mitochondrial DNA". *Journal of Molecular Evolution* **22(2)**, 160–174.

Hashimoto, D. T., Senhorini, J. A., Foresti, F. & Porto-Foresti, F. (2012). Interspecific fish hybrids in Brazil: management of genetic resources for sustainable use. *Reviews in Aquaculture* **4(2)**, 108–118.

Hattingh, J. (1972). Observations on the blood physiology of five South African freshwater fish. *Journal of fish biology* **4(4)**, 555–563.

Hattingh, J. (1996). Fluvial response to allocyclic influences during the development of the lower Sundays River, Eastern Cape, South Africa. *Quaternary International* **33**, 3–10.

Hay, C. J. (1991). The distribution of fish in the Fish River, Namibia, *Madoqua* **17**: 211–215.

He, A., Huang, W., He, Y. & Yang, J. (2015). *Cophecheilus brevibarbatus*, a new labeonine fish from Guangxi, South China (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwaters* **25(4)**, 299–304.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & de Waard, J. R. (2003a). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. B **270**, 313–322. (DOI 10.1098/rspb.2002.2218.)

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. (2003b). Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Procceedings of Royal Society of London B* 270, 96–99.

Hedges, S. B. & Kumar, S. (2009). The timetree of life. Oxford University Press: New York.

Henning, W. (1966). Phylogenetic systematics. University of Illinois Press, Urbana, Illinois.

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F. & Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**, 291–301.

Hitt, N. P., Frissell, C. A., Muhlfeld, C. C. & Allendorf, F. W. (2003). Spread of hybridization between native westslope cutthroat trout, Oncorhynchus clarki lewisi, and non-native rainbow trout, Oncorhynchus mykiss. *Canadian Journal of Fisheries and Aquatic Sciences*. **60**, 1440–1451.

Ho, S. Y. W., Phillips, M. J., Cooper, A. & Drummond, A. J. (2005). Time Dependency of Molecular Rate Estimates and Systematic Overestimation of Recent Divergence Times. *Molecular Biology and Evolution* **22(7)**, 1561–1568.

Hou, Z., Fu, J. & Li, S (2007). A molecular phylogeny of the genus *Gammarus* (Crustacea: Amphipoda) based on mitochondrial and nuclear gene sequences. *Molecular Phylogenetics and Evolution* **45**, 596–611.

Houde, E. D. & Zastrow, C. E. (1993). Ecosystem- and taxon-specific dynamic energetics properties of fish larvae assemblages. *Bulletin of Marine Science*. **53(2)**, 290–335.

Howes, G. J. (1991). Systematics and boiogeography: an overview. In *Cyprinid fishes Systematics, biology and exploitation*, (Winfield, I. J. & Nelson, J. S., eds), pp. 1–33. Chapman & Hall: London.

Huelsenbeck, J. P. & Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17(8)**, 754–5.

International Union for Conservation of Nature and Natural Resources. (2012). *IUCN Red List Categories and Criteria. Version 3.1* (2nd edition). Gland: IUCN

IUCN (2016). IUCN Red List of Threatened Species™. Version 2016-1. Available at http://www.iucnredlist.org [accessed May, 2016]).

Ivanova, N. V., Zemlak. T. S., Hanner, R. H. & Hebert, P. N. (2007). Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes* **7(4),** 544–548.

Jubb, R. A. (1961). *An Illustrated Guide to the Freshwater Fishes of the Zambezi River, Lake Kariba, Sabi, Lundi, and Limpopo Rivers. Stuart Manning*: Bulawayo.

Jubb, R. A. (1964). Freshwater fishes and drainage basins in Southern Africa. *South African Journal of Science* **60(1)**, 17–21.

Jubb, R. A. & Farquharson, F. L. (1965). The freshwater fishes of the Orange River drainage basin. *South African Journal of Science* **61**, 118–125.

Jackson, P. B. N. & Coetzee, P. W. (1982). Spawning behaviour of *Labeo umbratus* (Pisces: Cyprinidae). *South African Journal of Science* **78**, 293–295.

Kadye, W. T. & Booth, A. J. (2013). An invader within an altered landscape: One catfish, two rivers and an inter-basin water transfer scheme. *River Research and Applications* **29(9)**, 1131–1146.

Karssing, R. J. (2008). Status of the KwaZulu-Natal yellowfish *Labeobarbus natalensis* (Castelnau, 1861). In *Technical Report on the State of Yellowfishes in South Africa* (Impson, N. D., Bills, I. R. & Wolhuter, L. eds), pp. 31–45. Water Research Commission, Gezina, Pretoria.

King, L. A., (1963). *The South African Scenery*. Oliver and Boyd, London, United Kingdom, 308 pp.

Kottelat, M. (2000). Diagnosis of a new genus and 64 new species of fishes from Laos (Teleostei: Cyprinidae, Balitoridae, Bagridae, Syngnathidae, Chaudhuriidae and Tetraodontidae). *Journal of South Asian Natural History* **5**, 37–82.

Kottelat, M. & Hui, T. H. (2008). A synopsis of the genus Lobocheilos in Java, Sumatra and Borneo, with descriptions of six new species (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwater* **19**, 27–58.

Koblmüller, S., Schliewen, U. K., Duftner, N., Sefc, K. M., Katongo, C. & Sturmbauer, C. (2008). Age and spread of the haplochromine cichlid fishes in Africa. *Molecular Phylogenetics and Evolution* **49**, 153–169.

Krupp, F. & Budd, K. (2009). A new species of the genus *Garra* (Teleostei: Cyprinidae) from Oman. *aqua, International Journal of Ichthyology* 15, 2–15.

Kullander, S. & Fang, F. (2004). Seven new species of Garra (Cyprinidae: Cyprininae) from the Rakhine Yoma, southern Myanmar. *Ichthyological Exploration of Freshwater* **15**, 257–278.

Lambert, D. M. & Spencer H. G. (1995). *Speciation and the recognition concept: Theory and application*. Baltimore: Johns Hopkins University Press.

Lanave, C., Preparata, G., Saccone, C. & Serio, G. (1984). A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* **20(1)**, 86-93.

Laplace, P. S. (1812). Théorie analytique des probabilités. Paris: Paris, Ve. Courcier,.

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.

Laurenson, L. B. J. & Hocutt, C. H. (1985). Colonisation theory and invasive biota: The Great Fish River, a case history. *Environmental Monitoring and Assessment* **6**, 71–90.

Laurenson, L. J. B., Hocutt, C. H. & Hecht, T. (1989). An evaluation of the success of invasive fish species of the Great Fish River. *Journal of Applied Ichthyology* **5**, 28–34.

Leprieur, F., Brosse, S., Garcı´a-Berthou, E., Oberdorff, T., Olden, J. D. & Townsend, C. R. (2009). Scientific uncertainty and the assessment of risks posed by non-native freshwater fishes. *FISH and FISHERIES* **10**, 88–97.

Lévêque, C**.** (1990). Cyprinidae. In *Faune des poissons d'eaux douces et saumâtres d'Afrique de l'Ouest. Tome I. Coll. Faune Tropicale n° XXVIII* (Lévêque, C., Paugy, D. & Teugels, G. G. eds.) pp. 269–361. Musée Royal de l'Afrique Centrale, Tervuren, and O.R.S.T.O.M., Paris.

Li, W.X., Ran, J.C., Chen, H.M., 2006. A new genus and species of Labeoninae from Guizhou, China. *Journal of Zhejiang Ocean University*. **26**, 1–2.

Li, F. L., Zhou, W. & Fu, Q. (2008). *Garra findolabium*, a new species of cyprinid fish (Teleostei: Cypriniformes) from the Red River drainage in Yunnan, China. *Zootaxa* **1743**, 62–68.

Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25(11)**, 1451–1452. doi:10.1093/bioinformatics/btp187.

Lin, C., Lin, H., Wang, J., Chao, S. & Chiang, T. (2010). Phylogeography of *Hemibarbus labeo* (Cyprinidae): secondary contact of ancient lineages of mtDNA. *Zoologica Scripta* **39**, 23–35.

Lintermans, M. (2004). Human assisted dispersal of alien freshwater fish in Australia. *New Zealand Journal of Marine and Freshwater Research*. **38**, 481–501.

López, J. A, Chen, W. & Ortı´, G. (2004). Esociform phylogeny. *Copeia* **2004**: 449–469.

Lothongkham, A., Arbsuwan, S. & Musikasinthorn, P. (2014). *Garra waensis*, a new cyprinid fish (Actinopterygii: Cypriniformes) from the Nan River basin of the Chao Phraya River system, northern Thailand. *Zootaxa* **3790(4)**, 543–554.

Lourens, L., Hilgen, F., Shackleton, N.J., Laskar, J., Wilson, D., (2004) "The Neogene Period". In *A Geologic Time Scale 2004*. (Gradstein, F., Ogg, J. & Smith, A. G. eds.), Cambridge: Cambridge University Press.

Losos, J. B & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *TRENDS in Ecology and Evolution* **18(5)**, 220–227.

Lowenstein, J. H., Osmundson, T. W., Becker, S., Hanner, R. & Stiassny, M. L. J. (2011). Incorporating DNA barcodes into a multi-year inventory of the fishes of the hyperdiverse Lower Congo River, with a multi-gene performance assessment of the genus *Labeo* as a case study. *Mitochondrial DNA* **22**(sup1), 52–70. doi: 10.3109/19401736.2010.537748.

Lubbe, A., Jordaan, M., Bragg, C., Paxton, B. R., Schumann, M., Van der Walt, R. & Ngobela, T. (2015). *Labeo seeberi*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed June 2015).

Maake, P. A. (2014). Systematics and phylogeography of Mormyrid fishes in south africa. PhD Thesis, Rhodes University, Grahamstown, South Africa.

Machordom, A. & Doadrio, I. (2001). Evolutionary history and speciation modes in the cyprinid genus Barbus. *Proceedings of the Royal Society* B **1**, 1297–1306.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology* **46**, 523–536.

Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology & Evolution* **10**, 294–299.

Marshall, B. E. & Tweddle, D. (2007). *Labeo lunatus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org>. (Accessed August 2015).

Masters, J. C,. Rayner, R. J., McKay, I .J., Potts, A. D., Nails, D., Ferguson, J. W., Weissenbacher, B. K., Allsopp, M. & Anderson, M. L. (1987). The concept of species: Recognition versus isolation. *South African Journal of Science* **83**, 534–537.

Maud, R. R. & Partridge, T.C. (1987). Regional geomorphic evidence for climatic change in southern Africa since the Mesozoic. *Palaeoecology of Africa* **18**, 337–348.

Mayden, R. L. (1997). A hierarchy of species concepts: The denouement in the saga of the species problem. In *Species: The units of biodiversity* (Claridge, M. A., Dawah, H. A. & Wilson, M. R., eds.). pp. 381–424. Londo: Chapman and Hall.

Mayden, R. L. (1999). Consilience and a Hierarchy of Species Concepts: Advances Toward Closure on the Species Puzzle. *Journal of Nematology* **31(2)**, 95–116.

Mayden, R. L., Chen, W. J., Bart, H. L., Doosey, M. H., Simons, A. M., Tang, K. L., Wood, R. M., Agnew, M. K., Yang, L., Hirt, M. V., Clements, M. D., Saitoh, K., Sado, T., Miya, M. & Nishida, M., (2009). Reconstructing the phylogenetic relationships of the Earth's most diverse clade of freshwater fishes order Cypriniformes (Actinopterygii: Ostariophysi): a case study using multiple nuclear loci and the mitochondrial genome. Molecular Phylogenetics and Evolution **51**, 500–514.

Mayr, E. (1942). *Systematics and the origin of species*. New York: Columbia University Press.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cam bridge, Massachusetts: Belknap Press of Harvard University Press.

Meier, R. & Willmann, R. (2000). The Hennigian species concept. In *Species concepts and phylogenetic theory* (Wheeler, Q. D. & Meier, R., eds.). pp. 30–43. New York: Columbia University Press.

Mia, M. Y., Taggart, J. B., Gilmour, A. E., Gheyas, A. A., Das, T. K., Kohinoor, A. H. M., Rahman, M. A., Sattar, M. A., Hussain, M. G., Mazid, M. A., Penman, D. J. & McAndrew, B. J. (2005). Detection of hybridization between Chinese carp species (Hypophthalmichthys molitrix and Aristichthys nobilis) in hatchery broodstock in Bangladesh, using DNA microsatellite loci. *Aquaculture* **247**, 267–273.

Michener, C. D. (1970). Diverse approaches to systematics. *Evolutionary Biology* **4**, 1–38.

Miller, R. J. (1963). Comparative morphology of three cyprinid fishes: *Notropis cornutus*, *Notropis rubellus*, and the hybrid, *Notropis cornutus* X *Notropis rubellus*. *American Midland Naturalist* **69**, 1–33.

Mishler, B. D. (1985). The morphological, developmental, and phylo genetic basis of species concepts in bryophytes. *Bryologist* **88**, 207–214.

Moelants, T. (2010). *Labeo annectens*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed August 2015).

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49(4)**, 718–726.

Moritz, C., Dowling, T. E. & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics* **18**, 269–292.

Moritz, C. (1994). Defining 'Evolutionary Significant Units' for conservation. *TREE* **9(10)**, 373–375.

Morrison, C. L., Lemarié, D. P., Wood, R. M. & King, T. L. (2006). Phylogeographic analyses suggest multiple lineages of *Crystallaria asprella* (Percidae: Etheostominae). *Conservation Genetics* **7**, 129–147.

Muhlfeld, C. C., McMahon, T. E., Belcer, D. & Kershner, J. L. (2009). Spatial and temporal spawning dynamics of native westslope cutthroat trout, *Oncorhynchus clarkia lewisi*, introduced rainbow trout, *Oncorhynchus mykiss*, and their hybrids. Canadian Journal of Fisheries Aquatic Sciences. **66**, 1153–1168.

Mulder, P. F. S. (1973). Aspects on the ecology of *Labeo capensis* and *Labeo umbratus* in the Vaal River, *Zoologica Africana* **8**, 15–24.

Nei, M. & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Nelson, G. & Platnick, N. I. (1981). *Systematics and biogeography*. New York: Columbia University Press.

Nelson, J. S. (2006). *Fishes of the World*. John Wiley & Sons, Hoboken, NJ, USA.

Nixon, K. C. & Wheeler, Q. D. (1990). An amplification of the phylogenetic species concept. *Cladistics* **6**, 211–223.

Nosil, P. (2009). Ecological speciation. Oxford University Press, New York.

O'Keefe, J. H. & De Moor, F. C. (2006). Changes in the physico-chemistry and benthic invertebrates of the Great Fish River, South Africa, following an interbasin transfer of water. *Regulated Rivers Research and Management* **2(1)**, 39–55.

Pacheco, N. M., Congdon, B. C & Friesen, V. L. (2002). The utility of nuclear introns for investigating hybridization and genetic introgression: a case study involving *Brachyramphus* murrelets.*Conservation Genetics*, **3**, 175–182.

Partridge, T. C. (1993). The evidence for cainozoic aridification in southern Africa. *Quaternary International* **17**, 105–110.

Partridge, T. C. & Maud, R. R. (1987). Geomorphic evolution of southern Africa since the Mesozoic. *South African Journal of Geology* **90**, 179–208.

Partridge, T.C. & Maud, R.R. (2000). Macro-scale geomorphic evolution of southern Africa. *The Cenozoic of southern Africa* (ed. by T.C. Partridge and R.R. Maud), pp. 3–18. Oxford University Press, New York.

Partridge, T. C., Scott, L. & Hamilton, J. E. (1999). Synthetic reconstructions of southern African environments during the Last Glacial Maximum (21–18 kyr) and the Holocene Altithermal (8–6 kyr). *Quaternary International* **57/58**, 207–214.

Paterson, H. E. H. (1985). The recognition concept of species. In *Species and speciation* (Vrba, E. S., ed.). pp. 21–29. Pretoria: Transvaal Museum.

Patwardhan, A., Ray, S & Amit Roy, A. (2014). Molecular Markers in Phylogenetic Studies-A Review. *Phylogenetics & Evolutionary Biology* **2**, 131. doi:10.4172/2329-9002.1000131.

Paxton, B. R., Clark, B. M. & Brown, C. A. (2002). *An assessment of the effects of habitat degradation and exotic fish species invasions on the distribution of three endemic cyprinids: Barbus capensis, Barbus serra and Labeo seeberi in the Olifants and Doring rivers, Western Cape* (PB E000-00-1302). Reseach Gate: https://www.researchgate.net/profile/Cate_Brown/publication/267830188_AN_ASSESSMENT_OF_THE_EFFECTS_OF_HABITAT_DEGRADATION_AND_EXOTIC_FISH_SPECIES_INVASIONS_ON_THE_DISTRIBUTION_OF_THREE_ENDEMIC_CYPRINIDS_BARBUS_CAPENSIS_BARBUS_SERRA_AND_LABEO_SEEBERI_IN_THE_OLIFANTS_AND_DORING_RIVERS_WESTERN_CAPE_FINAL_REPORT_Prepared_for_Department_of_Water_Affairs_and_Forestry_and_Department_of_Agriculture/links/545b2e9b0cf2c16efbbbdb8d.pdf

Pereira, S. L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular Biology* **23**, 745–752.

Pinton, A., Agnèse, J-F., Paugy, D. & Otera, O. (2013). A large-scale phylogeny of *Synodontis*(Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. *Molecular Phylogenetics and Evolution* **66**, 1027–1040.

Pletzen, A. J. J. & Hattingh, J. (1975). Respiratory characteristics of the blood of *Labeo umbratus* (Smith) and *Labeo capensis* (Smith). *Zoologica Africana* **10(1)**, 47–52.

Posada, D. & Crandall, K. A. (2001). Intraspecific gene genealogies: trees grafting into networks. *TRENDS in Ecology & Evolution* **16(1)**, 37–45.

Posada, D. & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**, 793–808.

Potts, W. M., Booth, A. J., Hecht, T. & Andrew, T. G. (2005). Reproductive biology of a riverine cyprinid, *Labeo umbratus* (Teleostei: Cyprinidae), in small South African reservoirs. *African Journal of Aquatic* Science **30(2)**, 147–155.

Prüfer, K., Munch, K., Hellmann, I., Akagi, K., Miller, J. R., Walenz, B., Koren, S., Sutton, G., Kodira, C., Winer, R., Knight, J. R., Mullikin, J. C., Meader, S. J., Ponting, C. P., Lunter, G., Higashino, S., Hobolth, A., Dutheil, J., Karakoc, E., Alkan, C., Sajjadian, S., Catacchio, C. R., Ventura, M., Marques-Bonet, T., Eichler, E. E., André, C., Atencia, R., Mugisha, L., Junhold, J., Patterson, N., Siebauer, M., Good, J. M., Fischer, A., Ptak, S. E., Lachmann, M., Symer, D. E., Mailund, T., Schierup, M. H., Andrés, A. M., Kelso, J. & Paäbo, S. (2012). The bonobo genome compared with the chimpanzee and human genomes. *Nature* **486**, 527–531. doi:10.1038/nature11128.

Pusey, B. J., Kennard, M. & Arthington, A. (2004). Freshwater fishes of North-Eastern Australia. CSIRO Publishing: Australia.

Raffi, S., Steven M. Stanley, S. M. & Marasti, R. (1985). Biogeographic patterns and Plio-Pleistocene extinction of Bivalvia in the Mediterranean and southern North Sea. *Paleobiology* **11(4)**, 368–388.

Rainboth, W. J. (1991). Cyprinid fishes of Southeast Asia. In *Cyprinid Fishes: Systematics, Biology and Exploitation*. (Winfield, I. J. & Nelson, J. S. eds.). Chapman and Hall, London, pp. 156–210.

Rainboth, W. J. (1996). *Fishes of the Cambodian Mekong. FAO species identification field guide for fishery purposes*. FAO, Rome.

Rambaut, A. (2009). Website: http://tree. bio.ed.ac.uk/software/figtree. (Accessed February 2015).

Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. (2014). Tracer v1.6, Available fromhttp://beast.bio.ed.ac.uk/Tracer. (Accessed January 2015).

Ramoejane, M. (2010). The genetic integrity of *Labeo capensis* and *L. umbratus* (Cyprinidae) in South Africa in relation to inter-basin water transfer schemes. MSc Thesis, Rhodes University, Grahamstown, South Africa.

Ramsay, P. J. & Cooper, J. A. G. (2002). Late Quaternary sea-level change in South Africa. *Quaternary Research* **57**, 82–90.

Ray, L. L. (1974). *Geomorphology and Quaternary Geology of the Glaciated Ohio River Valley–A Reconnaissance Study* (No 74-600081). U.S. Department of Interior http://pubs.usgs.gov/pp/0826/report.pdf

Ridley, M. (1989). The cladistic solution to the species problem. *Biology & Philosophy* **4**, 1–16.

Reid, G.M. (1982). The form, function and phylogenetic significance of the vomeropalatine organ in cyprinid fishes. *Journal of Natural History* **16**, 497–510.

Reid, G. M. (1985). *A revision of African Species of Labeo* (VV Cramer). Cramer: Braunschweig.

Reist, J. D. (1985). An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Canadian Journal of Zoology* **63**,1429–1439.

Rieman, B. & Clayton, J. (1997). Wildfire and native fish: Issues of forest health and conservation of sensitive species. *Fisheries* **22**, 6–15.

Rieman, B., Lee, D., Chandler, D. & Myers, D. (1997). Do wildfires threaten the extinction of Salmonids? Responses of Redband trout and Bull trout following

recent fires on the Boise national forest. In *Proceedings on the symposium on fire effects on threaten and endangered species and habitats* (Green, L. ed.). International Association of Wildland Fire. Fairfield, WA.

Roberts, T. (1986). Review: A revision of African species of *Labeo* (Pisces: Cyprinidae) and a re-definition of the genus. G.M. Reid 1985. *Copeia*, 847–848.

Rodriguez, F., Oliver, J. L., Marin, A. & Medina, J. R. (1990). The general stochastic-model of nucleotide substitution. *Journal of Theoretical Biology* **142**, 485–501.

Rogers, J. (1985). Geomorphology, offshore bathymetry and Quaternary lithostratigraphy around the Bot River Estuary. *Transactions of the Royal Society of South Africa* **45**, 211–227.

Rogers, A. R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552–569.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Rosen, D. E. (1979). Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bulletin of the American Museum of Natural History* **162**, 267–376.

Roux, D., De Moor, F., Cambray, J. & Barber-James, H. (2002). Use of landscape-level river signature in conservation planning: South African case study. *Conservation Ecology* **6**, 6.

Rozas, J. & Librado, P. (2009). DnaSP v5, a comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Ruddock, A. (1947). Terraces in the lower part of the Sundays River valley, Cape Province. *Transactions of the Royal Society of South Africa* **31**, 347–370.

Rundle, H. D. & Nosil, P. (2005). Ecological speciation. *Ecology Letters* **8**, 336–352.

Ryder, O. A. (1986). Species conservation and systematics: the Dilemma of Subspecies. *Trends Ecology and Evolution* **1**, 9–10.

Schwartz, J. H. (2005). Molecular systematics and Evolution. In *Encyclopedia of molecular Cell Biology and Molecular Medicine*. (Meyers, R. A. ed). Pp. 515–540. Wiley-VCH Gmbh & co. KGaA: Weinhem.

Scribner, K. T., Page, K. S. & Bartron, M. L. (2001). Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* **10,** 293–323.

Shah, M. S., Ghosh, A. K., Rahi, M. L., Huq, K. A., Rahaman, S. M. B. & Sabbir, W. (2011). Production of Heterotic Hybrid in Rohu (*Labeo rohita*) Through Strain Crossing. *International Journal of Life Sciences* **5(1)**, 32–38. DOI 10.3126/ijls.v5i1.5035.

Shao, Z. & Wang, H. (2003). Interbasin transfer projects and their implications: A China case study. *International Journal of River Basin Management* **1(1)**, 5–14.

Simpson, G. G. (1951). The species concept. *Evolution* **5**, 285–298.

Siesser, W.G. & Dingle, R.V. (1981) Tertiary sea-level movements around southern Africa. *The Journal of Geology* **89**, 83–96.

Skelton, P. H. (1986). Fish of the Orange-Vaal System. In *The ecology of river systems* (Davies, B. R. & Walker, K. F., eds), pp. 143–161. Dordrecht: Dr W. Junk Publishers.

Skelton, P. H., Tweddle, D. & Jackson, P. B. N. (1991). Cyprinids of Africa. In *Cyprinid fishes Systematics, biology and exploitation*, (Winfield, I. J. & Nelson, J. S., eds), pp. 1–33. London: Chapman & Hall.

Skelton, P. (2001). *A complete guide to the freshwater fishes of southern Africa*. Cape Town: Struik Publishers.

Smith, A. (1841). Illustrations of the zoology of South Africa; consisting chiefly of figures and descriptions of the objects of natural history collected during an expedition into the interior of south Africa in 1834-36. Pages not numbered, pls 9–14 inclu. Dated 1841 (*L. umbratus*, *L. capensis*).

Smith, G. R. (1992). Introgression in fishes – significance for paleontology, cladistics, and evolutionary rates. *Systematic Biology.* **41**, 41–57.

Snaddon, C. D., Davies, B. R. & Wishart, M. J. (1999). *A global overview of inter-basin water transfer schemes, with an appraisal of their ecological, socio-economical and socio-political implications, and recommendations for their management*. Water Research Commission, Pretoria, South Africa, Report No. WRC TT120/00. pp. 137–166.

Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical taxonomy: The principles and practice of numerical classification*. San Francisco: W H. Freeman.

Sokal, R. R., and T. J. Crovello. 1970. The biological species concept: A critical evaluation. *American Naturalist* **104**, 127–153.

Sousa, V & Hey, J. (2013) Understanding the origin of species with genome-scale data: modelling gene flow. *Nature Reviews Genetics* **14(6)**, 404–414.

Stanley, S. M. (1986). Anatomy of a Regional Mass Extinction: Plio-Pleistocene Decimation of the Western Atlantic Bivalve Fauna. *PALAIOS* 1(1), 17–36.

Stauffer Jr1, J. R., McKaye, K. R. & Konings, A. F. (2002). Behaviour: an important diagnostic tool for Lake Malawi cichlids. *FISH and FISHERIES* **3**, 213–224.

Stepien, C. A. & Kocher, T. D. (1997). Molecules and Morphology in Studies of Fish Evolution. In *Molecular Systematics of Fishes* (Kocher, T. D. & Stepien, C. A. eds), pp. 1–11. Academic Press: USA.

Stiassny, M. L. J. & Getahun, A. (2007). An overview of labeonin relationships and the phylogenetic placement of the Afro–Asian genus *Garra* Hamilton, 1922 (Teleostei: Cyprinidae), with the description of five new species of *Garra* from Ethiopia, and a key to all African species. *Zoological Journal of Linnean Society* **150**, 41–83.

Stokes, D., Morgan, C., O'Neill, C. & Bancroft, I. (2007). Evaluating the utility of Arabidopsis thaliana as a model for understanding heterosis in hybrid crops. *Euphytica* **156**,157–171. DOI 10.1007/s10681-007-9362-1.

Su, R. F., Yang, J. X. & Chen, Y. R. (2000). A review on the Chinese species of *Crossocheilus*, with description of a new species (Ostariophysi: Cyprinidae). *Raffles Bulletin of Zoology* **48**, 215–221.

Su, R. F., Yang, J. X. & Cui, G. H. (2003). Taxonomic review of the genus *Sinocrossocheilus Wu* (Teleostei: Cyprinidae), with a description of four new species. *Zoological Studies* **42**, 420–430.

Sullivan, W. P., Christie, G. C., Cornelius, F. C., Fodale, M. F., Johnson, D. A., Koonce, J. F., Larson, G. L., McDonald, R. B., Mullett, K. M., Murray, C. K. & Ryan, P. A. (2003). The sea lamprey in Lake Erie: a case history. *Journal of Great Lakes Research*. **29** (Suppl. 1), 615–636.

Suárez-Díaz, E. & Anaya-Munoz, V. H. (2008). History, objectivity, and the construction of molecular phylogeies. *Studies in history and phylosophy of biological and biomedical sciences* **39**, 451–468.

Swartz, E. & Impson, D. (2007). *Labeo capensis* and *Labeo umbratus*. The IUCN Red List of Threatened Species. Version 2015.2. <www.iucnredlist.org >. (Accessed July 2015).

Swartz, E. R., Skelton, P. H. & Bloomer, P. (2007). Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redfins (*Pseudobarbus afer* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *Journal of Biogeography* **34**, 2086–2099.

Swartz, E. R., Skelton, P. H. & Bloomer, P. (2008). Phylogeny and biogeography of the genus *Pseudobarbus* (Cyprinidae): Shedding light on the drainage history of rivers associated with the Cape Floristic Region. *Molecular Phylogenetics and Evolution* **51(1)**, 75–84.

Swofford, D. L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Tamura K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10(3)**, 512–526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729.

Tankard, A. J. (1976). Cenozoic sea-level changes: a discussion. *Annals of the South African Museum* **71**, 1–17.

Tang, Q., Getahun, A. & Lui, H. (2009). Multiple in-to-Africa dispersal of labeonin fishes (Teleostei: Cyprinidae) revealed by molecular phylogenetic analysis. *Hydrobiologia* **632**, 261–271

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* (American Mathematical Society) **17**, 57–86.

Templeton, A. R. (1989). The meaning of species and speciation: A genetic perspective. In *Speciation and its consequences* (Otte, D & Endler, J. A., eds.). pp. 3–27. Sunderland, Massachusetts: Sinauer Associates.

Templeton, A. R, Crandall, K. A. & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633.

Templeton, A. R. (1998). Species and speciation: Geography, population structure, ecology, and gene trees. In *Endless forms: Species and speciation* (Howard, D. J. & Berlocher, S. H., eds.). pp. 32–43. New York: Ox ford University Press.

Thys van den Audenaerde, D. F. E. (1987). Review of Reid GM 1985, a revision of African species of *Labeo* (Pisces, Cyprinidae). *Revue de Zoologie Africaine* **100**, 487–492.

Tobe, S. S., Kitchener, A. C. & Linecre, A. M. T. (2010). Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome *b* and Cytochrome Oxidase Subunit I Mitochondrial Genes. PLoS ONE **5(11)**: e14156. doi:10.1371/journal.pone.0014156.

Tómasson, T., Cambray, J. A. & Jackson, P. B. N. (1984). Reproductive biology of four large riverine fishes (Cyprinidae) in a man-made lake, Orange River, South Africa. *Hydrobiologia* **112**, 179–195.

Truswell, J. F. (1977). *The Geological Evolution of South Africa*. pp. 151–190. Purnell, Cape Town.

Tshibwabwa, S. M. (1997). *Systématique des espéces africaines du genre Labeo (Teleostei, Cyprinidae) dans les régions ichtyogéo-graphiquesde Basse-Guinée et du Congo*. Vol I et II. Faculté des Sciences, Facultés Universitaires Notre-Dame de la Paix de Namur, Belgium. Presses Universitaires de Namur: Namur. p 530.

Tshibwabwa, S. M. & Teugels, G. G. (1995). Contribution to the systematic revision of the African cyprinid fish genus *Labeo*: species from the Lower Zaire river system. *Journal of Natural History* **29**, 1543–1579.

Tshibwabwa, S., Stiassny, M. L. J. & Schelly, R. C. (2006). Description of a new species of *Labeo* (Teleostei: Cyprinidae) from the lower Congo River. *Zootaxa* **1224**, 33–44.

Tweddle, D. & Davies, M. T. T. (1997). Preliminary observations on the fish populations of the recently- 98 African Zoology Vol. 47, No. 1, April 2012 impounded Katse reservoir, Lesotho. *Southern African Journal of Aquatic Sciences* **23**, 7–20.

Unmack, P. J. (2013). Biogeography. In *Ecology of Australian Freshwater Fishes*, (Humphries, P & Walter, K., eds), pp. 26–48. CSIRO publishing. Australia.

Van Rensburg, B. J., Weyl, O. L. F., Davies, S. J., Van Wilgen, N. J., Spear, D., Chimimba, C. T. & Peacock, F. (2011). Invasive vertebrates of South Africa. In *Biological invasions: Economical and environmental costs of alien plants, animal*

and microbe species. (David Pimentel ed). pp. 325–378. CRC Press, Taylor & Francis group: London.

Van Robbhroeck, T. P. C. (1979). Inter-basin water tranfers in South Africa. Department of Water Affairs. (Accessed October 2015).

Van Steenberge, M., Gajdzik, L., Chilala, A., Snoeks, J. & Vreven, E. (2014). *Labeo rosae* (Cypriniformes: Cyprinidae) in the Congo Basin: a relict distribution or a historical introduction? *Journal of Fish Biology*. doi:10.1111/jfb.12491. Available online at wileyonlinelibrary.com. (Accessed June 2015).

Van Valen, L. (1976). Ecological species, multispecies, and oaks. *Taxon* **25**, 233–239.

Van Vuuren, N. G., Mulder, P. F., Ferreira, J. T. & Van Der Bank, F. H. (1989). The identification of hybrids of *Barbus aeneus* X *B. kimberleyensis* and *Labeo capensis* X *L. umbratus* in Hardap Dam, SWA/Namibia. *Madoqua* **16**, 27–34.

Van Vuuren, N. G., Mulder, P. F., Ferreira, J. T. & Van Der Bank, F. H. (1990). A morphometric and electrophoretic analysis of *Labeo capensis* and *Labeo umbratus* from two localities in southern Africa. *Water South Africa* **16**, 135–146.

Van Zyl, B. J. (1991). The lower Orange River. *Madoqua* **17**, 155–157.

Vishwanath, W. & Kosygin, L. (2000). Garra elongata, a new species of the subfamily Garrinae from Manipur, India (Cyprinidae, Cypriniformes). *Journal of Bombay Natural History Society* **97**, 408–414.

Waples, R. S. (1991). Pacific Salmon, *Oncorhynchus* spp., and the Definition of "Species" Under the Endangered Species Act. Marine Fisheries Review **53**, 11–22.

Ward, R. D. (2000). Genetics in fisheries management. *Hydrobiologia* **420**: 191–201.

Waters, J. M. & Wallis, G. P. (2001). Cladogenesis and loss of the marine life-history phase in freshwater Galaxiid fishes (Osmeriformes: Galaxiidae). *Evolution* **55(3)**, 587–597.

Wellington, J. H. (1955). *Southern Africa: A geographical study*, V 1. Cambridge University Press: London.

Weyl, O. L. F., Booth, A. J., Winker, H., McCafferty, J., Wartenberg, R., Peel, R. & Traas, G. R. L. (2009). Assessment of the fishery potential for alien species on Darlington Dam. Sundays River Freshwater Fish Research Project. Report No. 5.

Wheeler, Q. D. & Meier, R. (2000). Species concepts and phylogenetic theory: A debate. New York: Columbia University Press.

Wiley, E. 0. (1978). The evolutionary species concept reconsidered. *Systematic Zoology* **27**, 17–26.

Wilson, C. C. & Bernatchez, L. (1998). The ghost of hybrids past: fixation of artic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology* **7**, 127–132.

Winker, H., Weyl, O. L. F., Booth, A. J. & Ellender, B. R. (2012). Life history strategy and population characteristics of an unexploited riverine cyprinid, *Labeo capensis*, in the largest impoundment in the Orange River Basin. *African Zoology* **47(1)**: 85–99.

Wirtz, P. (1999). Mother species–father species: unidirectional hybridization in animals with female choice. Animal Behaviour **58**, 1–12.

Wood, R. M. & Raley, M. E. (2000). Cytochrome *b* sequence variation in the Crystal Darter *Crystallaria asprella* (Actinopterygii: Percidae). *Copeia* **1**, 20–26.

Woodford, D. J., Hui, C., Richardson, D. M. & Weyl, O. L. F. (2013). Propagule pressure drives establishment of introduced freshwater fish: quantitative evidence from an irrigation network. *Ecological Applications* **23(8)**, 1926–1937.

Wright, S. (1940). The statistical consequences of Mendelian heredity in relation to speciation. In *The new systematics* (Huxley, J., ed.). pp. 161–183. London: Oxford University Press.

Yang, L., Arunachalamb, M., Sadoc, T., Levind, A. B., Golubtsove, A. S., Freyhoff, J., Frielg, J. P., Chenh, W., Hirti, M. V., Manickamb, R., Agnewa, M. K., Simonsj, A. M., Saitohk, K., Miyac, M., Maydena, R. L. & He, S. (2012). Molecular phylogeny of the cyprinid tribe Labeonini (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **65(2)**, 362–379.

Yang, L. & Mayden, R. L. (2010). Phylogenetic relationships, subdivision, and biogeography of the Cyprinid tribe Labeonini (sensu Rainboth, 1991) (Teleostei: Cypriniformes), with comments on the implications of lips and associated structures in the Labeonin classification. *Molecular Phylogenetics and Evolution* **54**, 254–265.

Yang, L., Sado, T., Hirt, M. V., Pasco-Viel, E., Arunachalam, M., Li, J., Wang, X., Freyhof, J., Saitoh, K., Simons, A. M., Miya, M., He, S. & Maydena, R. L. (2015). Phylogeny and polyploidy: Resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **85**, 97–116.

Zardoya, R. & Doadrio, I. (1999). Molecualr evidence on the evolution and biogeographical patterns of European cyprinids. *Journal of molecular evolution* **49**, 227–237.

Zhang, E. & Chen, Y. Y. (2002). *Garra tengchongensis*, a new cyprinid species from the upper Irrawaddy river basin in Yunnan, China (Pisces: Teleostei). *Raffles Bulletin of Zoology* **50**, 459–464.

Zhang, E. & Chen, Y. Y. (2004). Qianlabeo striatus, a new genus and species of Labeoninae from Guizhou Province, China (Teleostei: Cyprinidae). *Hydrobiologia* **527**, 25–33.

Zhang, E. & Fang, F. (2005). Linichthys: a new genus of Chinese cyprinid fishes (Teleostei: Cypriniformes). *Copeia* **2005**, 61–67.

Zhang, E. & Kottelat, M (2006). *Akrokolioplax*, a new genus of Southeast Asian labeonine fishes (Teleostei: Cyprinidae). *Zootaxa* **1225**, 21–30.

Zhang, E. & Zhou, W. (2012). *Sinigarra napoense*, a new genus and species of labeonin fishes (Teleostei: Cyprinidae) from Guangxi Province, South China. *Zootaxa* **3586**, 17–25.

Zhang, E., Kullander, S. O. & Chen, Y. Y. (2006). Fixation of the type species of the genus *Sinilabeo* and description of a new species from the Upper Yangtze River Basin, China (Pisces: Cyprinidae). *Copeia* **2006**, 96–102.

Zheng, L., Yang, J. & Chen, X. (2012). Phylogeny of the Labeoninae (Teleostei, Cypriniformes) based on nuclear DNA sequences and implications on character evolution and biogeography. *Current Zoology* **58(6)**, 837–850.

Zhou, W., Pan, X. F. & Kottelat, M. (2005). Species of *Garra* and *Discogobio* (Teleostei: Cyprinidae) in Yuanjiang (upper Red River) drainage of Yunnan Province, China, with description of a new species. *Zoological Studies* **44**, 445–453.