The genetic integrity of *Labeo capensis* and *L. umbratus* (Cyprinidae) in South Africa in relation to inter-basin water transfer schemes

Thesis submitted in fulfilment of the requirements for the degree of

Master of Science

at

RHODES UNIVERSITY

by

Mpho Ramoejane

February 2011

The genetic integrity of *Labeo capensis* and *L. umbratus* (Cyprinidae) in South Africa in relation to inter-basin water transfer schemes

I am aware of the University's policy on plagiarism. All of the work is my own. I have not included ideas, phrases, passages or illustrations from another person's work without acknowledging their authorship.

Name: Mpho Ramoejane

Student Number: g08r6192

Signed:

Date: /02/2011

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank God and my ancestors for always giving me wisdom, strength and protection. If it was not for you, this would not have been possible.

I thank my Mom Dibuseng Ramoejane for the support, encouragement, guidance and love. Even though she was far away, she was still doing what a loving mother would do for her child. Thank you Mom.

My supervisor, Dr Ernst Swartz, gave me support and guidance. He made Grahamstown feel like a second home, because I knew who to turn to when I needed help with anything (academic or non-academic). Most of all, he had faith in me.

I also thank my co-supervisor, Dr Olaf Weyl, for his support and encouragement. He always kept me on my toes and I had to do things a lot faster than I am used to, which helped me to finish my thesis.

My fellow students at the South African Institute for Aquatic Biodiversity’s molecular laboratory (Thabo Maake, Tshoanelo Moloi, Albert Chakona, Moqebelo Morallana and Poogendry Reddy who was also the laboratory assistant) gave me moral support and a helping hand in the lab, field and with some of the genetic analyses. I thank my fellow students Bruce Ellender and Henning Winker at the Department of Ichthyology and Fisheries Science at Rhodes University for showing me how to use different field gear to catch fish and for sampling some of my specimens for me. Wilbert Kadye helped me with my morphological analysis. Sherwyn Mack helped me create maps.

I thank Roger Bills for the opportunity to survey Hardap Dam in Namibia and the lower Orange River with him and for his assistance in finding useful morphological characters for my study. Dean Impson and Martine Jordaan (CapeNature) helped with my survey of the Sak River and Gordon O’Brain and his students from the University of Johannesburg helped me to source specimens from the Vaal River.

The Eastern Cape Department of Economic Development and Environmental Affairs, Northern Cape Department of Tourism, Environment and Conservation and CapeNature provided permits for this research.

I thank the South African Institute for Aquatic Biodiversity (SAIAB) for funding, facilities and space. I also thank the National Research Foundation of South Africa for my bursary and for funding my research.

ABSTRACT

The Orange-Fish and Cookhouse tunnels that are part of a major inter-basin water transfer scheme (IBT) act as a pathway for several fish species from the Orange River system to enter the Great Fish and Sundays River systems in South Africa. These include *Labeo capensis* and *L. umbratus*. *Labeo capensis* was restricted to the Orange River system before the inter-basin water transfer scheme. *Labeo umbratus* occurred naturally in the Orange River and in southern flowing river systems. Previous studies showed that the two species hybridise in Hardap Dam, located in a tributary of the Orange River system in Namibia. There are also unconfirmed reports of hybrids from Darlington Dam on the Sundays River system.

The aim of the thesis was to confirm hybridisation in Hardap Dam, assess whether hybridisation between *L. capensis* and *L. umbratus* has occurred in Darlington Dam and to gain a better understanding of the diversity of these two species. Morphology (morphometrics and meristics), a nuclear S7 intron and the mitochondrial cytochrome *b* gene were used to assess for hybridisation. A total of 275 specimens were analysed from across the geographical range of the two species.

The two species could be distinguished using morphometrics (dorsal fin base, inter-orbital width and operculum to eye distance) and meristics (lateral line, origin of the dorsal fin to lateral line, origin of the pelvic fin to lateral line and caudal peduncle scale counts) characters. Hybrids from Hardap and Darlington dams were placed between the two species clusters. *Labeo umbratus* from the Orange River and southern flowing rivers formed a single cluster. The two species could also be distinguished from each other with six nuclear DNA mutations and hybrids were heterozygous at such sites in both dams. *Labeo umbratus* populations from the Orange River and southern flowing rivers (Gouritz, Gamtoos, Sundays, Bushmans, Great Fish and Nahoon) formed a single lineage. Analysis of mitochondrial DNA, however, revealed that *L. umbratus* populations from the Orange River and southern flowing rivers were two lineages that differ from each other by 5 mutations. *Labeo capensis* could be differentiated from both these lineages. Being maternally inherited, mitochondrial DNA did not reveal hybridisation, but ten specimens with *L. capensis* haplotypes were found in the Darlington Dam. In Hardap Dam, however, it appears that only *L. capensis* mitochondrial DNA haplotypes persist, despite morphological and nuclear DNA analysis suggesting that both morphs and hybrids of the two species occur.

The genetic integrity of these *Labeo* species has therefore been compromised in at least Hardap and Darlington dams. The Great Fish and Sundays populations are considered to be under threat of complete introgression. The Kat River and Slagboom Dam populations that were isolated before the IBTs have to remain isolated to protect the genetic integrity of the southern lineage of *L. umbratus* in these two systems.

Keywords: Freshwater fish; conservation; Hardap Dam; Darlington Dam; inter-basin water transfer schemes; hybridisation; introgression; morphology; nuclear DNA; mitochondrial DNA

Chapter 1: General Introduction

1.1 Importance of the study

The genetic integrity of many freshwater fishes has been affected by hybridisation (Bolnick, 2009; Scribner *et al.*, 2001). Hybridisation is the process whereby genetically distinct species breed and produce intermediate forms (Rhymer and Simberloff, 1996). It seems to be especially prevalent in freshwater fishes (Billington, 2003) and according to Winfield and Nelson (1991) and Scribner *et al*. (2001), the speciose Cyprinidae is the fish family with the most records of hybridisation (Gante *et al.*, 2004). Hybridisation mostly occurs between closely related species (Allendorf *et al.*, 2001). Hybridisation can happen in a way that individuals of hybridising population interbreed with each other and this leads to hybrids backcrossing with one or both parental species, often referred to as introgression (Rhymer and Simberloff, 1996).

Hybridisation can be caused naturally or by anthropogenic events (Scribner *et al.*, 2001). Natural hybridisation happens due to range overlap or expansion of species distribution into other habitats. For example, this would have happened naturally in fish species after Pleistocene glaciations (Scribner *et al.*, 2001). As landscapes deglaciated, connectivity between aquatic habitats would have been restored and species that were isolated in glacial refugia would have mixed, leading to hybridisation (Hewitt, 1996; Hubbs, 1955). More recently however Scribner *et al*. (2001) estimated that nearly 50% of hybridisation events are being driven by anthropogenic impacts, such as aquaculture activities, species introductions or translocations and loss or alteration of habitat. Dam walls, for example, can limit spawning habitats by preventing migration of species (Gaigher and Bloemhof, 1975) resulting in forced sharing of spawning habitats and resultant inbreeding. Introduction of alien species is the main cause of hybridisation and can cause extinction of indigenous species in the recipient river system (Gozlan *et al.*, 2010).

When Allendorf *et al.* (2001) set conservation guidelines concerning hybrids, they recognized that hybridisation can happen naturally or be influenced by anthropogenic events. They recognised three final stages of hybridisation (which they called "hybridisation types" :

1) Hybridisation without introgression, which happens because F1 hybrids are sterile.
2) Widespread introgression, which is hybridisation that results in a hybrid swarm where some of the parental species still exist somewhere in the river tributaries.
3) Complete admixture, which is hybridisation with a hybrid swarm where few, if any pure populations remain and there does not appear to be any selection against the hybrids.

1.2 Impact of inter-basin water transfer schemes

The distribution of fish and other freshwater organisms have been affected by inter-basin water transfer schemes (IBTs). An IBT allows the translocation of water from a donor river system across a certain distance, usually via a tunnel to a recipient river system (Snaddon, 1999). In South Africa, IBTs were mainly built for irrigation purposes in semi-arid areas and to generate hydro-electricity (Davies, 1993). When the IBTs were planned, the focus was on the economic value of them, rather than the ecological and social impacts (Snaddon, 1999). Now they affect the ecosystem negatively by introducing aquatic species, notably fishes (resulting in genetic contamination, predation and competition), invertebrates and algae which become pests themselves or vectors of parasites in the recipient system. IBTs can also contribute to soil erosion and change the chemistry and sediment load of the water in the recipient river system (Davies *et al.*, 1992). According to Slabber (2007), there are 26 major IBTs in southern Africa.

The Orange-Fish (completed in 1975) and Cookhouse (completed in 1978) tunnels, act as a pathway for several fish species from the Orange River system to reach the Great Fish and Sundays River systems (Cambray and Jubb, 1977). These species include the Orange River mudfish *Labeo capensis* (Smith, 1846) and moggel *Labeo umbratus* (Smith, 1846). *Labeo capensis* only occur naturally in the Orange River system (Skelton, 2001). *Labeo umbratus* is not confined to the Orange River system, but has an indigenous distribution that includes southern flowing rivers from the Gouritz in the west to the Nahoon in the east (Gaigher and Bloemhof, 1975). Other species that were translocated through the IBTs are *Labeobarbus aeneus*, *Austroglanis sclateri* and *Clarias gariepinus*. Some species were also introduced deliberately to improve angling (e.g. *Clarias gariepinus*, *Cyprinus carpio*, *Micropterus salmoides*, *Lepomis macrochirus*, *Tilapia sparrmanii* and salmonids).

1.3 Introduction of the two study species

Labeo is a large complex genus of carp-like (cyprinid) fishes with sucker-like lips (Reid, 1985). There are six *Labeo* groups (*L. gregorii*, *L. macrostoma*, *L. umbratus*, *L. niloticus*, *L. coubie and L. forskalii*) recognised in Africa (Reid, 1985). *Labeo capensis* and *L. umbratus* belong to the *L. umbratus* group with *L. seeberi* and *L. rubromaculatus*. *Labeo* species are called mudfishes because of their well developed mouth suited to scrub algae, diatoms and detritus from sediments (Skelton, 2001). The present study concentrates on the two species that occur in sympatry in the Orange River system –i.e, *Labeo capensis* (Fig 1.1) and *Labeo umbratus* (Fig 1.2)

Fig. 1.1 *Labeo capensis* from Darlington Dam (photograph by O. L. F. Weyl, South African Institute for Aquatic Biodiversity SAIAB).

Fig. 1.2 *Labeo umbratus* from Darlington Dam (photograph by O. L. F. Weyl, SAIAB).

The two species occupy different ecological niches (Cambray and Jubb, 1977) and this may be due to their different mouth forms. *Labeo capensis* has a terminal mouth while *L. umbratus* has a subterminal one. *Labeo capensis* prefers habitats with fast flowing water while *L. umbratus* prefers slow flowing water (tributaries and dams) (Mulder, 1973). They both feeds on algae and detritus according to Skelton (1986) and this has been confirmed with stable isotopes (Winker, unpublished PhD thesis). Both species are highly fecund with eggs that hatch within 24 hours (Skelton, 1986). *Labeo capensis* does not undertake extensive spawning migrations, but spawn in inundated grass on the sides of river banks while migrating upstream in the main river (lateral migration). *Labeo umbratus* migrate upstream in the channels or large tributaries to flooded areas in the Orange River system (longitudinal migration) (Mulder, 1973). Cambray (1991) reported lateral and longitudinal migration in *L. umbratus* from the Gamtoos River system. There, they spawn on flooded gravel and vegetation where food and shelter is abundant for the young. Both species may spawn at the same time under similar environmental conditions (when the rivers floods in summer) (Tomasson *et al.*, 1984).

1.4 Problem statement

The introduction of *L. capensis* to the Great Fish and Sundays river systems from the Orange River may lead to hybridisation, possibly because of modified habitats or because the natural *L. umbratus* population have not been in contact with *L. capensis* before the inter-basin water transfer occurred (Fig 1.3). In dams, the two species may utilise the same areas for spawning. Introgression is possible between *L. umbratus* populations from the Orange, Great Fish and Sundays River systems. The introduced *L. umbratus* and *L. capensis* were initially restricted to the upper reaches of the Great Fish River system because of the Grassridge dam (Laurenson and Hocutt, 1985). *Labeo capensis* was first discovered in the Great Fish River in 1975 and the population was still relatively small by 1983 (Laurenson *et al.*, 1989). According to Laurenson *et al.* (1989), this showed that it had not adapted well in its new environment and individuals apparently do not reach maturity. In 2007, however, this species has been recorded from the middle reaches of the Great Fish River system and in the Sundays river systems. Further, it may be hybridising with *L. umbratus* in Darlington Dam (O. Weyl, personal communication).

Labeo capensis and *L. umbratus* are reported to hybridise in Hardap Dam (Namibia) (Gaigher and Bloemhof, 1975; Van Vuuren *et al.*, 1989 and 1990). These studies used morphology and allozymes to differentiate between the two species and hybrids. It is not clear how they identified pure versus hybrid individuals *a priori*, since they did not sample localities having only pure populations. Gaigher and Bloemhof (1975) also mention that hybrids between the two species were not confined to Hardap dam, because of records from the Caledon River (citing an Orange Free State Nature Conservation report from 1972/73) and from the Gariep dam (citing a personal communication with Dr. Kas Hamman). However, these claims could not be verified by them because no reference was made to voucher specimens. Van Vuuren *et al.* (1989) concluded that hybrids could not be identified based on morphology alone.

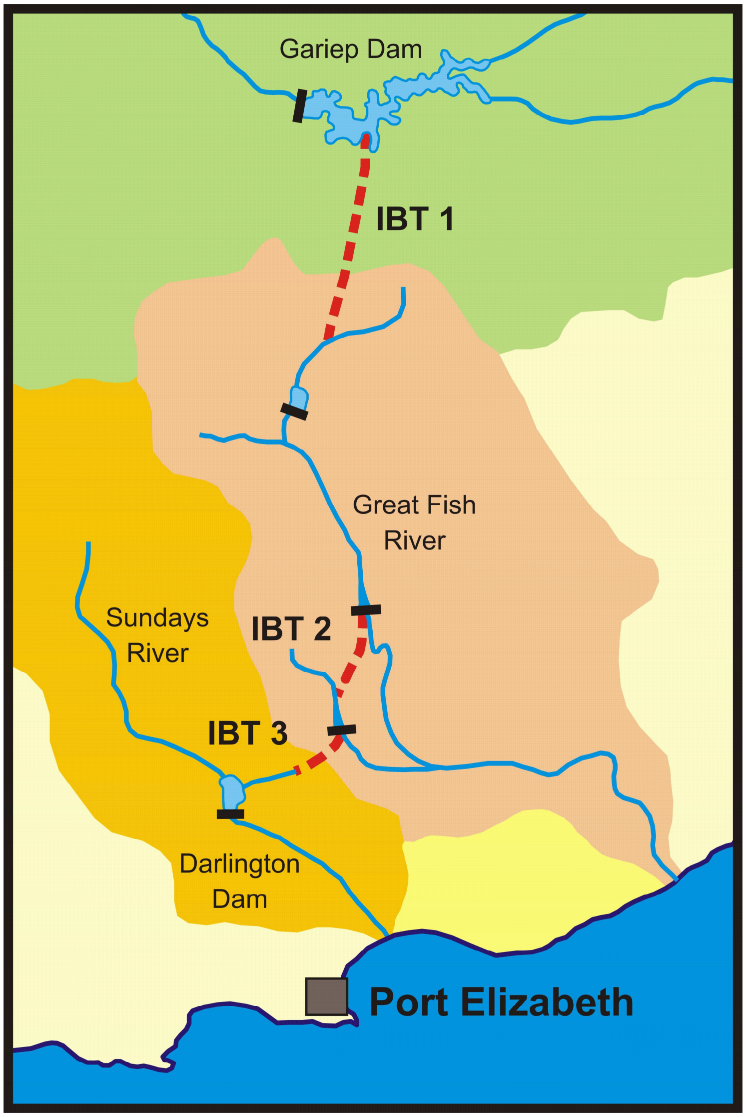

Fig. 1.3 The map showing inter-basin water transfer schemes (IBT's) and the river systems involved. IBT 1 = Orange Fish tunnel, IBT 2 and 3 = Cookhouse tunnels.

There are further problems with the analyses undertaken in these two studies. Apart from not being clear about which criteria their species identifications were based on, they only analysed samples from two localities. One of these localities was Hardap Dam itself. They therefore did not study enough reference populations to represent the overall variation of these two species across the Orange River system and therefore could not characterise pure populations accurately. Furthermore, Gaigher and Bloemhof, (1975) and Van Vuuren *et al.* (1989) lumped data and analysed populations. They did not analyse the specimens separately, so there is no way of knowing whether their original classification was correct.

Concerns have being raised about the genetic integrity and introgression of potentially unique genetic lineages of *L. capensis* and *L. umbratus*, due to the Orange-Fish and Cookhouse tunnels (Cambray and Jubb, 1977; Laurenson and Hocutt, 1985; Laurenson *et al*., 1989), but there is currently no clear evidence to confirm that hybridisation or introgression has indeed occurred apart for unconfirmed reports for Darlington Dam (O. Weyl, personal communication, SAIAB). Given that Orange River *L. capensis* and *L. umbratus* can hybridise in Hardap Dam (Van Vuuren *et al.*, 1989 and 1990; Gaigher and Bloemhof, 1975), one can expect that they could interbreed in the Great Fish and Sundays River systems as well (Fig 1.3).

1.5 Aims of the study and hypothesis investigated

The aim of the study was to differentiate between *L. capensis* and *L. umbratus* species, using morphology (morphometrics and meristics) and genetics (mitochondrial cytochrome *b* and nuclear S7 introns), to identify hybrids, and thus advise conservation authorities on how to best manage the genetic diversity of these two species in relation to existing inter-basin transfer schemes.

To achieve this, the following hypotheses were tested:

1. *Labeo capensis* and *L. umbratus* can be distinguished from each other morphologically and genetically.
2. *Labeo umbratus* from different river systems have been historically isolated.
3. *Labeo capensis* and *L. umbratus* do indeed hybridise in Hardap and Darlington dams.
4. Morphometrics, meristics, nuclear DNA and mitochondrial DNA are all useful procedures for identifying hybrids successfully.

The thesis is presented in six chapters. The general introduction (Chapter 1) is followed by a general chapter on methodology (Chapter 2). The two *Labeo* species are then assessed for variation in morphology (Chapter 3), nuclear DNA (Chapter 4) and mitochondrial DNA (Chapter 5). Finally, the morphological and genetic approaches are all combined to do an overall assessment of the genetic integrity of *L. capensis* and *L. umbratus* populations (Chapter 6).

Chapter 2: General materials and methods

2.1 Research design and sampling

A total of 371 fish specimens were collected from 14 localities in seven river systems across the geographic distributional ranges of *L. capensis* and *L. umbratus* (Table 2.1) in southern Africa. The largest number of sites was in the Orange River system due to its large size and because of the known hybridisation in the Hardap Dam (Fig 2.1).

Localities in the Orange River system

A ten day sampling survey was done in January 2008 at Gariep Dam (previously the Hendrik Verwoerd Dam) in the Orange River system. This dam has a surface area of 352 km^2 (SANCOLD, 2009) and is mostly populated by *Labeobarbus aeneus* and *Labeo capensis* (Ellender 2009, MSc thesis; Winker, unpublished PhD thesis). *Labeo umbratus* was difficult to collect as they were not abundant. Twenty-nine *L. capensis* of a range of sizes (6-35 cm) and twenty-two *L. umbratus* (9-40 cm) which were mostly juveniles, were collected from different areas in the dam to increase the chances of finding hybrids if any existed and to sample as much diversity as possible**.** An additional field trip was done at Gariep Dam in May 2008 for five days to increase the number of *L. umbratus* samples adding twelve specimens. The samples collected on these surveys were used for the preliminary study (see Chapters 3 and 5).

Hardap Dam (Namibia), Onseepkans (lower Orange) and Kanoneiland (middle Orange) were visited from 8 to 11 October 2008. Hardap Dam was built in 1962 and has a surface area of 25 km^2 (SANCOLD, 2009). Suspected pure *L. capensis*, pure *L. umbratus* and hybrids were collected from this dam (N = 81). Only *L. capensis* were found and collected from the middle (N = 11) and lower (N = 14) reaches of the Orange River system. Ten *L. umbratus* specimens were collected from the Brak River, a tributary of the middle Orange River.

No *L. capensis* and *L. umbratus* were collected from the Barrage (near locality 6 Fig 2.1) in the Vaal River, despite two days of gill netting there. However, *L. capensis* specimens (N = 28) were collected in the riffles below the Barrage as this places are close to the type locality of the neotypes for the two species. *Labeo umbratus* specimens (N=15) were collected from the Vaal Dam, where they were more abundant than *L. capensis*. The Vaal Dam was built in 1938 and has a surface area of approximately 322.755 km^2 (SANCOLD, 2009).

Great Fish River system localities

The Kat River Dam was surveyed in May 2008 and specimens of suspected pure *L. umbratus* (N = 27) were collected. This has a surface area of 2 km^2 (SANCOLD, 2009). The dam is on the Kat River, a tributary of the Great Fish River. Specimens were collected from the dam, because it was built before IBTs in 1969, which raises the possibility that the fish represents the original natural stock from the Great Fish River system. Another dam visited in this river system was Lake Arthur (1924) with the surface area of 8.9 km^2 (SANCOLD, 2009) which is in the Tarka River (tributary of Great Fish River). However, no *L. umbratus* specimens were found, possibly because the dam had dried completely in the past (according to local farmers).

Sundays River system localities

Three day sampling surveys were done on a monthly basis to Darlington Dam (built 1922), from April to October 2008. It has a surface area of 34.5 km^2 (SANCOLD, 2009) and 66 specimens were collected. Very few specimens of suspected pure *L. capensis* were found in Darlington Dam, despite considerable effort. Most of them were from one area in the dam (some were collected by Dr Olaf Weyl). Other dams visited in this river system were Slagboom (1955) and Van Ryneveldspas (1925). They were also surveyed, because they were built before the IBTs, which meant they could contain pure populations of indigenous *L. umbratus*. Van Ryneveldspas Dam is located near Graaff-Reinet and has a surface area of 10.3 km^2 (SANCOLD, 2009). No specimens of *L. umbratus* were found in Van Ryneveldspas Dam, probably because it had dried out completely in the past similar to Lake Arthur. However, suspected pure *L. umbratus* (N = 17) were collected from the Slagboom Dam that may be representative of the original stock from the Sundays River system. The dam is on the Wit River, tributary of the Sundays River below Darlington Dam, and has a surface area of 34.5 km^2 (SANCOLD, 2009).

Localities in other southern flowing river systems

The Bushmans River system was sampled at Amakhala Game reserve on the 4th November 2008 and 17 specimens were collected. Specimens (N = 8) were collected by Olaf Weyl above and below the Nahoon Dam in the Nahoon River system. Two localities were sampled in the Gamtoos River system at the farm Mont Pellier and the Perdegat pool near Steytlerville in December 2008 and 19 specimens were collected. Two localities were also sampled in the Gouritz River system at Stompdrift Dam and "Die Poort" in December 2008 and 21 specimens were collected.

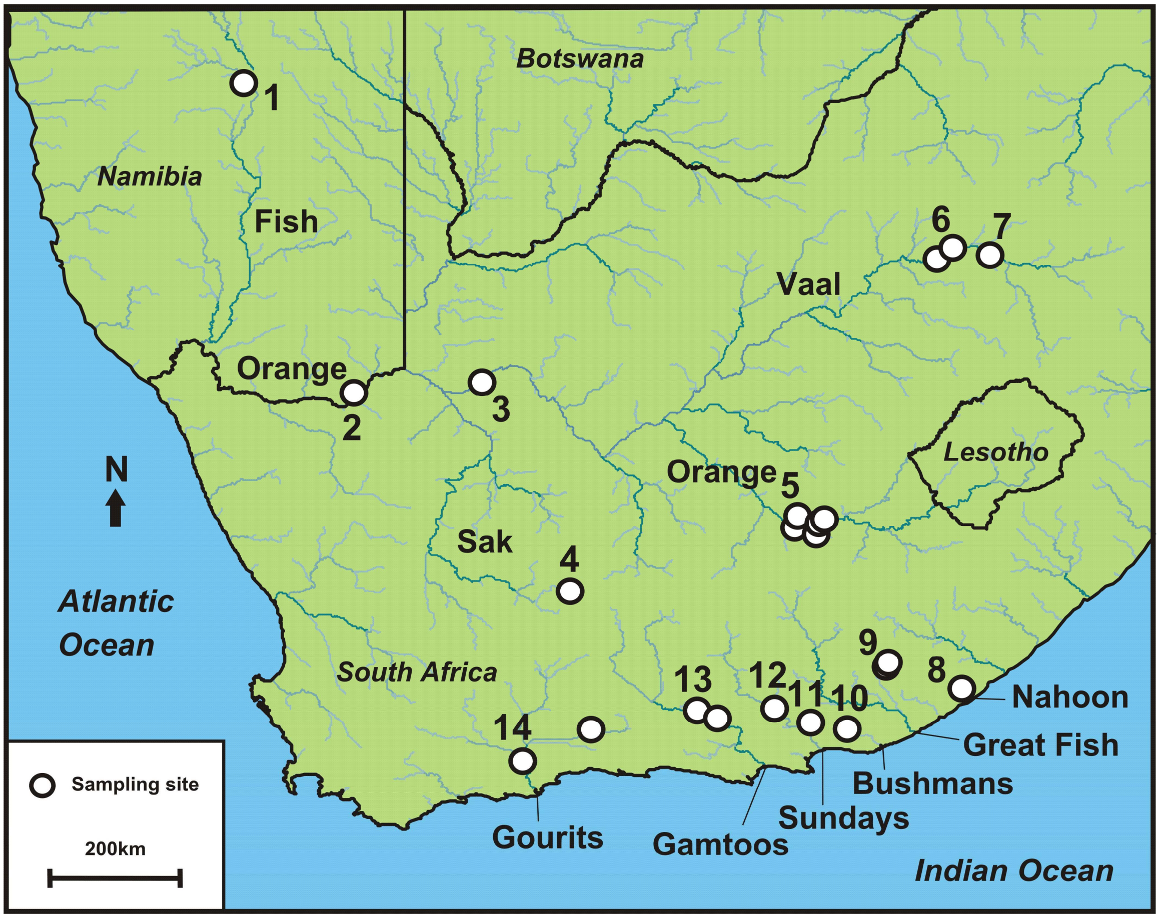

Fig 2.1 A map of southern Africa showing the areas where samples were collected in dams (1-Hardap, 5-Gariep, 7-Vaal, 9-Kat River, 11-Slagboom and 12-Darlington) and rivers (2-Onseepkans on the Orange, 3-Kanoneiland on the Orange, 4-Brak, 6-Vaal, 8-Nahoon, 10-Bushmans, 13-Gamtoos and 14-Gouritz)

Table 2.1 A list of locality information showing the number of fish specimens (of both study species) collected at each locality.

Locality	River system	Number of specimens collected and analysed	Co-ordinates	
			Latitude S/N	**Longitude W/E**
Hardap Dam	Orange	81	24° 28' 11.3"	17 °47' 51.9"
Onseepkans	Orange	14	28° 44' 14.5"	19 °18' 14.4"
Kanoneiland	Orange	11	28° 38' 05.7"	21 °05' 20.3"
Brak	Orange	10	31° 32' 26.4"	22 °20' 35.0"
Gariep Dam	Orange	10	30° 38' 38.2"	25 °33' 50.9"
		8	30° 38' 41.6"	25 °46' 49.2"
		11	30° 36' 65.9"	25 °47' 35.5"
		1	30° 35' 37.8"	25 °31' 49.8"
		6	30° 41' 92.9"	25 °44' 93.5"
		11	30° 38' 10.6"	25 °32' 36.1"
		2	30° 38' 42.5"	25 °31' 24.5"
		2	30° 43' 12.9"	25 °43' 02.3"
		9	30° 42' 84.5"	25 °43' 47.3"
		3	30° 39' 80.4"	25 °44' 08.2"
Vaal River	Orange- Vaal	28	26° 45' 57.6"	27 °40' 56.9"
Vaal Dam	Orange- Vaal	15	26° 51' 58.9"	28 °10' 14.3"
Kat River Dam	Great Fish	27	32° 33' 46.5"	26 °46' 43.0"
Darlington Dam	Sundays	32	33° 10' 82.2"	25 °07' 93.0"
		13	33° 11' 00.7"	25 °10' 11.0"
		21	33° 11' 32.3"	25 °09' 75.2"
Slagboom Dam	Sundays	17	33° 22' 31.1"	25 °40' 45.4"
Amakhala game reserve	Bushmans	17	33° 31' 02.5"	26 °07' 29.2"
East London	Nahoon	8	32° 54' 18.4"	27 °48' 32.2"
Gamtoos	Gamtoos	12	33° 18' 41.8	24 °20' 50.0"
		7	33° 13' 38.5	24 °09' 15.0"
Gouritz	Gouritz	15	33° 30' 42.3"	22 °36' 14.2"
		6	33° 58' 34.8"	21 °39' 19.0"

2.2 Survey gear

Passive gear

Most specimens were collected using gill nets (Fig 2.2) (30m long and 1.5m deep with mesh panels of 44, 60, 75, 100 and 144 mm). Gill nets were deployed using a boat in different bays of dams and along steep banks in the case of rivers. They were placed at a depth of about 3m. The depth was acquired using a fish finder or by testing the depth with a weighted rope. The gill nets were deployed in the late afternoon and were left overnight. They were then retrieved early in the morning. Fish were removed carefully from the gill nets to try and avoid damage to the specimens before preservation. Other gears used were fyke nets (Fig 2.3), which were used with otter guards when left unattended or without otter guards when they were under constant observation.

Fig 2.2 Gill nets as shown in the picture (Gariep Dam, Eastern Cape Province, South Africa) were used for sampling in deeper habitats (photograph by E. R. Swartz, SAIAB).

Fig 2.3 Fyke nets as shown in the picture (Lake Arthur in the Sundays River system, Eastern Cape Province, South Africa) were used for sampling in relatively shallow habitats (photograph by E. R. Swartz, SAIAB).

Active gear

Active gears like throw-, D- (Fig 2.4 A) and seine- (Fig 2.4 B) nets were used to actively collect specimens. Electric fishing was used in smaller water bodies (shallow runs and riffles of smaller rivers).

2.3 Voucher and tissue sample preservation

Prior to preservation, a piece of fresh muscle tissue or a fin clip was cut on the right side of each individual and placed in a 1.5ml tube containing 99 % ethanol. These tubes were later transferred to a -80Cº freezer for long-term storage after the ethanol was replaced.

Voucher specimens were tagged with labels that correspond with the genetic samples and were then fixed in 10% formalin. The body cavity was injected with 10% formalin to improve preservation of the whole specimen. They were later transferred through a gradient of concentrations (to prevent rapid dehydration minimise body shape changes) to an eventual 70% ethanol for long term preservation. Specimens are kept at SAIAB waiting to be accessioned.

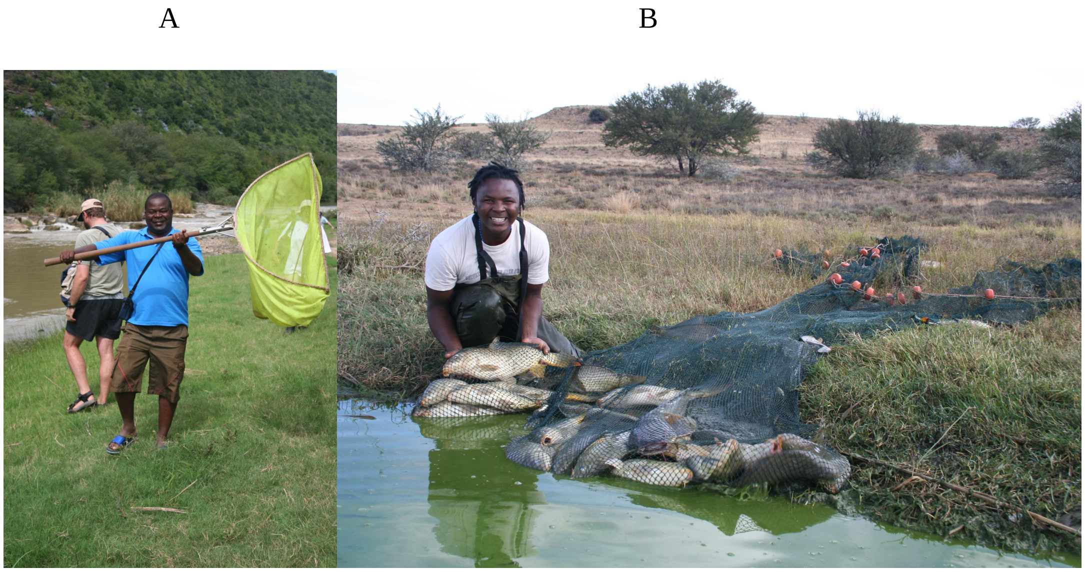

Fig 2.4 Other netting methods included the use of a D-net as seen in photograph A (Great Fish River, Eastern Cape Province, South Africa) and a seine net as seen in photograph B (Lake Arthur in the Sundays River system, Eastern Cape Province, South Africa) (photographs taken by E. R. Swartz, SAIAB).

Fig 2.5 Electric fishing as seen in the photograph (Western Cape Province, South Africa) was used to collect specimens in shallow habitats (photograph by K. Magellan, Rhodes University).

2.4 Genetic analysis in the laboratory

DNA extraction

DNA was extracted using a Promega purification kit. A small piece of tissue was placed in a 1.5 ml tube and mixed with 600µl of EDTA/ lysis solution and 15µl of Proteinase K to break down the tissue. The solution was incubated for 1-3 hours at 55°C to speed up digestion and was stirred with a vortex after every 30 min until the tissue was completely digested. RNAse (3 µl) was added to the solution, mixed with a vortex and then the solution was incubated for 15-30 mins at 37°C. Protein precipitation solution (200 µl) was added, stirred with a vortex for 20 sec and placed on ice for 5min to isolate DNA. The solution was centrifuged for 5 min at 13000-16000 rpm. The precipitated proteins formed a white pellet at the bottom of the tube. The liquid mixture was removed from the tube into a clearly labelled 1.5 ml tube containing 600 µl isopropanol. The solution was centrifuged for 2 min to force the DNA to form a pellet at the bottom of the tube and the supernatant was decanted. DNA was washed with 600 µl of room-temperature ethanol (70%) twice and supernatant was decanted in both occasions. The DNA was air dried for 10-15 mins and rehydrated with 100 µl of rehydration solution and stored in a -20°C freezer. Before Polymerase Chain Reaction (PCR), 5 µl of extracted DNA mixed with 2 µl of Bromophenol Blue was visualised by loading it on a 1% agarose gel containing Ethidium Bromide. Gel electrophoresis was used for 15–20 min at 100 mV while submerged in 1X TBE buffer. The gel was then visualised under ultra-violet light (UVP Translluminator), to verify the quality of the DNA extracted.

Polymerase chain reaction (PCR)

Polymerase chain reaction is described generally here but specific temperature and primer usage are described in chapters 4 and 5. The amplification of DNA involved three steps. Step 1 is the initial denaturation. The PCR reaction requires a single-stranded template. A high temperature is required for denaturation, the splitting of double-stranded template DNA into single-strands. The denaturation temperature was set to 92 ºC or 95 ºC for 1-2 mins. Step 2 was 35 cycles of denaturation at 92 ºC or 95 ºC for 30 sec, annealing at 55 °C or 58 °C for 30-60 sec and extension at 72 °C for 60 sec. The temperature must be lowered for primers to anneal at specific locations (primers complements) depending on which primers are used during annealing. During the extension step, which requires the temperature to be increased to 72 °C, taq polymerase adds nucleotides (adenosine, thymine, guanine and cytosine) contained in a solution of dNTPs. Lastly, one cycle of extension at 72 ºC was required for 10 min. The final product was viewed after gel electrophoresis as described above (section 2.4.1) to determine if the DNA amplified or not.

DNA purification

Most of the PCR products were sent to Macrogen (South Korea) for purification and sequencing, while others were purified using the QIAquick PCR Purification kit. PB buffer (500 µl) was added to PCR products before being poured into a QIAquick column to bind amplified DNA to the filter of the columns. PE buffer (750 µl) was used to wash the DNA. Thereafter, the column was placed in a labelled 1.5 ml tube and 50 µl EB buffer was poured into the column to elute the DNA.

DNA sequencing

Cycle sequencing was performed in 20 µl volumes containing 2 µl of purified DNA template, 3 µl of buffer, 2 µl of one of the above mentioned cytochrome *b* primers (m13F for cytochrome oxidase subunit I) and 2 µl of ABI Prism Big dye Terminator Cycle Sequencing Ready Reaction Kit (Southern Cross Biotechnology). A microcentrifuge tube (0.6 ml) was prepared containing 1 µl of 3M sodium acetate (NaOAc) with pH 4.6, 1 µl of 0.25 M EDTA with pH 8 and 50 µl of 99 % ethanol (EtOH). After the cycle sequencing product was added, the mixture was stirred with a vortex and left for an hour to precipitate. After the solution was centrifuged to allow the DNA to form a pellet at the bottom of the tube, the supernatant was decanted and discarded. The pellet was washed with 250 µl of 70 % ethanol and the supernatant was again discarded. The pellet was then dried at 60 °C for 5-8 mins. These products were sent to Rhodes University Micro biology laboratory for sequencing.

Sequence analysis

Sequences were edited manually in Lasergene v.7.2.0 SeqMan (DNA Star, Inc., Madison, WI) to check for mistakes and ambiguity codes were assigned to heterozygous positions.

Chapter 3: The use of morphology to differentiate between *Labeo capensis, L. umbratus* and hybrids of the two species

3.1 Introduction

Labeo capensis (Smith, 1841) and *L. umbratus* (Smith, 1841) are two large cyprinids closely resembling each other. They are believed to be sister species (Reid, 1985) with several morphological characters that overlap between them (Van Vuuren *et al*., 1990). Smith (1841) originally described the two species calling them *Abrostomus capensis* and *Abrostomus umbratus*. The exact type localities for these two species are not known (Reid, 1985), because Smith (1841) only gave a description of the broad area where the type material was from. He described the locality were the specimen of *A. Capensis* was found as many rivers of the Cape Colony, south to the Orange River, and for *A. umbratus* as north to the Orange River. Smith (1841) distinguished between the two species mainly on the basis of differences in lateral line scale counts (scales size). However, the type specimens now appear to be lost (Reid, 1985). According to Reid (1985), Smith's African specimens were sent to the United Kingdom and some of them made it to the British Museum of Natural History (BMNH), while others were deposited into the collection of the Musée d´Histoire Naturelle Paris (MHNP). No type localities for designated precisely by Smith (1841). This deficiency was remedied to some extent by Boulenger (Reid, 1985).

According to Reid (1985), Boulenger placed both species in the genus *Labeo* and he described neolectotypes for them from eight specimens of *L. capensis* from one locality and six specimens of *L. umbratus* from four localities. These are housed in the BMNH. Boulenger's redescription of *L. capensis* (Smith) highlighted 44-50 lateral line scales, 20-24 scales around the caudal peduncle, 10-11 dorsal rays with the anal fin nearly reaching the caudal fin. The redescription of *L. umbratus* emphasized 58-65 lateral line scales, 30-34 scales around the caudal peduncle, 8-10 dorsal rays with the anal fin not reaching the caudal fin. Boulenger's diagnosis of the two species was accepted by Gilchrist and Thompson (1913), Du Plessis (1963), Jubb (1967) and Reid (1985). When Reid (1985) revised the African *Labeo* species, he assigned neotypes for both species from the collection of neolectotypes specimens described by Boulenger from the Vaal River, Vredeford Road, Orange River colony, southern Africa (22° 07´S; 22 ° 07´E). The two neotype specimens were selected based on their superior condition and because they had a precise type locality (Reid, 1985).

Reid (1985) was able to distinguish between 13 individuals of *L. capensis* and 17 individuals of *L. umbratus* with both morphometric (ratio of head length, inter orbital width, eye diameter, pectoral fin length and dorsal fin base) and meristic characters (number of scales, vertebrae and dorsal fin rays). Van Vuuren *et al.* (1990) also compared populations of the two species by morphological means from Hardap Dam in Namibia (33 *L. capensis* and 30 *L. umbratus*) and the Vaal Barrage in South Africa (50 each), and found that of the characters used by Reid (1985) only the caudal peduncle scale, dorsal fin spine and lateral line scale counts were useful in distinguishing between the two species, but the lateral line scale count of Hardap *L. umbratus* overlapped with that of the two populations of *L. capensis*.

Gaigher and Bloemhof (1975) and Van Vuuren *et al.* (1989) conducted studies intended to identify hybrids in Hardap Dam. Gaigher and Bloemhof (1975) compared ten specimens of each species which they suspected to be pure and ten specimens of suspected hybrids. To distinguish between the three populations they counted the number of scales between lateral line and the origin of the dorsal fin, between the lateral line and origin of the pelvic fin, along the lateral line and around the caudal peduncle. They also used morphometric characters expressed as a percentage of standard length, namely head length, preopecular length, operculum length, inter orbital width, inter nasal width, dorsal fin length and anal fin length. They found that the intermediate forms had some characters that were closer to *L. capensis* (dorsal fin length, anal fin length, origin of dorsal fin to lateral line distance, origin of pelvic fin to lateral line distance and lateral line scale counts), some that were similar to *L. umbratus* (head length and the distance between the orbit and preopecular length) and some that were intermediate between the two species (operculum length, inter nasal length and caudal peduncle scale counts). One of the specimens that they identified as *L. umbratus* had a lateral scale count of 46, which is close to that of *L. capensis*.

Van Vuuren *et al*. (1989) used morphometric (standard length, body depth, head length, eye diameter and head width) and meristic (lateral line and caudal peduncle scales, dorsal and anal fin spines and rays, gill-rakers and vertebrae) characters to identify hybrids in

Hardap Dam. They found no significant differences between the three populations in gill-rakers, vertebrae and spines and rays of the anal fin. Similar to Gaigher and Bloemhof (1975), they also found that the intermediate forms were closer to *L. capensis* with regards to some characters (lateral line scale and dorsal fin spine counts), closer to *L. umbratus* in others (eye diameter, head width and head form) and some characters were intermediate between the two species (position of the mouth, number of scales around the caudal peduncle and head length). However these authors did not have a clearly defined classification system for pure fish and hybrids and they did not analyse specimens individually.

For these reasons, an attempt was made to morphologically differentiate between *Labeo* populations using samples from a wider range of localities, and by applying multivariate analysis to simultaneously consider the variation in several characters and thereby assess the similarities between individual samples (Turan, 1999). Multivariate analysis has been used successfully to assess morphological variation in fishes, e.g. cyprinids (McElroy and Douglas 1995 and Freyhof *et al*., 2005) and brown trout (Hermida *et al*., 2009) to assess the morphological variation. In the present study, more specimens are analysed across a range of localities that are more representative of the overall distribution of the two species, than in previous studies. Populations were assumed to be pure when there were no previous reports of hybridisation. Areas where hybridisation was suspected (i.e. Hardap and Darlington dams) were not classified *a priori* as in previous studies, but rather compared to pure populations that have been classified after analysis was done. Individuals were analysed separately and data were not lumped and analysed as populations as was the case with Gaigher and Bloemhof (1975) and Van Vuuren *et al.* (1989).

The aim of this chapter was to assess which morphological and meristic characters could be used to reliably identify *L. capensis* and *L. umbratus* and to reassess whether morphology is useful in identifying hybrids. The following hypotheses were tested based on the overall thesis hypotheses outlined in Chapter 1:

1) *L. capensis* and *L. umbratus* can be distinguished using morphology.
2) The two species hybridise in Hardap and Darlington dams.
3) Morphology can be used to identify hybrids.

3.2 Materials and methods

Sampling

A total of 275 specimens of *L. umbratus* (N = 115), *L. capensis* (N = 60) and from suspected hybrid populations (N = 100) were assessed for morphological variation. Specimens were analysed from across the geographic range of *L. umbratus* and *L. capensis* (Chapter 2, Table 2.1). Without access to the neolectotype specimens for both species at the British Museum of Natural History, specimens were collected at or as close as possible to the type locality in the Vaal River (Reid 1985). Measurements of specimens were taken after fixation in 10% formalin and transfer to 70% ethanol through a gradient of 10% and 50% ethanol to prevent rapid dehydration and shrinkage of the specimens.

Character measurements and counts

An initial analysis was done on ten *L. umbratus* and ten *L. capensis* specimens of different sizes from Gariep Dam to determine if specific characters or a combination of characters can be used to distinguish these two species. A total of 45 morphometric characters were measured, including the morphometric characters used in previous studies (Gaigher and Bloemhof, 1975; Reid, 1985; Hubbs and Lagler, 1947; Van Vuuren *et al*., 1989; Van Vuuren *et al*., 1990). Measurements greater than 20cm were taken using a plastic measuring tape to the nearest 1mm, while those less than 20cm were taken using Helios Vernier callipers to an accuracy of the nearest 0.1mm.

The morphometric characters were:

1. Standard length
2. Dorsal fin length
3. Dorsal fin base length
4. Caudal peduncle depth
5. Caudal peduncle length
6. Posterior dorsal fin to dorsal caudal fin base
7. Pectoral fin length, pelvic fin length
8. Pectoral fin to pelvic fin length

9. Anal fin length

10. Anal fin base length

11. Pectoral fin to anterior anal fin

12. Pectoral fin to posterior anal fin

13. Pelvic fin to anterior anal fin

14. Pelvic fin to posterior anal fin

15. Anterior dorsal fin to pectoral fin

16. Posterior dorsal fin to pectoral fin

17. Anterior dorsal fin to pelvic fin

18. Posterior dorsal fin to pelvic fin

19. Anterior dorsal fin to anterior anal fin

20. Posterior dorsal fin to anterior anal fin

21. Anterior dorsal fin to posterior anal fin

22. Posterior dorsal fin to posterior anal fin

23. Pectoral fin to ventral caudal fin base

24. Pectoral fin to dorsal caudal fin base

25. Pelvic fin to ventral caudal fin base

26. Pelvic fin to dorsal caudal fin base

27. Anterior anal fin to dorsal caudal fin base

28. Anterior anal fin to ventral caudal fin base

29. Posterior anal fin to dorsal caudal fin base

30. Posterior anal fin to ventral caudal fin base

31. Operculum to pre-operculum

32. Operculum to eye

33. Head length

34. Eye to snout length

35. Eye diameter

36. Inter-orbital width

37. Snout to anterior anal fin

38. Snout to posterior anal fin

39. Snout to pectoral fin

40. Snout to pelvic fin

41. Snout to posterior dorsal fin

42. Snout to anterior dorsal fin

43. Eye to nostrils

44. Left to right nostrils

Of these, characters 2, 6-32, 34, 35 and 37-44 had small distances that could not be measured accurately or that were not significant. These were excluded from further analysis.

A total of 14 meristic characters were recorded. A dissecting microscope had to be used to do counts for smaller specimens. X-rays were taken of 10 specimens to count the vertebrae and to do more accurate counts of rays and spines.

The meristic characters were:

1. Number of scales along the lateral line
2. Number of scales between the lateral line and origin of the dorsal fin
3. Number of scales between the lateral line and origin of the pelvic fin
4. Number of scales between the lateral line and origin of anal fin
5. Scales around the caudal peduncle
6. Dorsal fin spines
7. Pectoral fin spines
8. Pelvic fin spines
9. Anal fin spines
10. Dorsal fin rays
11. Pectoral fin rays
12. Pelvic fin rays
13. Anal fin rays
14. Number of vertebrae

Meristic characters 4 and 6-14 were not significantly different between the two species and were excluded from further analysis. Both the morphometric and meristic datasets were checked for errors by graphing the ratios for each character against standard length. Outliers were re-measured or counted. Those outliers that remained unchanged after re-measurement or counting were included in the analysis. As the specimens were of different sizes, the percentage ratio of each character against standard length was calculated to eliminate the effect of size before statistical analysis (Reist, 1985).

In an initial evaluation, all 44 morphometric and 14 meristic characters were compared between ten *L. capensis* and ten *L. umbratus* from the Gariep Dam. Only a few showed variation between the two species. These characters and the characters that Reid (1985) suggested were diagnostic, where then measured for all 275 specimens. The morphometric characters screened for all individuals were the standard length, dorsal fin base length, caudal peduncle depth, caudal peduncle length, head length, operculum to eye and inter-orbital width (Fig. 3.1). The meristic characters screened for all individuals were the lateral line scales, lateral line to the origin of the dorsal fin scales, lateral line to pelvic fin scales and caudal peduncle scales (Fig. 3.1).

Statistical analyses

Principal component analyses (PCA) of morphometric and meristic characters were performed using the computer software paleontological statistics (PAST) version 1.3. The raw data collected from the meristic characters were log transformed in PAST to reduce count bias. PCA was used to reduce the multidimensionality of variables in the data to fewer, more significant uncorrelated components of variation. Eigenvalues [a measure of variance accounted for by eigenvectors (components) as a percentage of variance accounted for by the components] were recorded (Hammer *et al*., 2001). The Jolliffe cut-off value was used to select which components were significant (Jolliffe, 1986). The highest variation is accounted for by the first principle and that makes it the best component to correlate variables with either of the other components. These loadings were used to check which variables (characters) correlated best with the components.

Scatter plots were drawn to show the variation in the data. Basic statistics (mean, standard error mean, median, 95% confidence interval and permutation tests) were calculated. Discriminant function analysis was applied to the Orange River populations above the Augrabies Falls to support the PCA results by maximizing the difference between the two data sets (Hammer *et al*., 2001). A t-test parametric analysis was performed to test significance of characters.

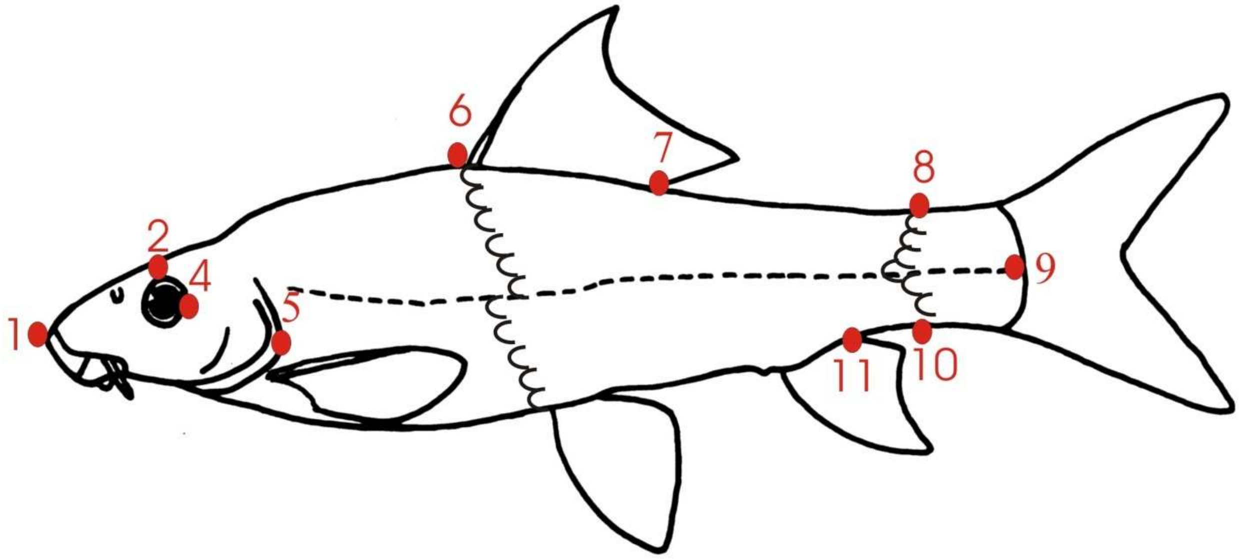

Fig. 3.1 A picture showing homologous points where measurements were taken on each *Labeo* specimens. The drawn scales indicate the position where scale counts were taken (excluding the scale count that followed the lateral line). The morphometric characters measured were head length (point 1-5), inter-orbital width (point 2 to the other side of the fish in the same position), eye to operculum (point 4-5), dorsal fin base (point 6-7), caudal peduncle depth (point 8-10) and caudal peduncle length (point 9-11).

For the first set of PCAs, *L. capensis* and *L. umbratus* populations from Gariep dam (juveniles vs. juveniles, adults vs. adults and species vs. species) were compared, since they are considered to be pure. Specimens above 25cm were considered adult, because Skelton (2001) suggested that *L. capensis* reach sexual maturity at 22cm for males and 24cm for females and that *L. umbratus* reach sexual maturity at 15cm for males and 25cm for females. A comparison was also done within species (adults vs. juveniles) to test for the influence of sexual maturity. PCAs were also performed on all suspected pure populations in the Orange River system above the Augrabies Falls (Orange River at Kanoneiland, Brak River, Vaal River and Gariep Dam). This was to assess whether an increased sample size would add to within species variation, thereby lowering confidence in distinguishing the two species, and to test whether the Gariep Dam individuals could be considered the same as other individuals above the Augrabies Falls. PCAs were also performed to compare Orange River *L. umbratus* from above the Augrabies Falls to populations of the same species from the southern flowing river systems of the Eastern Cape Province and eastern parts of Western Cape Province (Kat River Dam in the Great Fish River system, Slagboom Dam in the Sundays River system and sites in the Bushmans, Nahoon, Gamtoos and Gouritz river systems). Darlington Dam specimens were initially excluded as hybridisation was suspected.

PCAs were then carried out to identify possible hybrids. The first comparison was among populations across the entire Orange River system, including Hardap dam where hybrids have been recorded in the past and the lower Orange where some of these hybrids could have established. Secondly, a comparison was done among populations from the Orange River system above Augrabies Falls, the Great Fish and Sundays river systems as the IBT only involves these three river systems. This was done specifically to test whether hybrids could be identified in Darlington Dam. Finally, all specimens were compared to assess overall morphological differentiation and to test whether any additional hybrids could be identified.

3.3 Results

Natural variation in morphometric characters

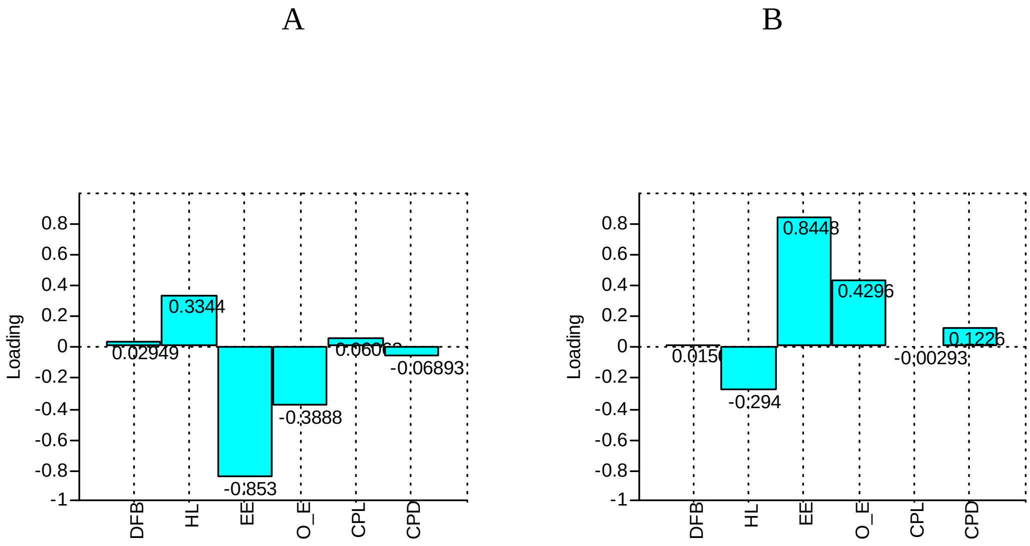

Fig. 3.2 Graphs showing PCA loading variation within Gariep (A) and populations from the Orange above Augrabies Falls (B) for *L. capensis* (adults vs. juveniles). DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

There was no significant difference between the Gariep populations when compared to other Orange River population above the Augrabies Falls based on the PCA loadings of within species (Fig. 3.2 and 3.3), adults vs. adults (Fig. 3.4), juveniles vs. juveniles (Fig. 3.5) and species vs. species (Fig. 3.6) comparisons, but there was a significant difference between adults vs. juveniles within *L. umbratus*. This was due to significant relationship ($p < 0.001$) and high positive values of inter-orbital width and operculum to the eye distance (Table 3.1).

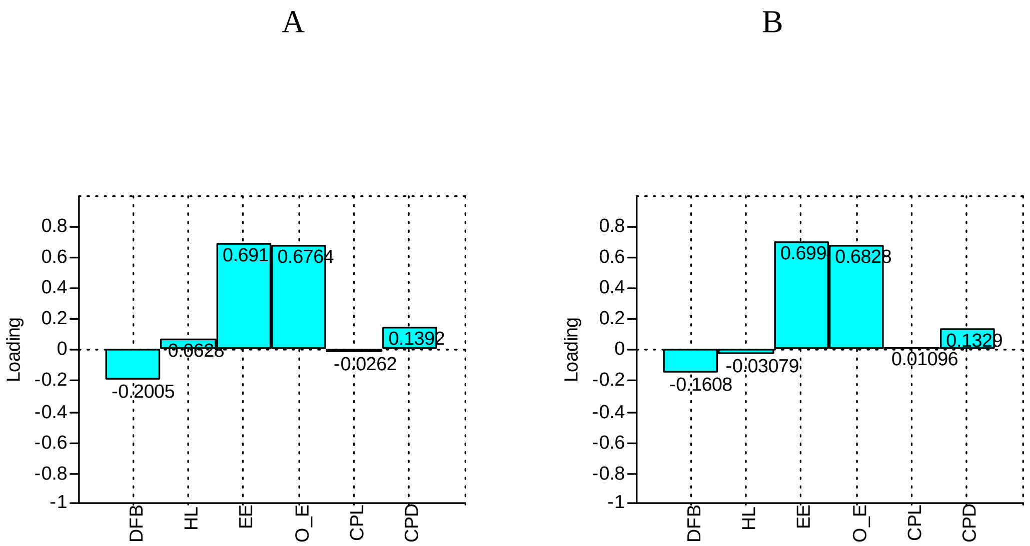

Fig. 3.3 Graphs showing PCA loading variation within Gariep (A) and populations from the Orange above Augrabies Falls (B) for *L. umbratus* (adults vs. juveniles). DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

Component one and two accounted for most of the variance > 88% (Table 3.2) with Jolliffe cut-off of 2.43 for the Eigenvalues (Jolliffe, 1986). The *L. capensis* and *L. umbratus* clusters were significantly different (Fig. 3.7) with six components accounting for the variation (Table 3.2). The Jolliffe cut-off was 6.53 for the Eigenvalues, meaning that only component 1 and 2 were significant and accounting for 92.37% of the variation (Eigenvalues 1 and 2; 42.8 and 8.9, % variance; 76.536 and 15.831) respectively (Table 3.3).

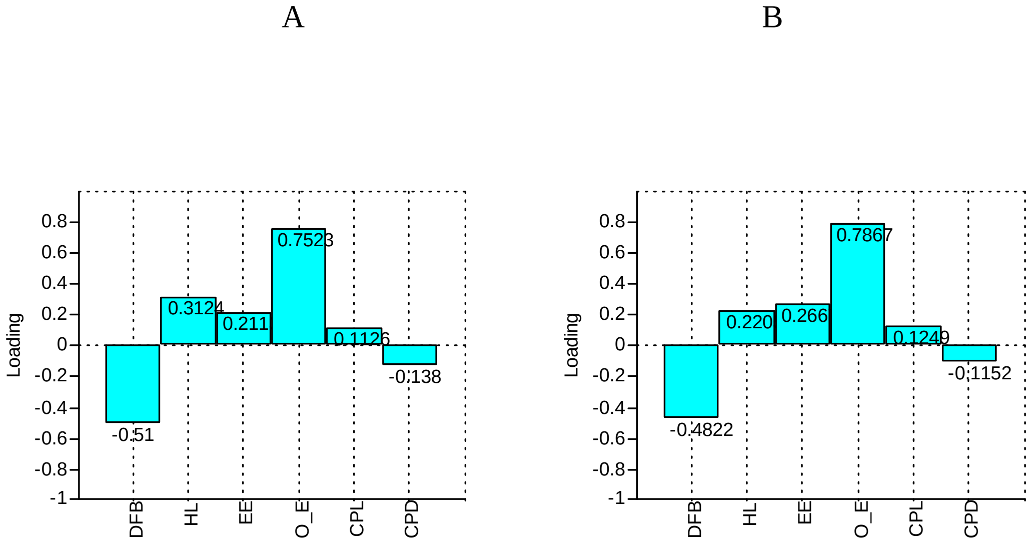

Fig. 3.4 Graphs showing PCA loading variation between Gariep (A) and populations from the Orange above Augrabies Falls (B) for *L. capensis* and *L. umbratus* (adults). DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

The discriminant function analysis revealed that a 100% of individuals of the two species could be correctly classified as belonging to that specific species. As component 1 and 2 were considered significant, the correlating variants with high values (dorsal fin base, inter-orbital width, operculum to the eye and head length) were regarded as significant as well (Table 3.4). The t-test also indicated that all these characters were significantly different between *L. capensis* and *L. umbratus* (all tests with p<0.001) (Table 3.5).

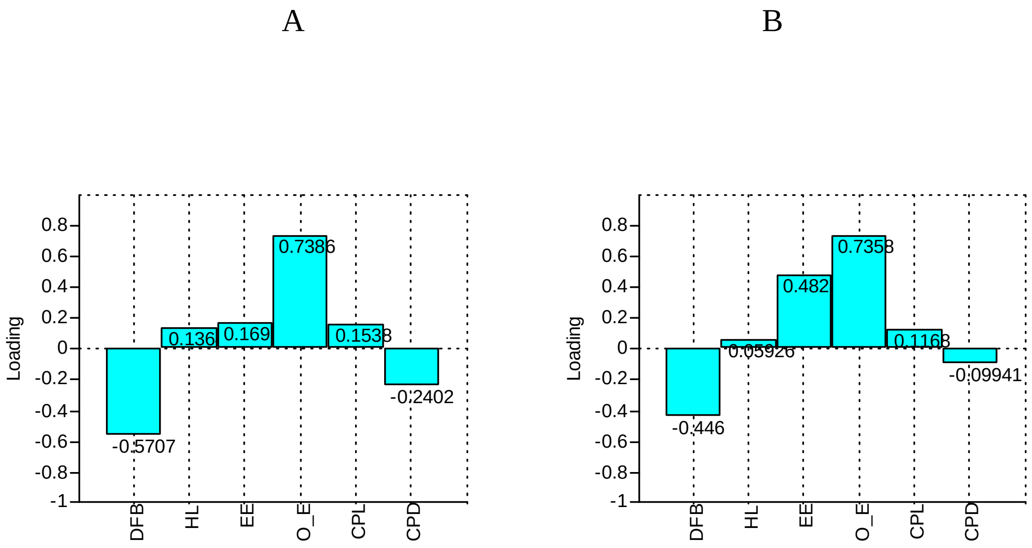

Fig. 3.5 Graphs showing PCA loading variation between Gariep (A) and populations from the Orange above Augrabies Falls (B) *L. capensis* and *L. umbratus* (Juveniles). DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

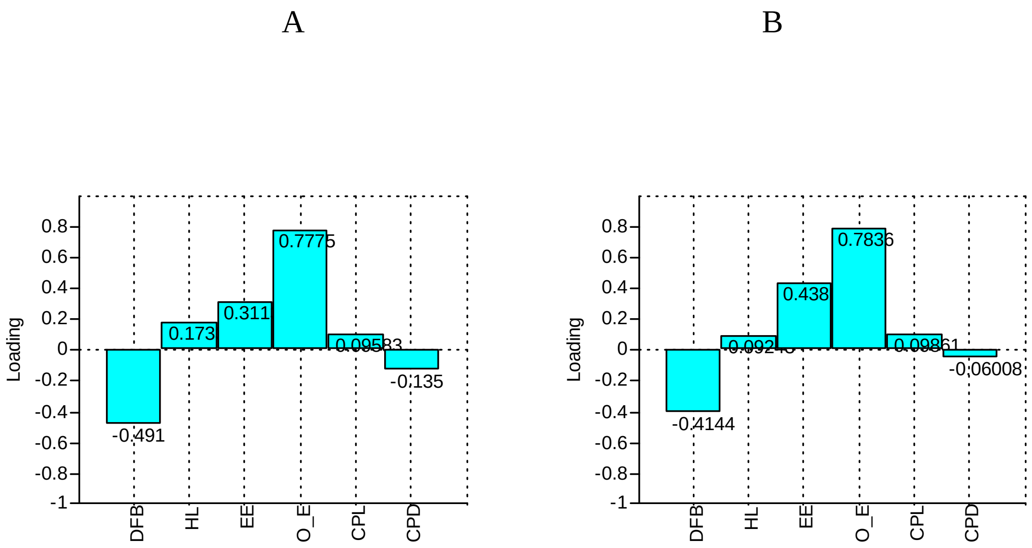

Fig. 3.6 Graphs showing PCA loading variation within Gariep (A) and populations from the Orange above Augrabies Falls (B) samples DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

Table 3.1 PCA loading demonstrating the correlation between the six principle components and variants using *L. umbratus* from the Orange River system above Augrabies Falls and with bold numbers indicating the best correlations. DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6
DFB	-0.160	0.060	-0.185	0.162	**0.816**	**0.495**
HL	-0.031	0.264	**0.865**	**0.384**	-0.016	0.183
EE	**0.699**	**-0.666**	0.148	0.120	0.114	0.136
O_E	**0.683**	**0.686**	-0.166	-0.134	-0.012	0.144
CPL	0.011	0.037	**-0.412**	**0.850**	**-0.315**	0.087
CPD	0.133	0.108	0.034	0.269	**0.471**	**-0.821**

Table 3.2 Eigenvalues from PCA analysis for the six principle components and their percentage variance of *L. umbratus* populations from the Orange River above Augrabies Falls.

PC	Eigenvalues	% Variance
1	16.001	76.819
2	2.476	11.883
3	1.154	5.538
4	0.616	2.955
5	0.310	1.487
6	0.275	1.319

Table 3.3 Eigenvalues from PCA analysis for the six principle components and their percentage variance for populations of *L. capensis* and *L. umbratus* from the Orange River above Augrabies Falls.

PC	Eigenvalues	% Variance
1	42.838	76.536
2	8.861	15.831
3	1.819	3.249
4	0.973	1.739
5	0.849	1.517
6	0.631	1.128

Table 3.4 PCA loading demonstrating the correlation between the six principle components and variants using populations from the Orange River system above Augrabies Falls and numbers in bold indicate strong correlations. DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6
DFB	**-0.414**	**0.379**	**-0.633**	-0.034	**-0.304**	**-0.437**
HL	0.092	**-0.415**	0.110	**0.344**	**-0.828**	-0.057
EE	**0.439**	**0.767**	**0.368**	0.020	-0.271	-0.096
O_E	**0.784**	-0.147	**-0.576**	0.060	0.118	-0.123
CPL	0.099	-0.113	-0.033	**-0.919**	**-0.329**	0.156
CPD	-0.060	0.248	**-0.346**	0.181	-0.161	**0.870**

The PCA (Fig3.8) of *L. umbratus* adults from the Orange River system above the Augrabies Falls compared to those from southern flowing river systems suggests morphological differentiation between these two areas, but there was overlap between *L. umbratus* adults from above Augrabies Falls and the Slagboom Dam population. The Great Fish specimens had a slight overlap with other southern populations. There was no difference among juveniles from the Orange above Augrabies Falls and southern flowing river systems, suggesting that morphometric differences between these areas can only be observed in later stages of development.

Table 3.5 Basic descriptive statistics and t-test results for each character investigated in the two *Labeo* species from the Orange River system above Augrabies Falls, showing the difference in mean and the significance of each character using p-values. DFB = Dorsal fin base length, HL = Head length, EE = Inter-orbital width, O_E = Operculum to eye, CPL = Caudal peduncle length and CPD = Caudal peduncle depth

	Labeo capensis n=50			*Labeo umbratus* n=53		
	Mean	Std. error	p-value	Mean	Std. error	p-value
DFB	21.236	0.150	<0.001	15.386	0.119	<0.001
HL	22.534	0.200	<0.001	24.585	0.147	<0.001
EE	46.143	0.525	<0.001	49.341	0.411	<0.001
O_E	37.065	0.335	<0.001	45.971	0.405	<0.001
CPL	15.379	0.155	<0.001	16.842	0.113	<0.001
CPD	10.184	0.152	<0.001	8.670	0.107	<0.001

Another indication of this is that all specimens from the Bushmans and Nahoon were juveniles that grouped with juveniles from other areas, while the Slagboom specimens were all adults grouping with adults from other areas. These populations therefore appear different in PCAs where adults and juveniles are mixed. Morphological variation within *L. capensis* was lower than in *L. umbratus*. The latter had significant differences between adults and juveniles from the Orange River system above Augrabies Falls and the southern flowing river systems. This was mainly due to two morphometric characters of the head (inter-orbital and operculum to eye measurements), suggesting an increase in the head to body size ratio (Fig. 3.8 and Table 3.1). The samples that are clustered with *L.*

umbratus adults, but that are indicated as juveniles in Fig 3.8, are samples that were close to the size of sexual maturity and can therefore be considered to be sub-adults (Fig. 3.8).

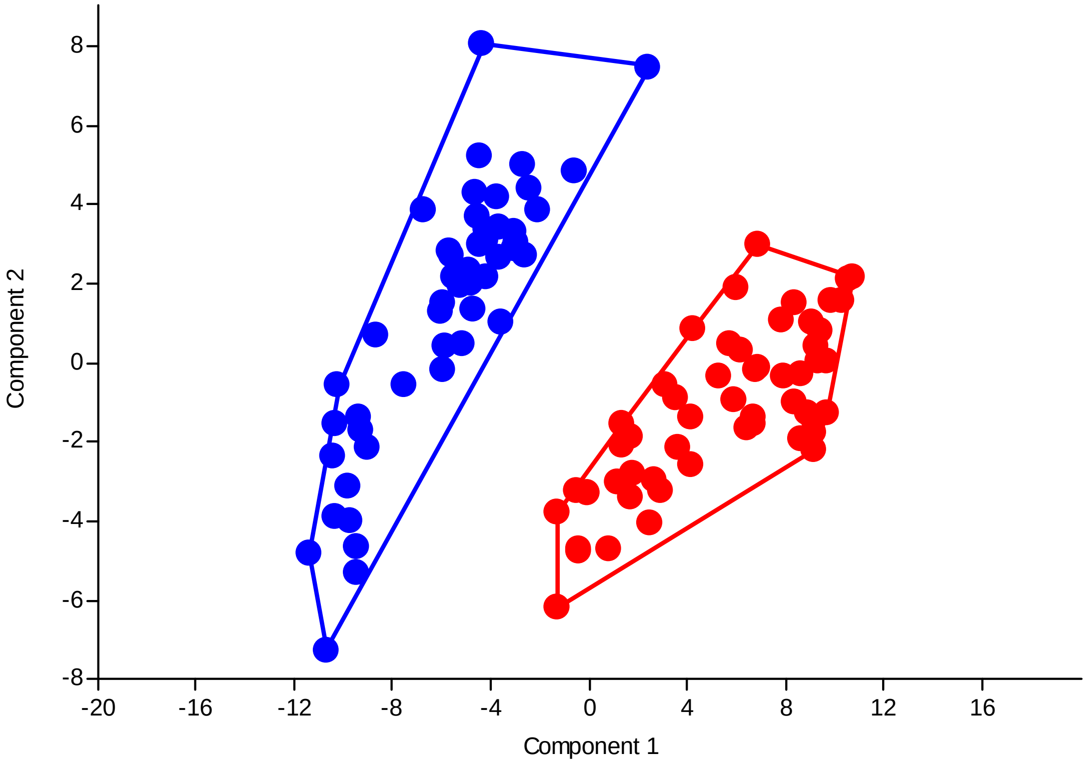

Fig 3.7 PCA plot comparing morphometric variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls, showing the differentiation between the two species.

Comparison of populations of the two species from the Orange above Augrabies Falls are analysed with the southern flowing river populations, all *L. umbratus* populations form one group that is distinct from the *L. capensis* populations (Fig. 3.9).

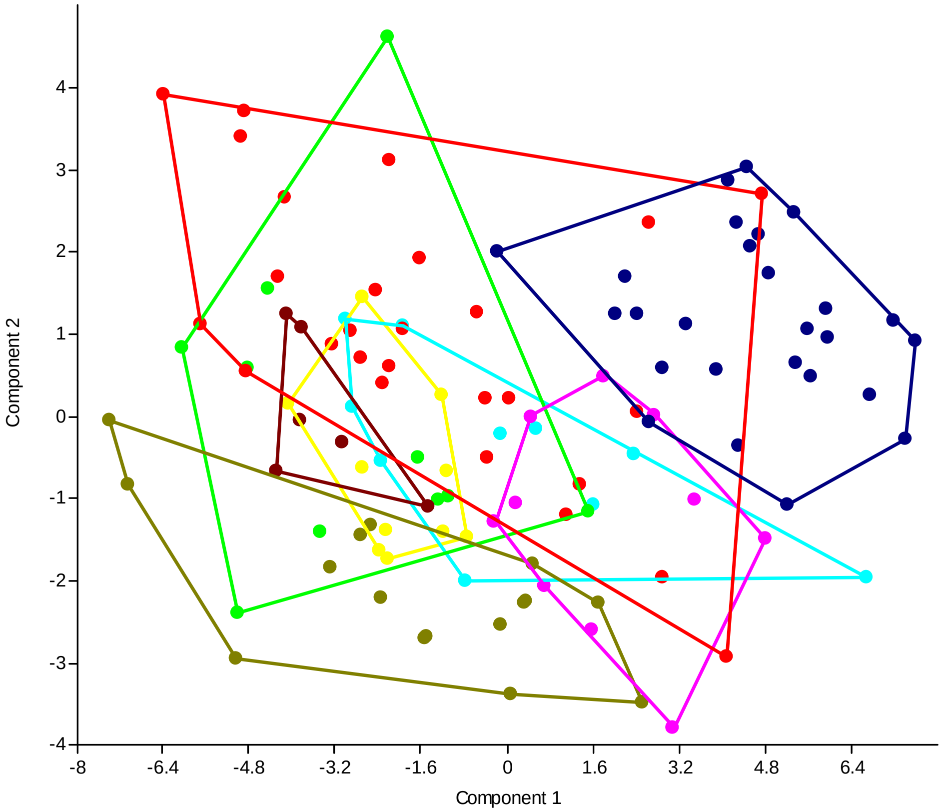

Fig. 3.8 PCA plot comparing morphometric variation in characters of *L. umbratus* from the Orange River above Augrabies Falls (**red**) and *L. umbratus* from southern flowing river systems, namely the Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**).

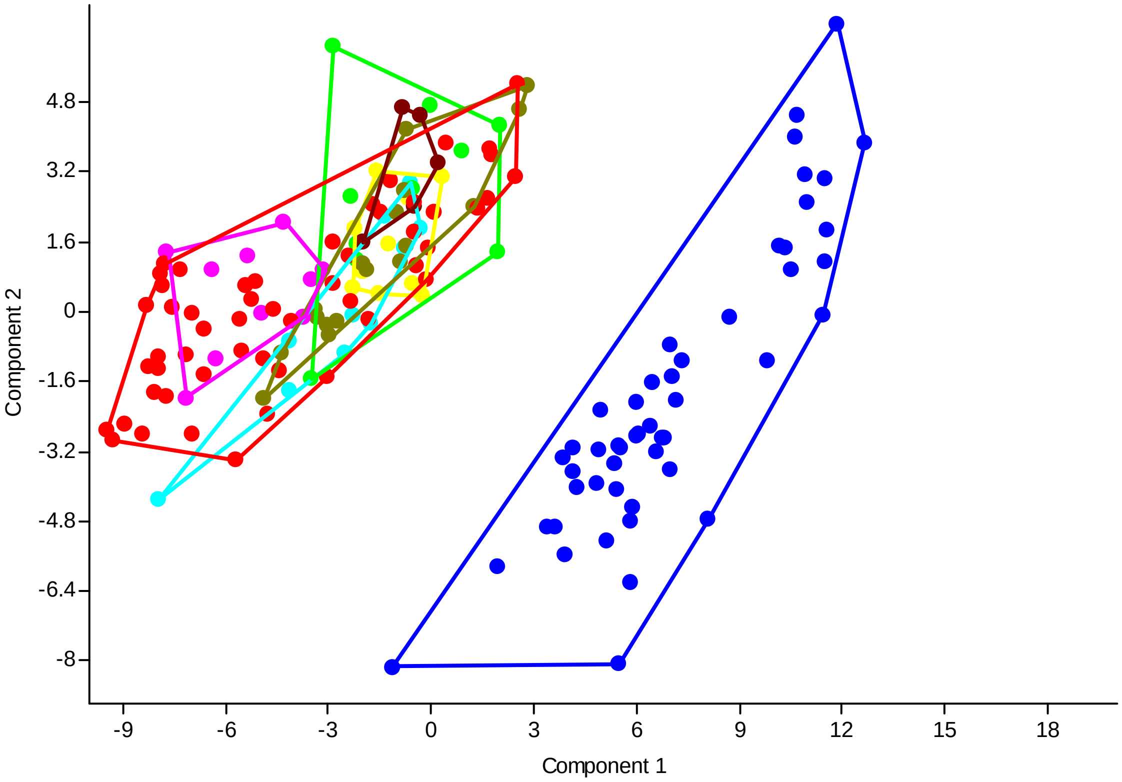

Fig. 3.9 PCA plot comparing morphometric variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from Orange River above Augrabies Falls and *L. umbratus* from southern flowing river systems, namely the Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**).

Hybrid identification using morphometric data

The PCA analysis of all Orange River system specimens (adults and juveniles) including Hardap Dam and the lower Orange River revealed two distinct clusters representing *L. capensis* and *L. umbratus* respectively (Fig. 3.10). All the specimens from the lower Orange clustered with *L. capensis*. Some specimens from Hardap Dam clustered with *L. capensis* (N=10) and some clustered with *L. umbratus* (N=7). However, most Hardap Dam specimens (N=35) clustered between what was considered the two parental species clusters and were therefore regarded as potential hybrids.

The PCA (Fig. 3.11) comparing both species from the Orange above Augrabies Falls to both species in the Great Fish and Sundays river systems also revealed two distinct clusters that can be ascribed to pure specimens of the two species. Most of the specimens (N=33) from Darlington Dam clustered with *L. umbratus*. Some specimens clustered with *L. capensis* (N=11), with only four specimens that were intermediate between the two clusters and therefore classified as potential hybrids.

When all the specimens were analysed together to show the overall variation, the two clusters representing potentially pure *L. capensis* and *L. umbratus* specimens were observed with potential hybrids between them (Fig. 3.12).

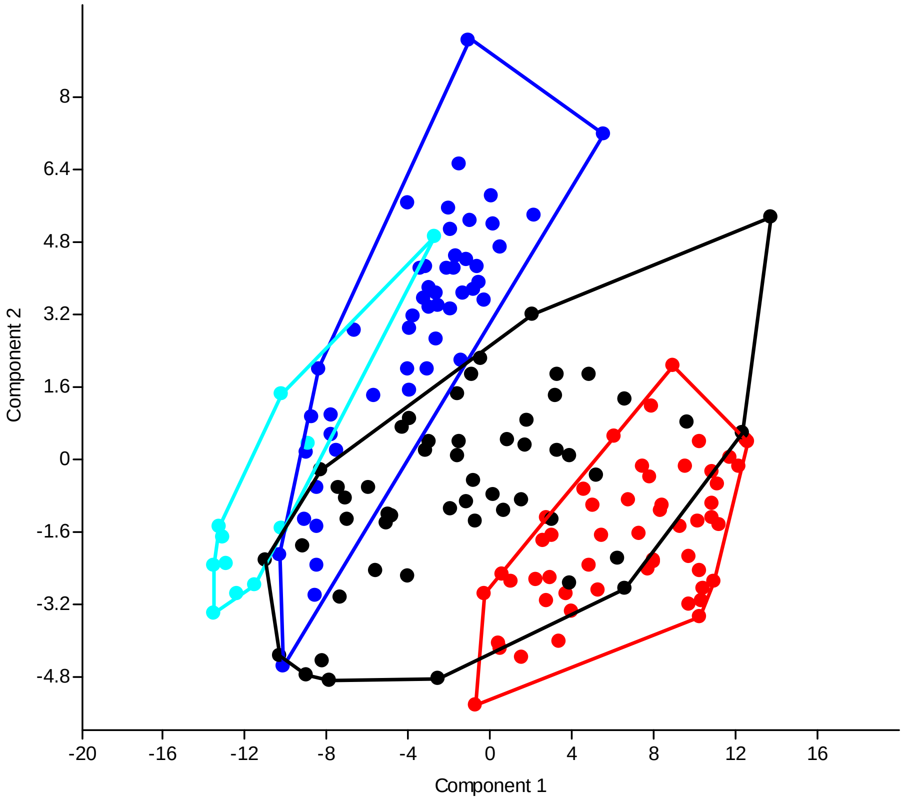

Fig. 3.10 PCA plot comparing morphometric variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls, *L. capensis* from the lower Orange River (**sky blue**) and Hardap Dam *Labeo* specimens (**black**), showing the difference between the two species and hybrids between them.

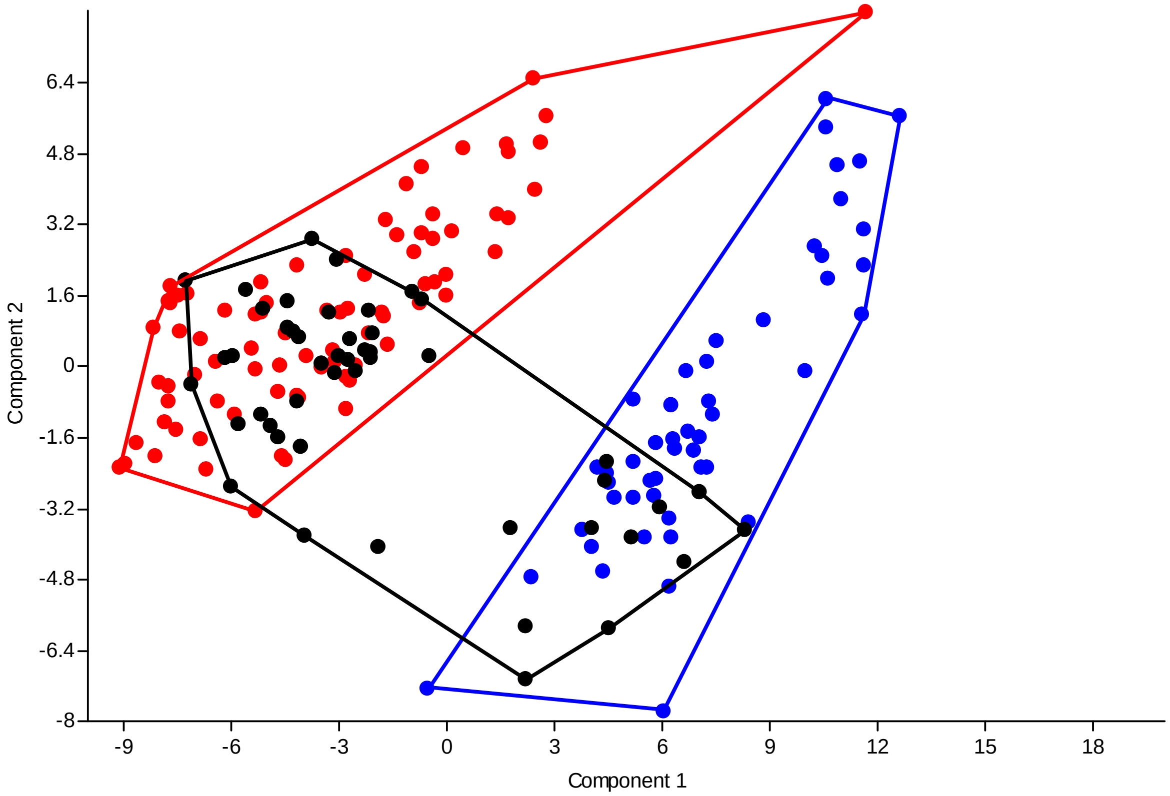

Fig. 3.11 PCA plot comparing morphometric variation in characters of *L. capensis* from the Orange River above Augrabies Falls (**blue**), *L. umbratus* (Orange, Great Fish and Sundays river systems) (**red**) and *Labeo* specimens from Darlington Dam (**black**), illustrating pututive hybrids between the two species.

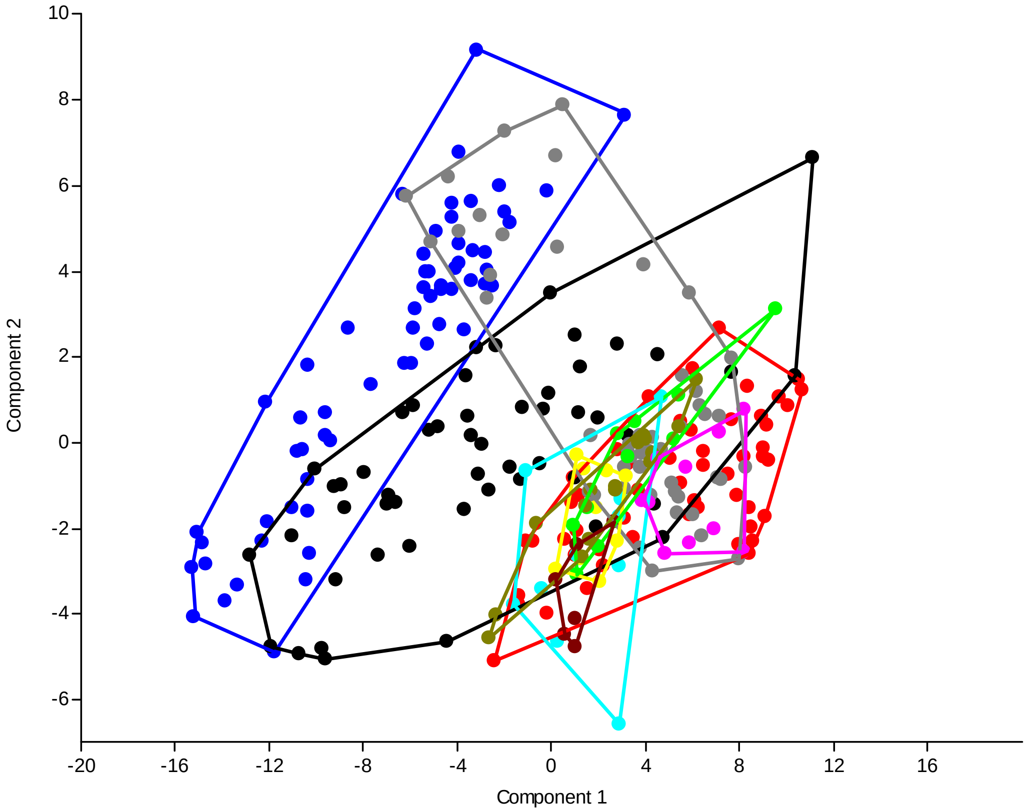

Fig. 3.12 PCA plot comparing morphometric variation in characters showing overall variation between all populations, namely *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls**,** *L. umbratus* from southern flowing river systems (Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**)), Darlington Dam (**grey**) and Hardap Dam (**black**).

Natural variation in meristic characteristics

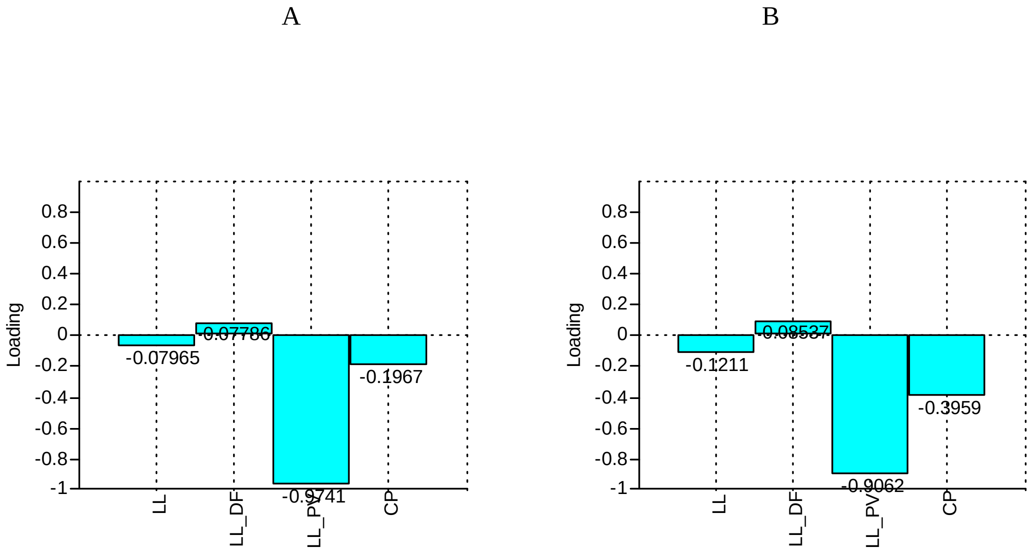

Fig. 3.13 Graphs showing PCA loading variation within Gariep (A) and the Orange above Augrabies Falls (B) for *L. capensis* (adults vs juveniles). LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

There was no difference in PCA loadings between the PCA of the Gariep population and the PCA of the population from the Orange above Augrabies Falls (Fig. 3.13 and 3.14), adults vs. adults (Fig. 3.15), juvenile vs. juvenile (Fig. 3.16) and species vs. species (Fig. 3.17). However, there was increased variation because of the increased population size. Size and sexual maturity did not influence the results based on the PCA loading results.

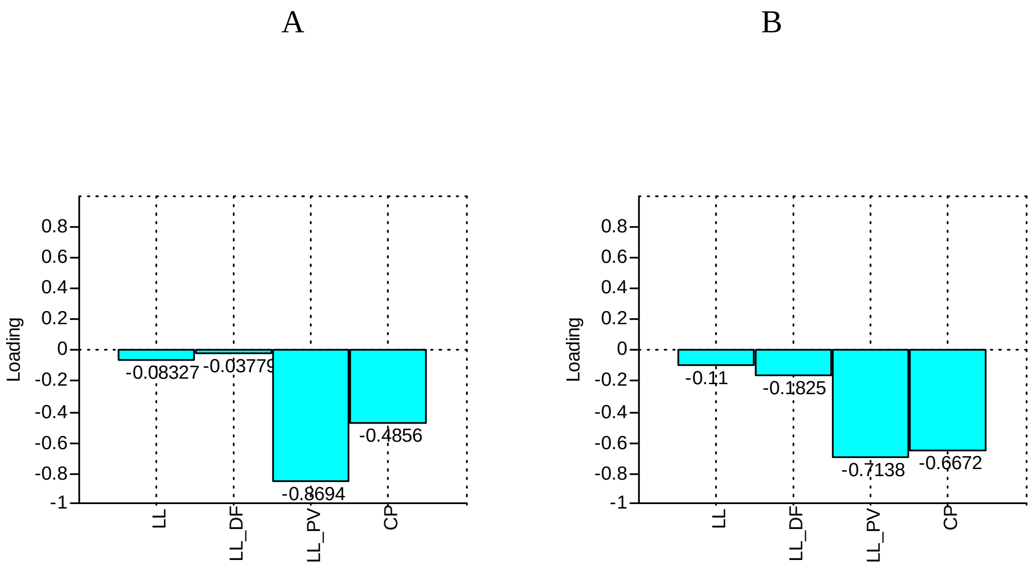

Fig. 3.14 Graphs showing PCA loading variation within Gariep (A) and the Orange above Augrabies Falls (B) for *L. umbratus* (adults vs. juveniles). LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

Of all the meristic characters investigated only two (scale counts from the lateral line to the pelvic fin and around the caudal peduncle) were significant and accounted for most of the variation within species. Loadings within species (*L. capensis* and *L. umbratus*) indicated that the variance was almost spread evenly amongst the components with components 1 the highest at 44.47 (48.06%). Similar to the morphometric results (see Fig 3.7), all individuals of the two species from the Orange above Augrabies Falls clustered separately. The two species clusters were significantly different (Fig. 3.18) with *L. umbratus* showing more variation. This was probably because of the large number of small scales it has. Four components accounted for the overall variance (Table 3.6). The

Jolliffe cut-off was 0.004 for the Eigenvalues, meaning that only component 1 was significant and accounted for 90.50% of the variance (Eigenvalue of 0.018 and variance of 90.50).

Component 1 and 3 were chosen as they seemed to show the best result (see page 32 for explanation). The discriminant function analysis also revealed a 100% correct classification for individuals of the two species. All the meristic characters tested significant from the loadings in Table 3.7. The t-test for the characters were all significant (all with $p < 0.001$), suggesting that all the characters contributed significant variation (Table 3.8).

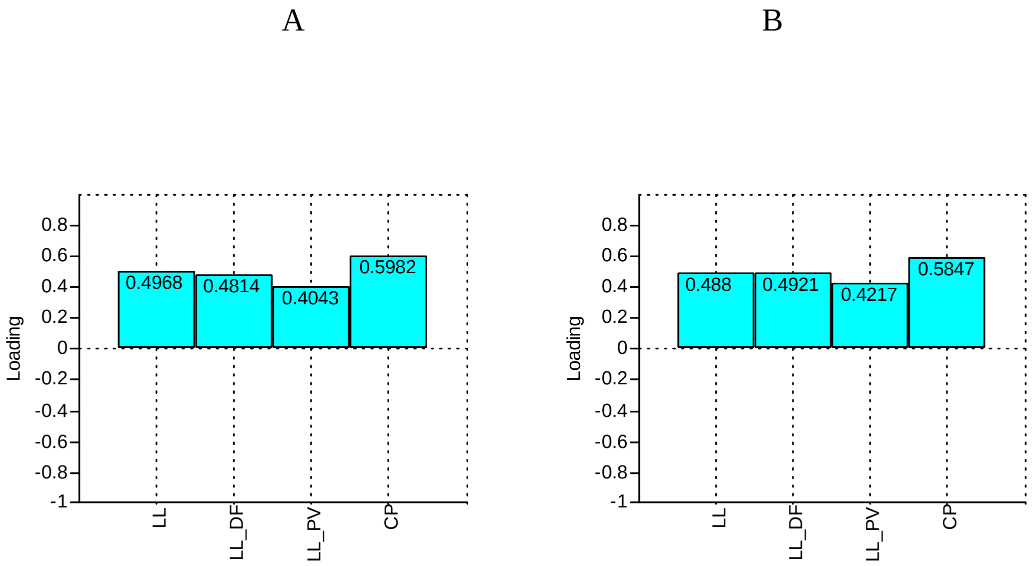

Fig. 3.15 PCA loading graphs showing variation between Gariep (A) and the Orange above Augrabies Falls (B) for *L. capensis* and *L. umbratus* (adults). LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and

origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

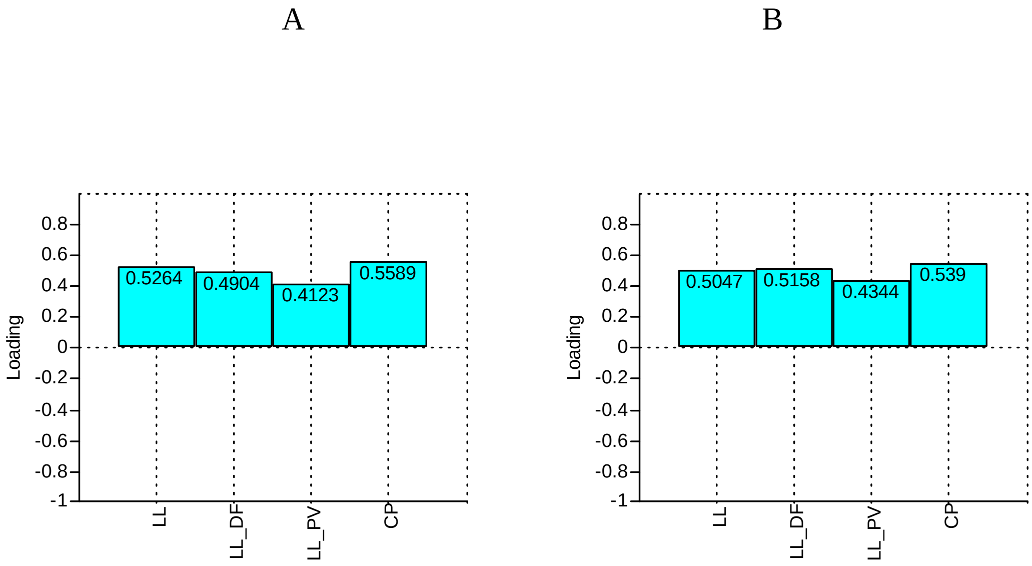

Fig. 3.16 Graphs showing PCA loading variation between Gariep (A) and Orange above Augrabies Falls (B) *L. capensis* and *L. umbratus* (juveniles). LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

The PCA of *L. umbratus* meristic variation from the Orange above Augrabies Falls compared to specimens from southern flowing rivers systems (Fig. 3.19) revealed no differences unlike the morphometric results (Fig. 3.8). However, there was a significant difference between some of the southern flowing rivers. The Gamtoos population is significantly different from those in the Great Fish and Bushmans, but overlapped with the Nahoon. The Gouritz and the Gamtoos populations were different, but one of the Gouritz samples clustered with specimens from the Great Fish.

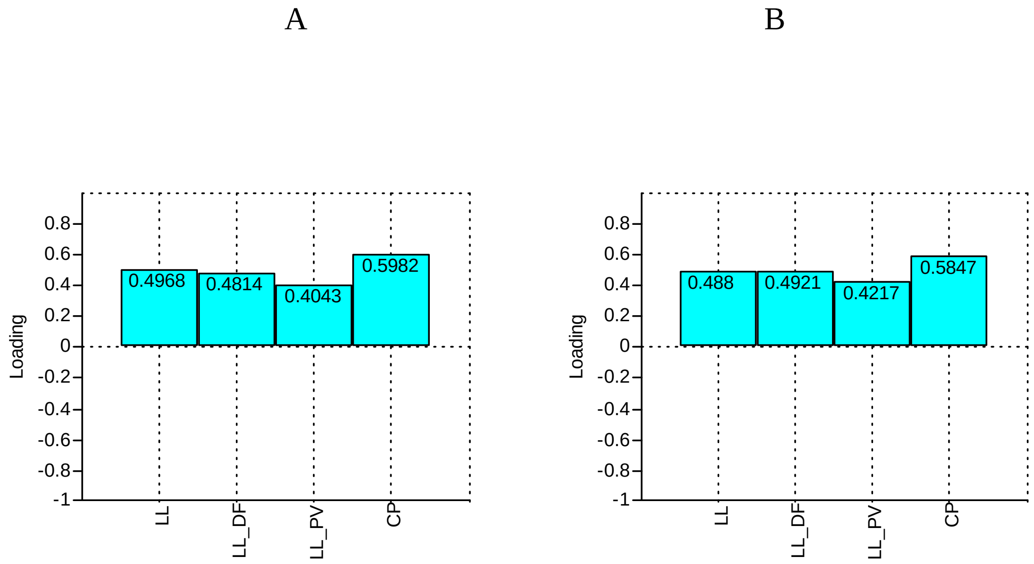

Fig. 3.17 Graphs showing PCA loading variation within Gariep (A) and Orange above Augrabies Falls (B) *L. capensis* and *L. umbratus*. LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

Table 3.6 Eigenvalues for the four principle component (PC) analyses and the percentage variance of individuals of both species from the Orange above Augrabies Falls.

	Eigenvalues	% Variance
PC1	0.018	90.50
PC2	0.001	4.888
PC3	0.0001	3.266
PC4	<0.001	1.345

Table 3.7 PCA loading demonstrating the correlation between the four principle components and variants using species individuals from the Orange above Augrabies Falls and with bold numbers indicating the best correlations. LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

	Axis 1	Axis 2	Axis 3	Axis 4
LL	**0.488**	-0.026	**0.499**	**0.716**
LL_DF	**0.492**	**-0.306**	**0.464**	**-0.670**
LL_PV	**0.422**	**0.882**	-0.117	-0.175
CP	**0.585**	**-0.357**	**-0.723**	0.092

Table 3.8 Basic descriptive statistics and t-test results for each character of individuals of the two *Labeo* species from the Orange River system above Augrabies Falls, showing the difference in mean and the significance of each meristic character with p-values. LL = number of scales along the lateral line, LL_DF = number of scales between the lateral line and origin of the dorsal fin, LL_PV = number of scales between the lateral line and origin of the pelvic fin and CP = scales around the caudal peduncle.

	LL		LL_DF		LL_PV		CP	
	cap	umb	cap	umb	cap	umb	cap	umb
N	50	52	50	52	50	52	50	52
Mean	1.633	1.764	0.955	1.086	0.835	0.941	1.357	1.506
Std. error	0.002	0.003	0.003	0.003	0.004	0.005	0.004	0.005
p-value	<0.001		<0.001		<0.001		<0.001	

cap = *L. capensis* and umb = *L. umbratus*

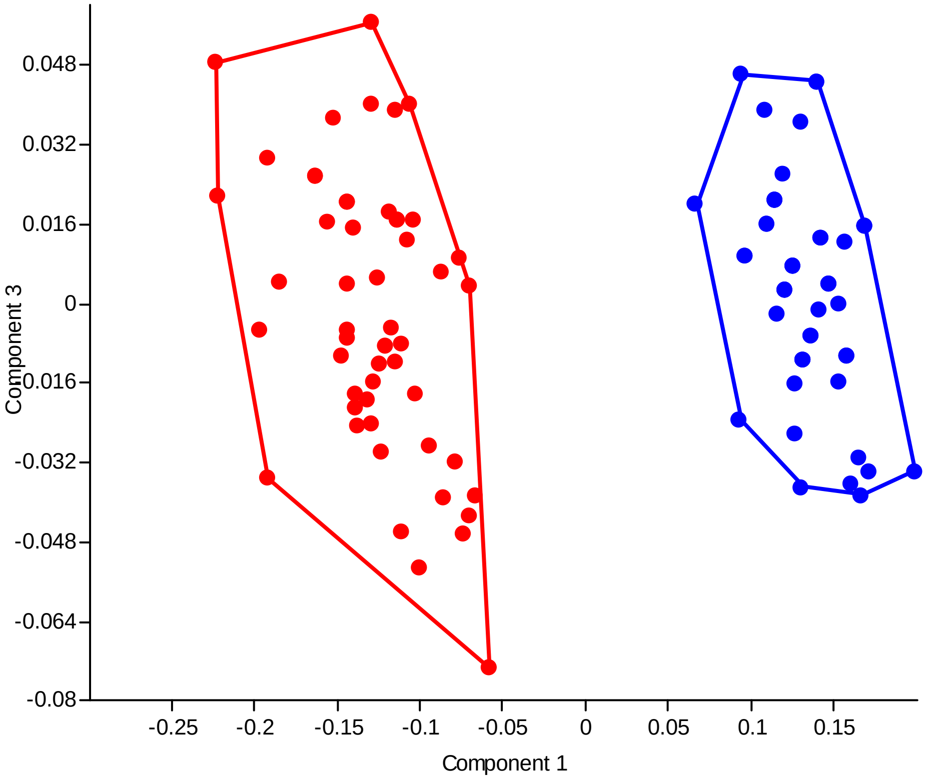

Fig. 3.18 PCA plot comparing meristic variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River system above Augrabies Falls showing differentiation between the two species.

This difference was due to lower scale counts (along the lateral line, between the lateral line to dorsal fin and around the caudal peduncle) in Gouritz and Gamtoos specimens compared to Great Fish, Bushmans and Nahoon specimens. *Labeo umbratus* from the Orange above Augrabies Falls showed a large variation in scale counts. Specimens from the Sundays were intermediate to all the other populations.

When comparing specimens of *L. capensis* and *L. umbratus* from the Orange upstream of Augrabies Falls with specimens from the southern flowing river systems (Fig. 3.20), two significant clusters were apparent. *Labeo capensis* from the Orange above Augrabies Falls clustered separately from *L. umbratus* from the Orange above Augrabies Falls and the southern flowing rivers. This separation was mainly due to the relatively larger scales and resulting lower scale counts for *L. capensis* compared to *L. umbratus*. The lower scale counts of specimens of *L. umbratus* from the Gouritz and Gamtoos (Fig. 3. 19) made the two species clusters appear more similar than it would have had these populations been excluded (Fig. 3.20).

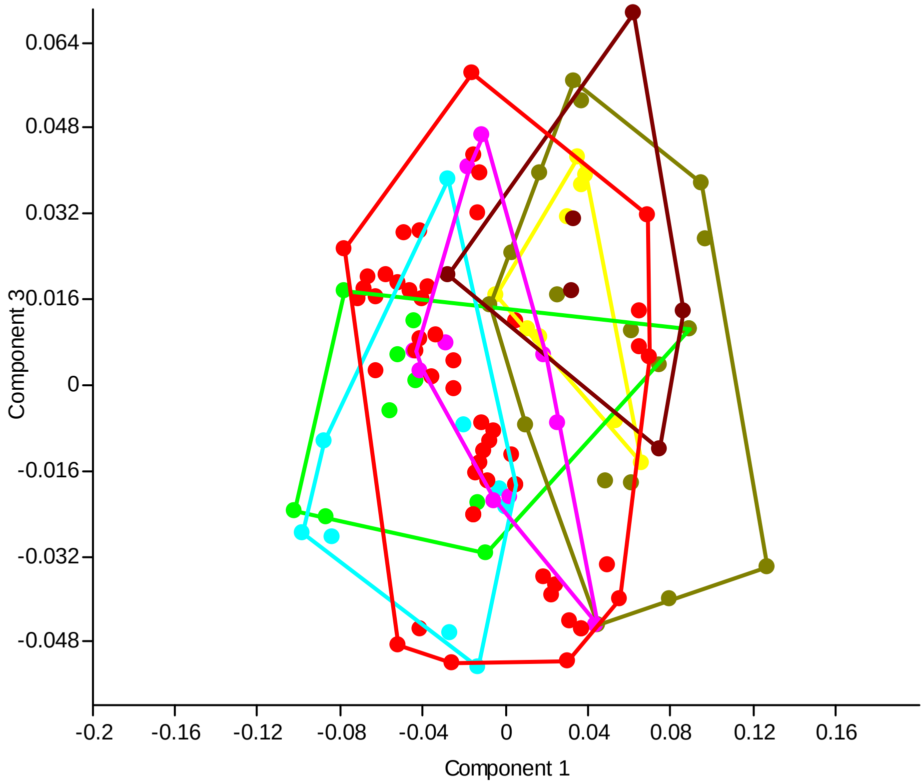

Fig. 3.19 PCA plot comparing meristic variation in characters of *L. umbratus* (**Red**) from the Orange River above Augrabies Falls and *L. umbratus* from southern flowing rivers (Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**)) showing the variation among the populations.

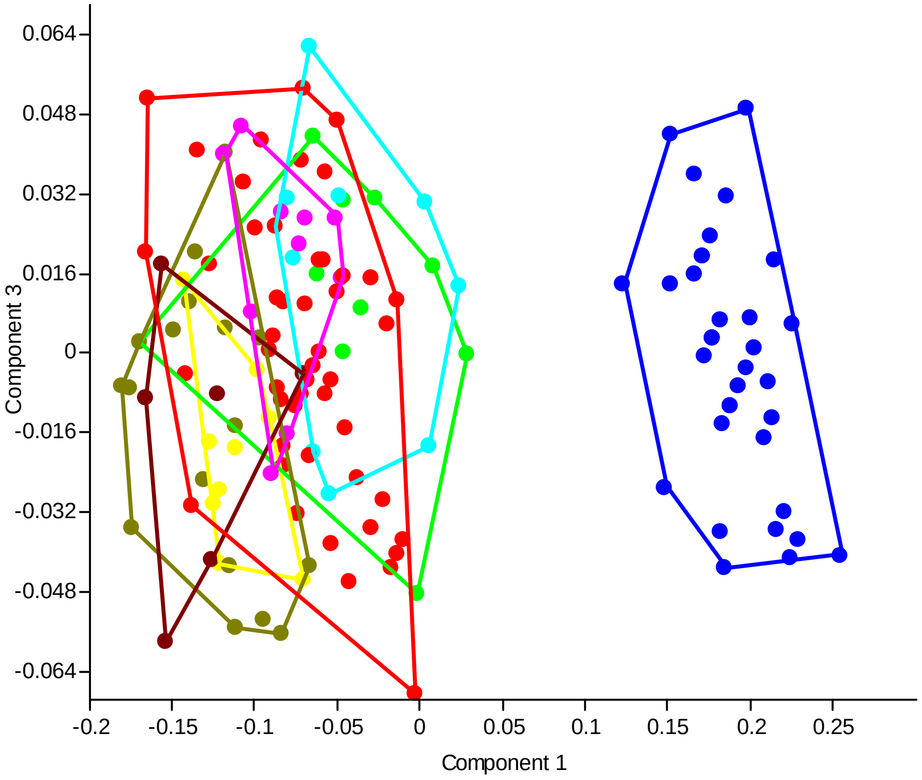

Fig. 3.20 PCA plot comparing meristic variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls and *L. umbratus* from southern flowing river systems, namely the Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**).

Hybrid identification using meristic characters

Similar to the morphometric results, the PCA analysis of meristics characters of all Orange River system specimens revealed two distinct clusters representing *L. capensis* and *L. umbratus* respectively (Fig. 3.21). Most of the specimens (N=29) from Hardap Dam clustered with *L. capensis* with only one clustering with *L. umbratus.* The rest of the specimens (N=14) from Hardap dam were between the two parental species clusters and were therefore classified as potential hybrids.

The PCA of specimens from the Orange above Augrabies Falls, Great Fish and Sundays revealed two distinct clusters assigned to *L. capensis* and *L. umbratus* (Fig. 3.22). Most of the specimens (N=30) from Darlington Dam clustered with *L. umbratus*. Six specimens were intermediate between the two clusters and ten clustered with *L. capensis* (Fig. 3.22). The PCA of the combined samples revealed two clusters representing potentially pure *L. capensis* and *L. umbratus* individuals with hybrids between them (Fig. 3.23)

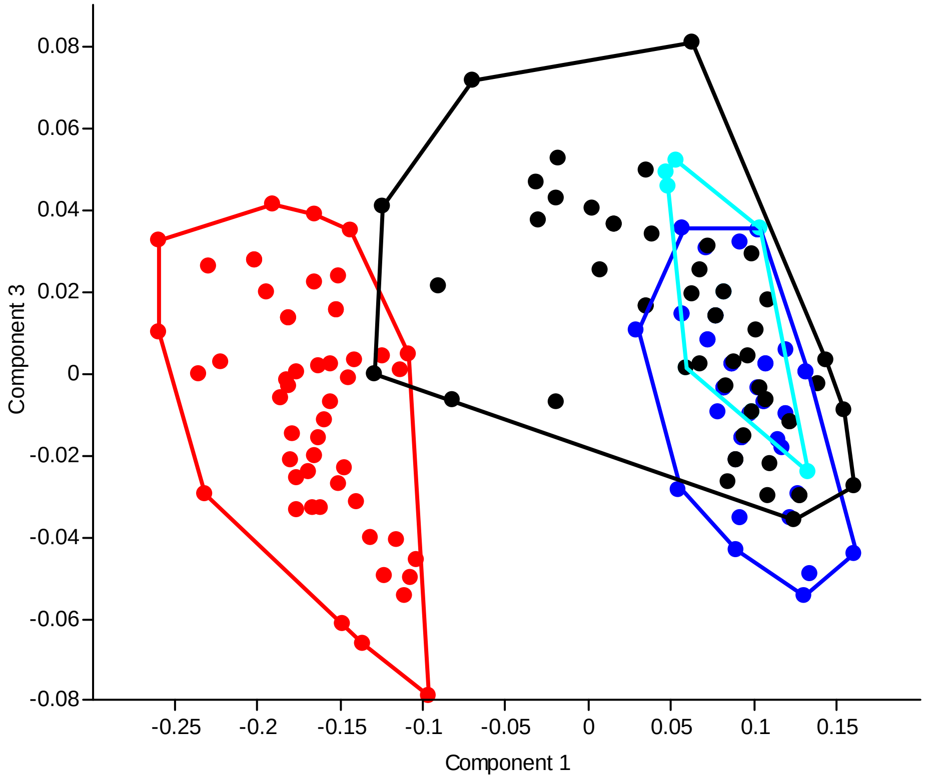

Fig. 3.21 PCA plot comparing meristic variation in characters of *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls, *L. capensis* from the lower Orange River (**sky blue**) and *Labeo* specimens from Hardap Dam (**black**), showing the difference between the two species and potential hybrids between them.

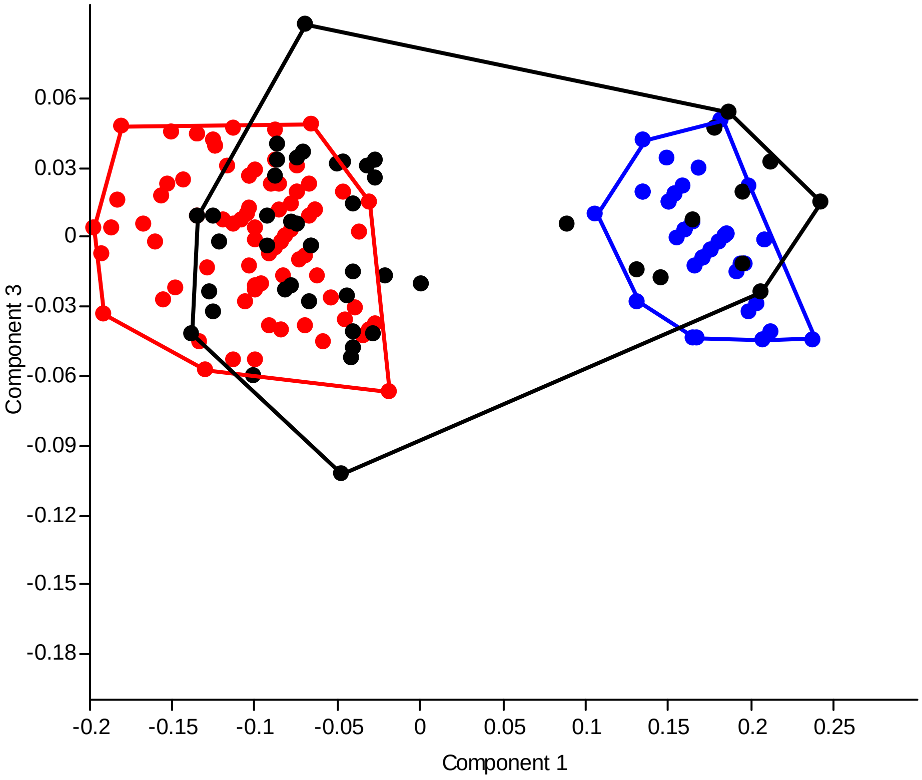

Fig. 3.22 PCA plot comparing meristic variation in characters of *L. capensis* (**blue**) from the Orange River above Augrabies Falls, *L. umbratus* (Orange, Great Fish and Sundays river systems) (**red**) and *Labeo* specimens from Darlington Dam (**black**), showing potential hybrids between the two species.

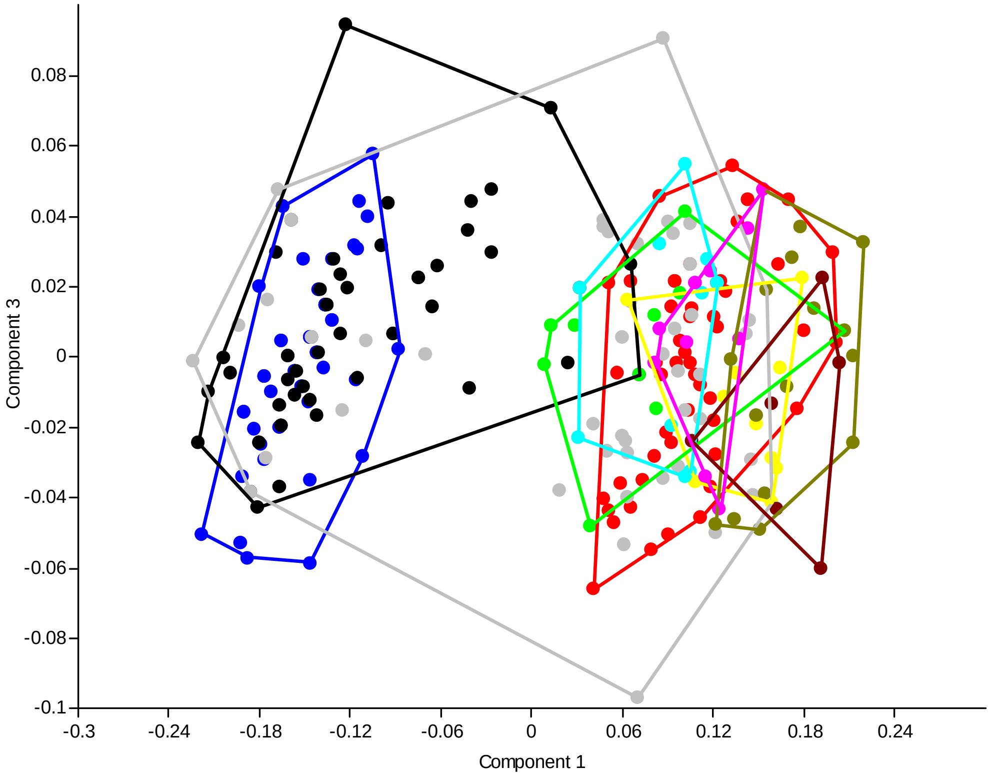

Fig. 3.23 PCA plot for meristic characters showing overall variation between all populations, namely *L. capensis* (**blue**) and *L. umbratus* (**red**) from the Orange River above Augrabies Falls, *L. umbratus* from southern flowing rivers (Great Fish (**olive green**), Sundays (**pink**), Bushmans (**yellow**), Nahoon (**brown**), Gamtoos (**sky blue**) and Gouritz (**green**)), *Labeo* specimens from Darlington Dam (**grey**) and *Labeo* specimens from Hardap Dam (**black**).

3.4 Discussion

Contrary to the conclusion reached by Gaigher and Bloemhof (1975) and Van Vuuren *et al*. (1989), the present study reveals that morphological characters are a useful tool to diagnose possible hybrids. Other studies on the family Cyprinidae have also shown that morphology is a valid tool to identify hybrids (Dowling *et al*., 1984; Freyhof *et al*., 2005; Gante *et al*., 2004; Hayden *et al*., 2010). While previous authors have suggested that these two species could be diagnosed by morphometric and meristic characters (Reid, 1985; Van Vuuren *et al*., 1990; Skelton, 2001; Du Plessis, 1963), the present study is the first comprehensive assessment of morphological diversity within and between these two species across their distributional range with a relatively large sample size.

Pure specimens of *L. capensis* and *L. umbratus* could be diagnosed using four morphometric (dorsal fin base length, inter-orbital width, head length and eye to operculum distance) and four meristic (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) characters. Of these, inter-orbital width and eye to operculum distance were not previously noted as diagnostic characters. According to Van Vuuren *et al*. (1989, 1990), the two species can be distinguished based on dorsal fin spines, but that is not the case as *L. capensis* has 3 spines and *L. umbratus* has 3 or 4 spines.

One of the significant differences between the two species is the increase in the head to body size ratio in *L. umbratus*, which was not observed in *L. capensis*. Gilchrist and Thompson (1913) also noted this, but they suggested that it occurs in both species, suggesting that adults of the two species were macrocephalic. The reason for this is unclear and further research is required to understand this condition. The head to body size ratio was important in distinguishing between specimens from the Orange above Augrabies Falls and those from southern flowing rivers.

The meristic data provided further evidence of structuring in *L. umbratus*. The Gamtoos and Gouritz populations could be distinguished from the Great Fish, Bushmans and Nahoon populations using meristic characters. The two groups of populations had different scale counts with the Sundays being intermediate between them. It therefore seems as if *L. umbratus* from nearby river systems share similar meristic characteristics, suggesting congruence with geographic distributions. This also suggests that the isolation in different river systems is contributing to morphological differentiation. *Labeo umbratus* specimens from the Orange above Augrabies Falls showed a large variation in scale counts, which is probably also a reflection of differentiation across a wide geographical range or large sample size.

There is clear evidence for hybridisation in Hardap Dam. Some specimens were closely associated, but did not completely cluster with suspected pure species from the Orange above Augrabies Falls. These could be due to backcrossing with pure specimens of the parental species or subsequent introgression. Very few pure *L. umbratus* appear to occur in Hardap Dam, since most of the specimens identified in the field as potential *L. umbratus* were classified as potential hybrids after the morphological analysis. It is therefore possible that pure *L. umbratus* is being replaced by hybrids in Hardap Dam. It is also possible that the ongoing hybridisation in Hardap Dam has not spread to the lower Orange River, since all of the lower Orange specimens from Onseepkans clustered with pure *L. capensis*. However, more samples and additional localities will have to be analysed to confirm this, since only ten specimens were collected and analysed. *Labeo umbratus* has never been recorded from below the Augrabies Falls (Hay, 1991; Skelton, 1991; Van Zyl, 1989). This may suggest that *L. umbratus* is unable to establish in the lower Orange, which could also be a reason why hybridisation has not spread from Hardap Dam to this part of the system.

There is also evidence of hybridisation in Darlington Dam from morphometric characterisation. In comparing Darlington Dam specimens to other populations affected by the IBTs, only a few specimens were intermediate between the suspected pure populations of the two species. The low number of hybrids suggests that hybridisation is not yet extensive in this dam. There are very few *L. capensis* in the dam, but their presence confirms that they were able to move from the Orange River into the Great Fish River system and then into Sundays River systems through the IBT system. It seems inevitable that *L. capensis* will establish throughout the Great Fish and Sundays River systems outside areas protected by dams that were built before the IBTs opened. It has been successful in establishing itself well in Gariep Dam (Hamman, 1980) and

Vanderkloof Dam (Gaigher *et al.*, 1981) after impoundment. As a result, the extent of hybridisation in Darlington Dam might therefore increase in future.

In summary external morphological characters can be used successfully to distinguish between *L. capensis* and *L. umbratus* and can provide evidence of hybridisation between the two species. Backcrossing could make it difficult to identify all hybrid individuals, but it is possible that at least all F1 hybrids can be identified with confidence using morphology. Pertinently, the data from Hardap Dam suggests that backcrossing is happening and could be happening extensively. The unique morphological variation detected in *L. umbratus* from southern flowing river systems is being threatened by the introduction of *L. capensis* and *L. umbratus* from the Orange River system via the IBTs.

The null hypotheses that 1) *L. capensis* and *L. umbratus* can not be distinguished using morphology, 2) these two species do not hybridise in Hardap and Darlington dams and 3) morphology can not be used to identify hybrids, can therefore be rejected based on the morphometric and meristic results of this chapter.

Chapter 4: The use of nuclear DNA to differentiate between *Labeo capensis, L. umbratus* and hybrids of the two species

4.1 Introduction

Nuclear DNA is organised within the nucleus of the cell, is responsible for the overall functioning of the cell. Nuclear DNA is diploid, produces heterozygotes, is inherited maternally and paternally, recombines, is affected by natural selection, undergoes insertions and deletions and evolves slower with lower levels of intraspecific divergence compared to mitochondrial DNA (Zhang *et al*., 2003). It can be used to answer questions at different taxonomic levels. Several studies have used nuclear DNA markers to investigate phylogenetics, phylogeography and introgressive hybridisation in freshwater fish (e.g. Alam *et al*., 2002; Clabaut *et al*., 2005; Gante *et al*., 2004; He *et al*. 2008; Lavoué *et al*., 2003; Simonsen *et al*., 2004).

In the past, allozyme electrophoresis was a popular method to investigate population level questions. It has also been used successfully to investigate hybridisation (e.g. Alam *et al*., 2002; Scribner *et al*., 2001; Simonsen *et al*., 2004; Van Vuuren *et al*., 1990; Van Vuuren *et al*., 1989). Van Vuuren *et al*. (1990) and (1989) assessed potential hybridisation between *Labeo capensis* and *L. umbratus* from the Hardap Dam and Vaal Barrage, using 10 allozyme loci. Their study revealed fixed allelic differences at two loci between the two species from the Vaal Barrage where no hybridisation was expected, suggesting that allozyme electrophoresis can be a good tool in distinguishing between the two species. They were able to identify hybrids in Hardap Dam, since hybrid individuals were heterozygous for the loci that had fixed differences between the two species. One of the problems with their studies was that they did not use enough reference pure populations as control representatives of the overall variation within the Orange River system. In addition, it is not clear which criteria they used to identify species and hybrids *a priori* in the field.

In this chapter, the first intron of the S7 ribosomal protein-coding gene was used to assess hybridisation between *L. capensis* and *L. umbratus*. Studies have been conducted on cyprinids (He *et al*., 2008: Xuzhen *et al*., 2002) and other families (Guo *et al*., 2010; Lavoué *et al*., 2003) using this genetic marker. He *et al*. (2008) and Xuzhen *et al*. (2002) conducted studies using the S7 ribosomal protein coding gene to assess phylogenetic relationships of the family Cyprinidae. This marker was successful at resolving relationships even between closely related species. The gene consists of introns and exons in series. Most studies use the first intron as it provides a longer base-pair length and is more variable. Introns are more effective than exons in phylogenetic studies for closely related species, due to their higher variability (Lavoué *et al*., 2003).

The aims of this chapter are therefore to firstly assess whether a nuclear gene intron (such as the first intron of S7) can be useful in distinguishing between *L. capensis* and *L. umbratus*, and secondly, useful in identifying hybrids between them. The following null hypotheses were tested based on the overall thesis hypotheses outlined in Chapter 1:

1) *L. capensis* and *L. umbratus* can not be distinguished using nuclear DNA.
2) *L. umbratus* from different river systems have not been historically isolated.
3) The two species do not hybridise in Hardap and Darlington dams.
4) Nuclear DNA can not be used to identify hybrids.

4.2 Materials and methods

Sampling

Details about the sampling procedures can be found in Chapter 2. Out of a possible 275 samples analysed in Chapter 3, only 77 where used for nuclear DNA analysis (Table 4.1). This was done to save costs and to specifically target the suspected hybrid areas.

DNA extraction

Total genomic DNA was extracted using the protocol of the Promega DNA purification kit (Madison, New York) (see Chapter 2 for further details).

Table 4.1 The number of specimens analysed and 13 localities where they were collected from. Suspected hybrid populations are shown with an asterisk.

Locality	River system	No of specimens
Gariep Dam	Orange	4
Vaal	Orange	4
Brak	Orange	2
Mid Orange	Orange	2
Lower Orange	Orange	2
*Hardap Dam	Orange	26
Kat River Dam	Great Fish	2
Slagboom Dam	Sundays	2
*Darlington Dam	Sundays	25
Bushmans	Bushmans	2
Nahoon	Nahoon	2
Gamtoos	Gamtoos	2
Gouritz	Gouritz	2

Polymerase chain reaction (PCR), purification and sequencing

Extracted genomic DNA was used to amplify the first intron of the nuclear gene coding for the S7 ribosomal protein, using the following primers described by Chow and Hazama (1998): forward primer S7RPEX1F (5' TGG CCT CTT CCT TGG CCG TC 3') and reverse primer S7RPEX3R (5' GCC TTC AGG TCA GAG TTC AT 3'). The PCR was performed in a final volume of 50µl consisting of 6µl of DNA, 5µl of 1x buffer, 5µl of 2mM $MgCl_2$, 5µl of 0.2mM dNTPs, 1µl each 20mM primer, 0.2 of 5u/µl Super-therm DNA polymerase and 26.8µl of double distilled water. The PCR conditions were as follows : initial denaturation 92ºC for 2 min; followed by 35 cycles of denaturation at 92ºC for 60sec, annealing at 58ºC for 60sec and extension at 72ºC for 60sec and finishing with one cycle of extension at 72ºC for 7min.

Genetic analysis

Alignment and editing of S7 sequences were done according to methods described in Chapter 2. The aligned sequences were collapsed to unique nuclear sequence fragments "haplotypes" with the program COLLAPSE (Posada, 2004). A model of nucleotide substitution that best fits the data was selected from 56 models with the Akaike test in MODELTEST version 3.7 (Posada and Crandall, 1998). With the same program, base frequencies, Ti:Tv ratio, 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 neighbour-joining in the program PAUP

(Swofford, 2002) and to create a neighbour-joining tree with *Labeo senegalensis* as an outgroup. Identification was done on an individual bases, firstly with pure populations covering the natural distribution of the two species, and then, with hybrids from suspected hybrids populations.

4.3 Results

Sequence variation

The 77 nuclear sequences were represented by 27 unique nuclear sequence orthologous fragments. The model that best fits the variation between these sequence fragments was F81+I. Gaps were treated as a 5th base. Of the 605 base pairs used, 578 sites were invariable and 24 were variable. Out of the 24 variable sites, 12 were parsimony informative and 12 were autapomorphic. Only the variable sequence base pairs are shown in Table 4.2.

Similar sequence fragments were grouped together to reflect their difference to other groups. Three groups were apparent. Sequences from populations in areas where no hybridisation was suspected, were separated from those that were only found in hybrid areas. As all the sequences from the second group with suspected pure populations were from areas were only *L. umbratus* is found, they were classified as pure *L. umbratus* sequences. The first group was therefore associated with *L. capensis* and the third group was associated with hybrids between the two species as it had alleles from both species (groups one and two).

The C/T (Y) heterozygotes in position 1 for *L. capensis* from suspected hybrid areas, either suggests an extra *L. capensis* allele or a hybrid with an allele that was not detected in other populations. In position 2, pure species are A or G or A/G (R). Hybrids cannot be identified based on this position, because all combinations are possible. In position 3, A/C (M) in both pure and *L. capensis* from suspected hybrids areas and C/T (Y) in pure *L. umbratus* and suspected hybrids, either suggests extra alleles, or hybrids with unknown alleles. There are fixed differences in position 4 between the two species with G in *L. capensis* and T in *L. umbratus*. The hybrids have the heterozygote G/T (K) combination. In positions 5, 17, 19, 21 and 23, gaps in the sequence cause relatively rare alleles. Positions 6 and 8 has the heterozygote A/G (R) in *L. capensis* from suspected hybrid areas either suggesting an extra *L. capensis* allele or hybrids with an unknown allele. In position 7, A/C (M) in pure *L. capensis* suggests an extra *L. capensis* allele.

Table 4.2 A list of variable nuclear sequence fragments showing different base pairs between the *L. capensis* and *L. umbratus*, and hybrids.

Sequence no	Variable sequence fragments	Rivers and Dams
Position	0000 0 00001111111 1 1 1 2 2 2 2 2	
	1234 5 67890123456 7 8 9 0 1 2 3 4	
	Pure *L. capensis*	
Sequence 1	CACG 1 GCATGACTGGG - - - - - - - A	Vaal
Sequence 2	CAMG 1 GCATRWCTGGK 2 T 3 C 4 T 5 A	Mid Orange
Sequence 3	CACG 1 GCATGACTGGG 2 T 3 T 4 T 5 A	Mid Orange
Sequence 4	CAMG 1 GMATGACTGGG 2 T 3 C 4 T 5 A	Vaal
Sequence 5	CACG 1 GCATGWCTGKG 2 T 3 C 4 T 5 A	Gariep
	Labeo capensis from suspected hybrid areas	
Sequence 6	CACG 1 GCATGACTGGG 2 T 3 C 4 T 5 A	Lower Orange
Sequence 7	CAMG 1 RCATGACTGGG 2 T 3 C 4 T 5 A	Lower Orange
Sequence 8	CACG 1 GCATGACTGGG 2 T 3 Y 4 T 5 A	Darlington Dam
Sequence 9	YACG 1 GCATGACTGGG 2 T 3 T 4 T 5 A	Darlington Dam
Sequence 10	CACG 1 GCRTGACTGGG 2 T 3 Y 4 T 5 A	Darlington Dam
Sequence 11	CACG 1 GCATGACYGGG 2 T 3 Y 4 T 5 A	Darlington Dam
Sequence 12	YACG 1 GCRTRWCTGGG 2 T 3 Y 4 W 5 A	Darlington Dam
Sequence 13	CACG 1 GCRTGACTGGG 2 T 3 C 4 T 5 A	Hardap Dam
Sequence 14	CACG 1 GCGTGACTGGG 2 T 3 C 4 T 5 A	Hardap Dam
	Pure *L. umbratus*	
Sequence 15	CGCT 1 GCATGTCTCGG 2 G 3 C 4 T 5 G	Brak
Sequence 16	CRYT 1 GCATGTCTCGG 2 G 3 C 4 T 5 C	Slagboom Dam
Sequence 17	CRYT 1 GCATGTCTCGG 2 G 3 C 4 T 5 S	Slagboom Dam
Sequence 18	CGCT 1 GCATGTCTCGG 2 G 3 C 4 T 5 S	Katriver Dam
Sequence 19	CGCT - GCATGTCTCGG 2 G 3 C 4 T 5 C	Katriver Dam
Sequence 20	CGCT 1 GCATGTCTCGG 2 G 3 C 4 T 5 C	Bushmans
Sequence 21	CGCT 1 GCATGTYTCGG 2 G 3 C 4 T 5 C	Nahoon
Sequence 22	CGCT 1 GCAKGTCTCGG 2 G 3 C 4 T 5 C	Gamtoos
	Labeo umbratus from suspected hybrid areas	
Sequence 23	CGYT 1 GCATGTCTCGG 2 G 3 C 4 T 5 C	Darlington Dam
Sequence 24	CRYT 1 GCAKGTCTCGG 2 G 3 C 4 T 5 C	Darlington Dam
	hybrids	
Sequence 25	CRCK 1 GCAKGWCTSGG 2 K 3 Y 4 T 5 M	Darlington Dam
Sequence 26	CRCK 1 GCATGWCTSGG 2 K 3 C 4 T 5 R	Hardap Dam
Sequence 27	CRCK 1 GCRTGWCTSGG 2 K 3 C 4 T 5 R	Hardap Dam

1 = GCTCTAAGATAGTCTAAAATGCCT

2 = AACGATGATGTTAC

3 = GGTTGA

4 = TCCTTCAG

5 = CTACA

\- = Deletion over corresponding sequence length

In position 9, G/T (K) in pure *L. umbratus* from suspected hybrid areas and hybrid suggests an extra allele. Heterozygous A/G (R) in pure, *L. capensis* from suspected hybrid areas in position 10 suggests an extra allele. In position 11, A and T alleles are present in pure *L. capensis*, but it is interesting that *L. capensis* has mostly A, *L. umbratus* has mostly T and hybrids has mostly W. The heterozygous C/T (Y) in pure *L. umbratus* in position 12 suggests an extra *L. umbratus* allele. In position 13, C/T (Y) in *L. capensis* from suspected hybrid areas either suggests an extra *L. capensis* allele or a hybrid with an unknown allele. Position 14 shows fixed differences between the two species with G in *L. capensis*, C in *L. umbratus* and G/C (S) in hybrids. In position 15 and 16, G/T (K) in pure *L. capensis* suggests an extra *L. capensis* allele. Position 18 shows fixed differences between the two species with T in *L. capensis* and G in *L. umbratus* with G/T (K) in hybrids and a gap causes a relatively rare allele in *L. capensis*. In position 20, homozygous gap, C or T and heterozygous C/T (Y) in pure and *L. capensis* from suspected hybrid areas either suggests extra *L. capensis* alleles or hybrids with unknown alleles. Position 22 has A/T (W) in *L. capensis* from suspected hybrid areas, either suggesting an extra *L. capensis* allele or hybrid with an unknown allele. Position 24 shows fixed differences between the two species with A in *L. capensis*, G and C and G/C (S) in *L. umbratus* with hybrids either A/G (R) or A/C (M).

There were six fix differences between individuals of the two species from suspected pure populations at positions 2, 4, 11, 14, 18 and 24. Of these, only positions 4, 14, 18 and 24 showed fixed differences when the suspected hybrid populations were added and are therefore the only sites that can be used to identify hybrids. Sequences 25-27 were therefore identified as sequence fragments of hybrid origin.

Geographic distribution of nuclear DNA variation

All *L. capensis* from the Orange River above Augrabies Falls (sequence 1-5) grouped together, compared to *L. umbratus* from Orange River above Augrabies Falls (sequence 15) and southern flowing rivers (sequence 16-22) that grouped together as a second lineage (Fig 4.1). There was very low differentiation between *L. umbratus* from the Orange and southern flowing rivers. When suspected hybrid populations were added (Hardap and Darlington dams), some individuals grouped with the *L. capensis* (sequence 6-14) lineage, others grouped with the *L. umbratus* lineage (sequence 23 and 24), while the rest grouped between these two lineages (sequence 25-27) (Fig 4.2). The latter three sequence fragments are therefore identified as hybrids between the two species. Sequences 8-12 were from specimens of *L. capensis* from Darlington Dam, but they grouped with specimens from the Orange River system. Individuals from suspected hybrid populations had more sequence variation. This is because of the low number of individuals sequenced in pure populations compared to the high number of individuals sequenced in suspected hybrid areas and not necessarily due to the effect of hybridisation.

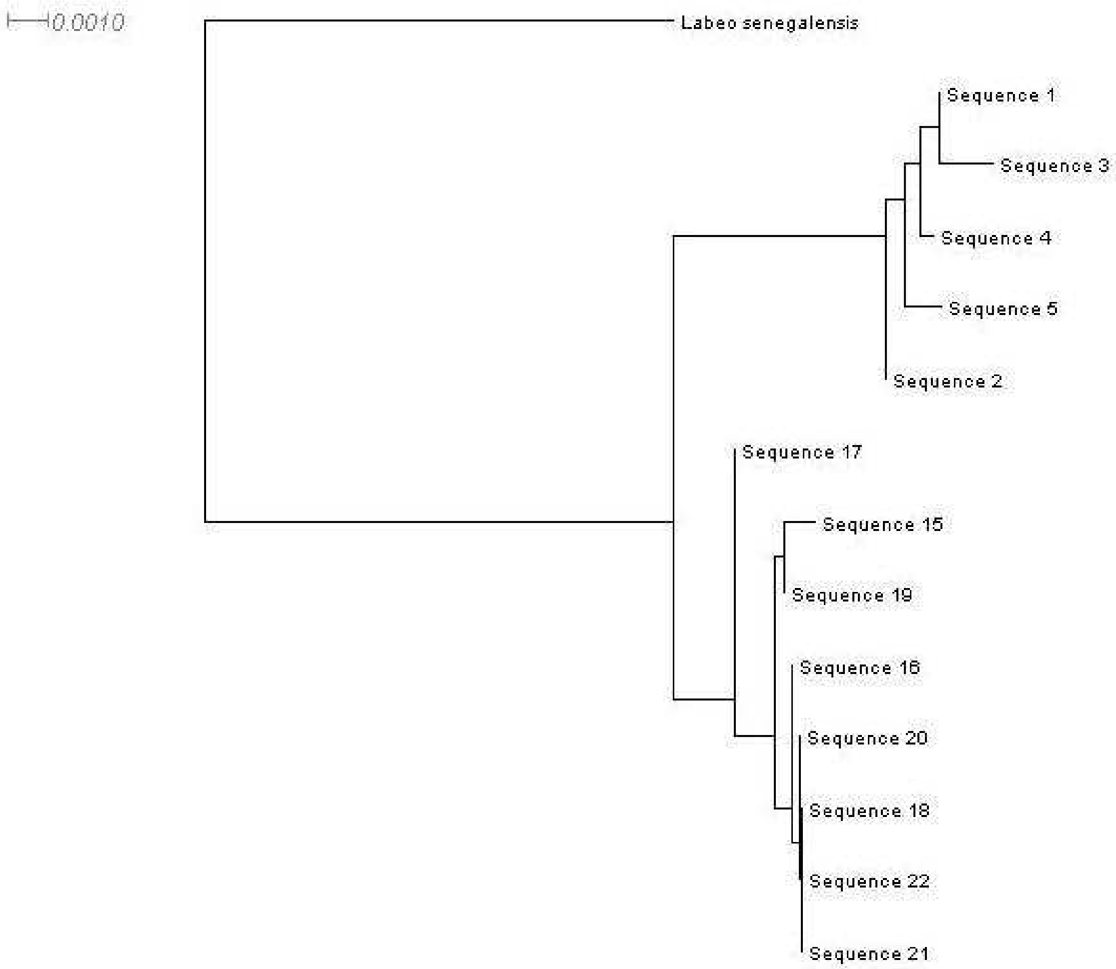

Fig 4.1 A neighbour joining phylogram showing the divergence between *L. capensis* (1-5) and *L. umbratus* (15-22) from suspected pure populations.

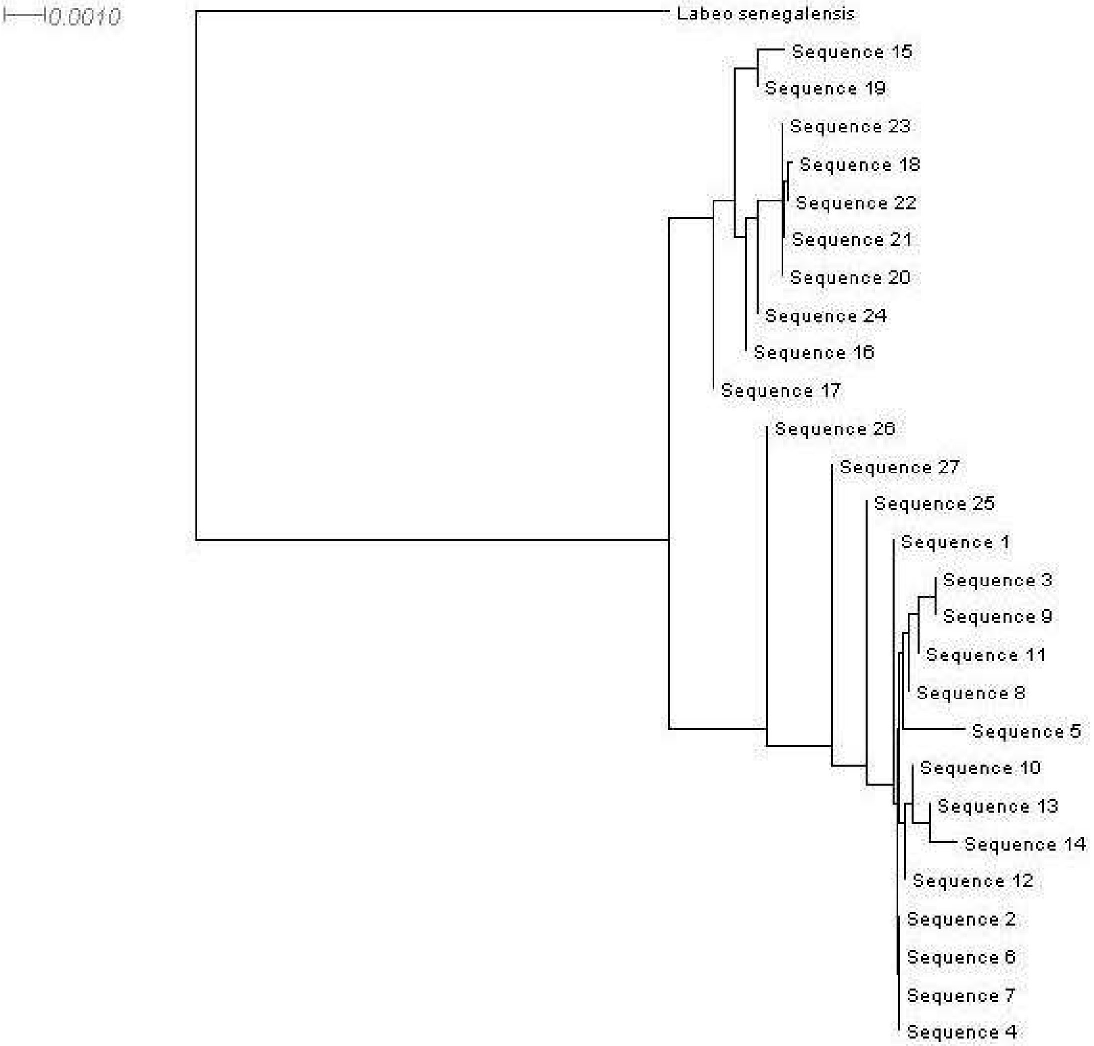

Fig 4.2 A neighbour joining phylogram showing the relationships between suspected pure *L. capensis* (1-5), *L. capensis* from suspected hybrid areas (6-14), suspected pure *L. umbratus* (15-22), *L. umbratus* from suspected hybrid areas (23 and 24) and confirmed hybrids (25-27). Hybrids are placed between the two species lineages due to heterozygote bases at variable sites.

5.4 Discussion

The present study confirms that nuclear markers (in this case specifically the first intron of the S7 ribosomal protein coding gene) can be used to distinguish between *L. capensis* and *L. umbratus*. After all sequences were compared, only three were clearly of hybrid origin containing alleles from both species. As judged by available evidence, all the *L. capensis* sequences from suspected hybrid areas sequences seem to be pure *L. capensis* sequences that were not found before and all the *L. umbratus* sequences from suspected hybrid areas (sequence 23 and 24) seem to be pure *L. umbratus* sequences that were also not found before. That does not mean that the individuals that carry these sequences are pure, because they could be back-crosses or offspring of hybrids that inherited pure genes (Scribner *et al*., 2001).

Nuclear DNA is therefore only useful in identifying hybrids under three different scenarios. The first is when the one parent is a pure *L. capensis* individual and the other a pure *L. umbratus* (F1 hybrids). The second scenario is when one parent is from a pure species and the other a hybrid, where the hybridisation happens in such a way that the hybrid donates an allele of the opposite species (back-cross). For example, if a *L. capensis* parent has an A/A homozygote, a hybrid parent has an A/G heterozygote and the hybrid donates A, then it will not be possible to detect that the offspring is a hybrid. However, if the hybrid donates G, the offspring will be AG and clearly a hybrid. The third scenario is when the parents are two hybrids individuals and where the hybridisation happens in such a way that they donate alleles of opposite species (fertile hybrids breeding with each other). For example, if hybrid 1 with A/G donates A and hybrid 2 with A/G donates G, then the offspring will be A/G and can therefore be identified as a hybrid. However, if alleles are inherited in such a way that the offspring are homozygous for A/A or G/G, it will not be possible to distinguish the offspring from pure specimens.

The two species occur together naturally in the Orange River system where they appear not to hybridise (Tomasson *et al.*, 1984), except in Hardap Dam (Van Vuuren *et al*., 1989 and 1990). Two sequence fragments from Hardap Dam (26 and 27) were heterozygous at the four sites that showed fixed differences between suspected pure specimens of the two species, confirming that they are of hybrid origin. Van Vuuren *et al*. (1989) and (1990) had similar results using allozyme electrophoresis, with heterozygotes in hybrids at loci that showed fixed differences between the two species from the Vaal Barrage.

The presence of *L. capensis* sequences in Darlington Dam is an indication of successful invasion of Orange River *L. capensis* into the Great Fish and the Sundays River systems. Their occurrence in Darlington dam has led to hybridisation between *L. capensis* and *L. umbratus* as confirmed by the presence of a sequence of hybrid origin (Sequence 25).

The null hypotheses that 1) *L. capensis* and *L. umbratus* can not be distinguished using nuclear DNA and 2) the two species do not hybridise in Hardap and Darlington dams can be rejected based on the nuclear DNA results. The nulll hypothesis that nuclear DNA can not be used to identify hybrids, can also be rejected, but nuclear DNA will fail to identify some backcrosses. The use of additional genetic markers and especially more nuclear loci will be necessary to identify more of the backcrossed hybrids. The null hypothesis that *L. umbratus* from different river systems have not been historically isolated could not be rejected using nuclear DNA, due to the low levels of differentiation in the S7 introns among currently isolated river systems.

Chapter 5: The use of mitochondrial DNA to differentiate between *Labeo capensis*, *L. umbratus* and hybrids of the two species

5.1 Introduction

Mitochondrial DNA cannot directly be used to identify hybrids, because it is only maternally inherited, as opposed to nuclear DNA and morphology that is influenced by both parents. However, while the morphological and nuclear analyses of *Labeo capensis* and *L. umbratus* (Chapters 3 and 4) suggested the presence of hybrids in Hardap and Darlington dams, mitochondrial DNA can provide different insights about which specimens have a hybrid history and how the hybridisation occurred when analysed by these other methods (Chapter 6).

The mitochondrion is the organelle in the cell that is responsible for the production of adenosine triphosphate (ATP), which transports chemical energy in cells for metabolism (Moritz *et al*., 1987). It contains circular DNA molecules (mitochondrial DNA), which have 37 genes. Of these, 13 code for proteins involved in oxidative phosphorylation (seven NADH dehydrogenase subunits, three cytochrome oxidase subunits, ATPase 6, ATPase 8 and cytochrome *b*). Twenty-two genes code for transfer ribonucleic acids (tRNAs) and two code for ribosomal RNAs (12S and 16S). There is also a non-coding region called the control region (Wilson *et al*., 1985).

Mitochondrial DNA generally does not recombine, it is mainly maternally inherited, evolves faster compared to nuclear DNA, is small and easy to characterise and there are several copies in each cell (Wilson *et al.*, 1985; Avise *et al.*, 1987, Moritz *et al.*, 1987; Pereira, 2000). Mitochondrial DNA has therefore found application in systematic biology of fish as a tool to investigate population structure (e.g. Swartz *et al*., 2007), population history (Templeton *et al*., 1995), gene flow (Templeton, 1998), hybridisation (Aubert and Solignac, 1990) and phylogenetic relationships (Swartz *et al.*, 2009). The most popular mitochondrial protein-coding gene used in fish systematics has historically been cytochrome *b* (Hebert *et al.*, 2003a), but the recent Fish Barcode of Life project has increased the number of cytochrome oxidase I (COI) sequences available in Genbank (Ward *et al.*, 2009).

According to Esposti *et al*. (1993), cytochrome *b* appeared early in the evolution of organisms because of its presence in nearly all the eukaryotes and prokaryotes and it is one of the best characterised proteins in terms of its structure and function. It is regarded as a useful tool for a variety of systematic questions, because it consists of slow and fast evolving 1st, 2nd and 3rd codon positions. The gene's structure also allows for conservative (e.g. the outer surface of the protein) and variable (inner surface) regions (Lydeard and Roe, 1997). Mitochondrial cytochrome *b* has been widely used to resolve phylogenetic relationships among cyprinids (Machordom and Doadrio, 2001; Bielawski and Gold, 2001; Swartz *et al.*, 2009). Cytochrome oxidase I has similar characteristics to cytochrome *b* and was chosen as a standard for animal barcoding, because it can be amplified in many different taxonomic groups (Hebert *et al.*, 2003b). These genes were both chosen for the present analysis because they are protein-coding genes that are not constrained by insertions and deletions which makes sequence alignment difficult (Doyle *et al.*, 2000). In addition, cytochrome *b* was chosen because it has been used widely in the past and COI was chosen because it is the barcoding standard. Control region usually has more variation, but was not chosen because of undesirable mutation patterns.

Previous studies that assessed hybridisation between *L. capensis* and *L. umbratus*, used morphology (Gaigher and Bloemhof, 1975; Van Vuuren *et al*., 1989 and 1990; Chapter 3) and allozyme electrophoresis analysis as methods (Van Vuuren *et al*., 1989 and 1990). These were useful to distinguish between the two species and the authors were able to identify hybridisation in Hardap dam. The present study found structuring within *L. umbratus* populations using morphology (Chapter 3) and nuclear DNA (Chapter 4). As mitochondrial DNA cytochrome *b* is a fast evolving protein-coding gene, one would expect it to be able to distinguish between the two species. One can also expect to uncover structuring within at least *L. umbratus*, since this species occurs across several currently isolated river systems (see Swartz *et al*., 2007 and 2009).

Since mitochondrial DNA is only maternally inherited, the aim of this chapter was not to identify hybrids, but rather to map the distribution of lineages and to assess which lineages were present in hybrid areas. Only in combination with other methods (Chapter 6), will mitochondrial DNA be useful to identify hybridisation.

The hypotheses that were therefore tested were:

1) *L. capensis* and *L. umbratus* can be distinguished from each other using mitochondrial DNA.

2) *L. umbratus* from different river systems have been historically isolated.

3) Alleles from both species are present in hybrid areas.

5.2 Materials and methods

Table 5.1 Samples of specimens of Labeo from total of 13 localities

Locality	River system	No of specimens
Gariep Dam	Orange	58
Vaal	Orange	20
Brak	Orange	10
Middle Orange	Orange	10
Lower Orange	Orange	10
Hardap Dam	Orange	57
Katriver Dam	Great Fish	16
Slagboom Dam	Sundays	10
Darlington Dam	Sundays	47
Bushmans	Bushmans	10
Nahoon	Nahoon	7
Gamtoos	Gamtoos	10
Gouritz	Gouritz	10

Sampling

Sampling procedures are described in Chapter 2. Out of 389 samples collected, 275 were used for mitochondrial DNA analysis (Table 5.1).

DNA extraction

Total genomic DNA was extracted according to methods described in Chapter 2.

DNA amplification and sequencing

Extracted DNA was used to amplify sections of the COI and cytochrome *b* genes in a polymerase chain reaction (PCR). The following primers were used to amplify the COI gene: VF2_tl forward 5'- TGT AAA ACG ACG GCC AGT CAA CCA ACC AAG ACA TTG GCA C-3' and VR1_tl reverse 5'- CAG GAA ACA GCT ATG CTT CTG GGT GGC CAA AGA ATC A-3'. For cytochrome *b*, two different primer combinations were used. Initially, GluF forward 5'- AAC CAC CGT TGT ATT CAA CTA CAA - 3' with ThrR reverse 5'- ACC TCC GAT CTT CGG ATT ACA AGA CCG - 3' were used and later Gcyt-Glu forward 5'- GAA AAA CCA CCG TTG TTG TTA TTC A - 3' with Gcyt-Thr reverse 5'- CGA CTT CCG GAT TAC AAG ACC - 3' were used when the former primer combination failed to amplify. The PCR was performed in a final volume of 50µl, consisting of 2µl of DNA, 5µl of 1x buffer, 5µl of2mM $MgCl_2$, 5µl of 0.2mM dNTPs, 1µl each 20mM primer, 0.2 of 5u/µl Taq polymerase and 30.8µl of double distilled water.

The PCR conditions were as follows for cytochrome oxidase subunit I: initial denaturation 95ºC for 60sec; followed by 35 cycles of denaturation at 94ºC for 30sec, annealing at 55ºC for 30sec and extension at 72ºC for 60sec and finishing with one cycle of extension at 72ºC for 10min. The conditions for the cytochrome *b* gene was: initial denaturation 94ºC for 2min, 35 cycles of denaturation at 94ºC for 45sec, annealing at 55ºC for 1min and extension at 72ºC for 1min and finishing with extension at 72ºC for 5min.

Genetic analysis

Alignment and editing of sequences were done according to methods described in Chapter 2. A model of nucleotide substitution that best fits the data was selected from 56 models with the Akaike test in MODELTEST version 3.7 (Crandall and Posada, 1998). With the same program, base frequencies, Ti:Tv ratio, 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 neighbour-joining in the program PAUP (Swofford, 2002).

DnaSP 5.10 (Rozas and Librado, 2009) was used to collapse the sequences into alleles and to calculate gene (δ) and nucleotide (π) diversity within and among lineages. With the same program, invariable, variable, parsimony informative and autapomorphic sites were identified. The programs Network 4.5.1.0 (Polzin and Daneschmand, 2003) was used to construct networks and genealogical relationships among the sequences.

Cytochrome *c* oxidase subunit I

In a pilot study, the CO I gene was sequenced for 12 individuals each for *L. umbratus* and *L. capensis* from Gariep Dam, but there was only one mutation found between them in an alignment of 603 base pairs. COI was therefore not considered variable enough for the study and was not further analysed. Cytochrome *b* was more variable than COI in the pilot study, with four mutations between the two species based on 29 samples of each species from Gariep dam in an alignment of 730 base-pairs. It was therefore decided to screen the rest of the samples using only this gene segment.

5.3 Results

Sequence variation and lineage diversity

Of 730 bases analysed, 699 sites were invariable and 31 were variable. Of the 31 variable sites, 19 were parsimony informative and 12 were autapomorphic. Most of the variation was between three unique lineages, one most likely associated with *L. capensis* (hereafter "Orange lineage A") and two most likely associated with *L. umbratus* (hereafter "Orange lineage B" associated with *L. umbratus* from the Orange River system and "southern lineage" associated with *L. umbratus* from southern flowing river systems). There were significant differences in nucleotide and haplotype diversity within populations and lineages (Table 5.2 and 5.3). The southern lineage had the highest nucleotide and haplotype diversity and Orange River lineage B had the lowest.

Table 5.2 Gene (δ) and nucleotide (π) diversity among the different lineages**.**

Lineage	**No sequences**	**No alleles**	**Gene (δ)**	**Nucleotide *π*)**
Orange A	127	15	0.399	0.001
Orange B	53	3	0.075	<0.001
Southern	95	12	0.798	0.003
All individuals	275	30	0.820	0.004

Geographic distribution of alleles and lineages

Distribution of lineages

A total of 30 unique sequences (alleles) were found (see Table 5.4). There were three groups of similar alleles (lineages), representing the two Orange River *Labeo* species (lineages A (alleles 1-15) and B (alleles 16-18)) and populations from southern flowing rivers (lineage C (allele 19-30)) (Fig. 5.1). Lineage C is 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* were found in the middle and lower Orange River therefore Orange lineage A is associated with *L. capensis* and Orange lineage B is therefore associated with *L. umbratus*. Some of the alleles from the Orange River lineage A and B were found in Darlington Dam which is part of the southern flowing rivers (Fig.5.1). The dominant alleles in the dam were part of lineage C. In Hardap Dam it appears as if only Orange lineage A alleles associated with *L. capensis* persisted (Table.5.5).

Table 5.3 Difference in gene (δ) and nucleotide (π) diversity between populations of *L. capensis*, and *L. umbratus*.

Population	**No alleles**	**No sequences**	**Gene (δ)**	**Nucleotide (*π*)**
Orange lineage A (associated with *L. capensis*)				
Gariep	7	29	0.818	0.001
Vaal	5	10	0.756	0.002
Middle Orange	4	10	0.711	0.001
Lower Orange	4	10	0.533	0.001
Orange lineage B (associated with *L. umbratus* from the Orange River system)				
Gariep	2	29	0.069	<0.001
Vaal	2	10	0.200	<0.001
Brak	1	10	0.000	0.000
Southern lineage (associated with *L. umbratus* from southern flowing river systems)				
Katriver Dam	3	10	0.650	0.001
Slagboom	4	10	0.711	0.002
Bushmans	3	10	0.711	0.001
Nahoon	2	7	0.571	0.002
Gamtoos	3	10	0.378	0.001
Gouritz	2	10	0.200	<0.001

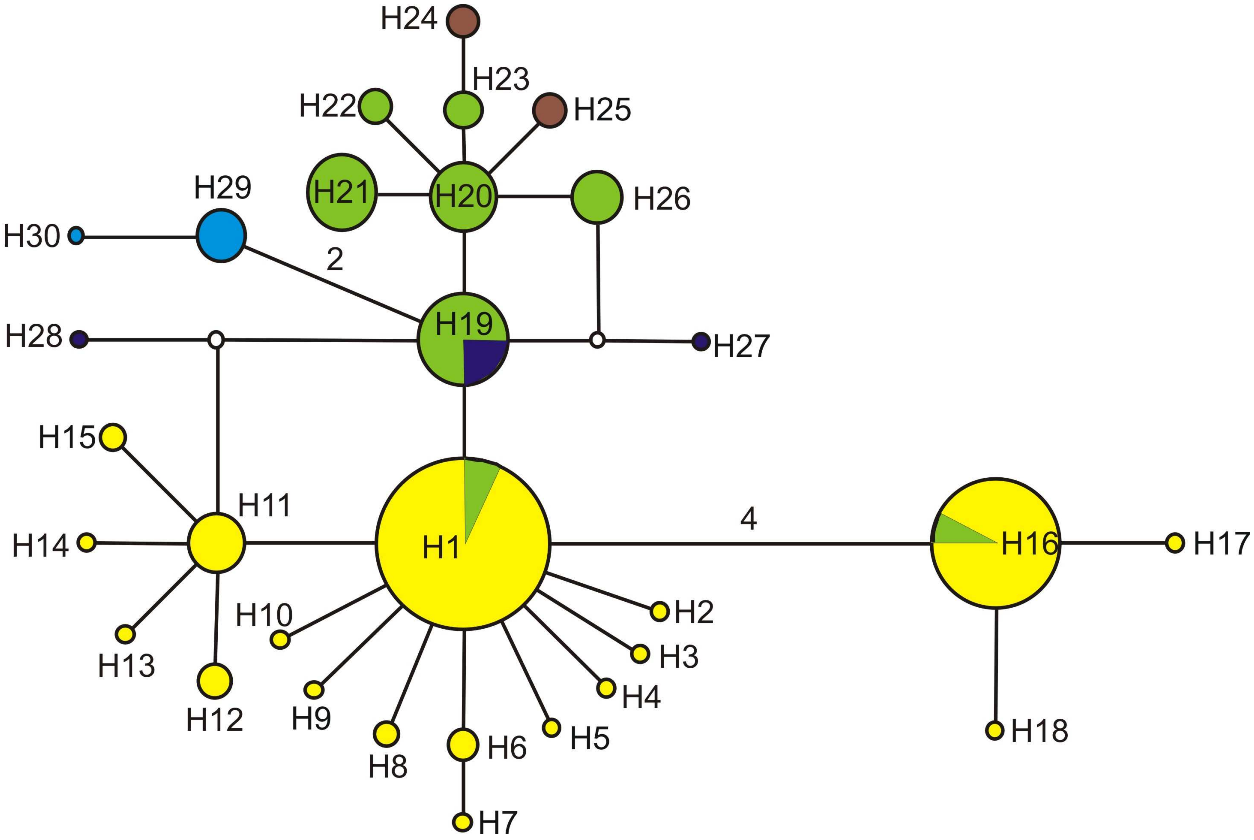

Fig 5.1 A network median joining tree showing where the *Labeo* alleles were recorded. Colour codes: yellow = Orange River system, light green = Great Fish, Sundays and Bushmans river systems; blue = Gamtoos River system; light blue = Gouritz River system; brown = Nahoon River system; white = missing alleles. The number of mutations between alleles is one unless indicated by a number.

Alleles restricted to the Orange River system

Most of the alleles were confined to the Orange River basin, because of the natural occurrence of two species and probably as a result of more sampling locations. Only 16 alleles were found in southern flowing river systems, most of them associated with the *L. umbratus* that naturally occurred there before. Alleles 2-11, 13-14, 17 and 18 were restricted to the Orange River system, because of the historical isolation from southern flowing river systems. Of these, alleles 4-5, 9 and 13 were restricted to lower parts of the Orange River system below the Augrabies Falls. This 56 m knick point (SANParks) is a barrier that prevents fish from migrating upstream.

Alleles from southern flowing river systems

Alleles (19-30) are restricted to southern flowing river systems. Some alleles were shared between southern flowing river systems, namely allele 21 (Sundays and Bushmans), 19 (Sundays and Gamtoos), 26 (Sundays and Great Fish), 20 (Sundays, Great Fish and Bushmans) and 23 (Bushmans and Great Fish) (Table 5.4). Some alleles were restricted to single southern river systems, namely in the Gouritz (24 and 27), Gamtoos (25 and 26) and Nahoon (29 and 30).

Table 5.4 Alleles of *L. capensis* and *L. umbratus*, and their distribution in 13 localities.

Locality	**River system**	**Alleles**
Gariep Dam	Orange	1-3, 6-7, 11 and 16-17
Vaal	Orange	1, 8, 10-12, 16 and 18
Brak	Orange	16
Middle Orange	Orange	1, 11, 14 and 15
Lower Orange	Orange	1, 9, 11 and 13
Hardap Dam	Orange	1, 4 and 5
Kat River Dam	Great Fish	20, 23 and 26
Slagboom Dam	Sundays	19-21 and 26
Darlington Dam	Sundays	1, 12, 15-16 and 19-22
Bushmans	Bushmans	20-21 and 23
Nahoon	Nahoon	24 and 25
Gamtoos	Gamtoos	19, 27 and 28
Gouritz	Gouritz	29 and 30

Alleles that are shared among areas.

Alleles 1, 12 and 15-16 were found in fish from the Orange River system as well as in fish from Darlington Dam in the Sundays River system. Allele 12 and 15 that belong to lineage A, associated with *L. capensis* from the Orange River system, were not found in fish from Gariep Dam which is where the Orange-Fish tunnel starts, but were found in Darlington Dam.

Differentiation between *Labeo* lineages

The Orange River lineage B associated with *Labeo capensis* differs from the Orange lineage A associated with *L. umbratus* by four mutations (0.6% divergence), but from lineage C, associated with *L. umbratus* from southern flowing river systems, by only one mutation (Fig.5.1). Orange lineage B differs from the southern lineage by five mutations (Fig.5.1). The divergence within Orange lineage A (alleles 1-15) was 0.1-0.6%, compared to 0.1-1% within Orange lineage B (alleles 16-18) and 0.1-1.3% within the southern lineage (alleles 19-30) (Table 5.5).

Table 5.5 Data matrix with percentage divergences between alleles of Orange River lineage A associated with *L. capensis* (1-15), Orange River lineage B associated with *L. umbratus* (16-18) and southern flowing rivers lineage C *L. umbratus* (19-30).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	Allele1																														
2	Allele2	0.3																													
3	Allele3	0.1	0.1																												
4	Allele4	0.1	0.4	0.3																											
5	Allele5	0.1	0.4	0.3	0.3																										
6	Allele6	0.1	0.4	0.3	0.3	0.3																									
7	Allele7	0.1	0.4	0.3	0.3	0.3	0.3																								
8	Allele8	0.1	0.4	0.3	0.3	0.3	0.3	0.3																							
9	Allele9	0.1	0.4	0.3	0.3	0.3	0.3	0.3	0.3																						
10	Allele10	0.1	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3																					
11	Allele11	0.1	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3																				
12	Allele12	0.3	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.1																			
13	Allele13	0.3	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.1	0.3																		
14	Allele14	0.3	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.1	0.3	0.3																	
15	Allele15	0.3	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.1	0.3	0.3	0.3																
16	Allele16	0.6	0.9	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.9	0.9	0.9	0.9															
17	Allele17	0.7	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	1.0	1.0	1.0	0.1														
18	Allele18	0.7	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	1.0	1.0	1.0	0.1	0.3													
19	Allele19	0.1	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.7	0.9	0.9												
20	Allele20	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.3	0.4	0.4	0.4	0.4	1.0	1.2	1.2	0.3											
22	Allele22	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6										
21	Allele21	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6	0.6									
23	Allele23	0.6	0.9	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.9	0.9	0.9	0.9	1.2	1.3	1.3	0.4	0.7	0.4	0.1								
24	Allele24	0.3	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.6	0.6	0.6	0.6	0.9	1.0	1.0	0.1	0.4	0.4	0.4	0.6							
25	Allele25	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6	0.3	0.6	0.7	0.1						
26	Allele26	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6	0.6	0.6	0.7	0.1	0.3					
27	Allele27	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.9	0.9	0.3	0.6	0.6	0.6	0.7	0.1	0.3	0.3				
28	Allele28	0.6	0.9	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.9	0.9	0.9	0.9	1.2	1.3	1.3	0.4	0.7	0.7	0.7	0.9	0.3	0.4	0.4	0.4			
29	Allele29	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6	0.6	0.6	0.7	0.1	0.3	0.3	0.3	0.1		
30	Allele30	0.4	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	1.0	1.2	1.2	0.3	0.6	0.6	0.6	0.7	0.1	0.3	0.3	0.3	0.4	0.3	

5.4 Discussion

It is likely that Orange lineage A is associated with *L. capensis* and that Orange lineage B is associated with *L. umbratus*. However for the purposes of this Chapter, it will not be assumed that the alleles associated with these lineages are restricted to a single species. Using cytochrome *b*, it seems as if *L. umbratus* and *L. capensis* can be distinguished in the Orange River system and it is possible that *L. capensis* does not share alleles with the southern lineage of *L. umbratus* either. This will, however, have to be investigated with the combined analysis in Chapter 6.

The low number of mutations between *L. capensis* and *L. umbratus* in cytochrome *b* (one to four mutations among all lineages) and cytochrome oxidase I (one mutations between the two Orange lineages), suggest that speciation between them occurred relatively recently. *Labeo umbratus* can be regarded as a single species, despite the low divergence between Orange lineage B and the southern lineage. It does, however, suggest that there has been isolation between populations occurring in the Orange River system and those in southern flowing rivers. Further investigation needs to be done to assess which evolutionary processes led to the genetic diversity patterns observed and relate the population history to known climatic and geological events.

Surprisingly, the southern lineage of *L. umbratus* appears to be more closely related to the Orange River lineage associated with *L. capensis*, rather than the Orange River lineage associated with *L. umbratus*. The present analysis therefore suggests that *L. umbratus* is polyphyletic with regards to mitochondrial DNA alleles. It is therefore possible that incomplete lineage sorting is playing a role. According to Funk and Omland (2003), this is a process whereby the two daughter species would contain alleles that are similar to each other after speciation, but over time, most of the similar alleles will be lost by genetic drift and new ones will be formed by mutation, until only one ancestral lineage survives. At this point the sorting will be complete and this will ultimately result in a gene tree that reflects the species tree.

Even though it is possible that *L. capensis* and *L. umbratus* do not share alleles anymore (complete lineage sorting), it is possible that the gene tree (Orange lineage A associated with *L. capensis* grouping with the southern lineage associated with *L. umbratus* from southern flowing rivers) does not reflect the species tree (lineages associated with *L. umbratus* should group together), because of the initially incomplete lineage sorting between these three lineages. It does suggest that the divergence between Orange lineage A (probably *L. capensis*), Orange lineage B (probably Orange *L. umbratus*) and the southern lineage (probably southern *L. umbratus*) happened very recently. It might be difficult to assess the sequence of events, because of the similar divergence times between these lineages and the effect of incomplete lineage sorting. The same scenario was observed by Koblmüller *et al.* (2010) who investigated the evolutionary history of the endemic Lake Tanganyika cichlid tribe Tropheini. They found inconsistencies between the phylogenetic reconstructions from the data sets of mitochondrial DNA and nuclear DNA that they attributed to incomplete lineage sorting.

The presence of Orange River alleles in Darlington Dam is an indication of a successful translocation through the IBTs. The mitochondrial DNA results suggest that both *L. umbratus* from the Orange River system and *L. capensis* are now present in Darlington Dam as indicated by the presence of four Orange River alleles from lineage A and B (Table 5.5). This shows that Orange River individuals from both species made it through the Orange-Fish and Cookhouse tunnels into Darlington Dam. Two of the four alleles thought to be of translocation origin that were found in Darlington Dam, were not found in fish from Gariep Dam, which is where Orange-Fish tunnel starts. This could be because they are rare in the Gariep Dam. Apart from hybridisation with *L. capensis*, introgression between the different lineages of *L. umbratus* is very likely and an additional threat to the genetic integrity of *L. umbratus* in the Great Fish and Sundays River systems.

The structuring detected in mitochondrial DNA suggests that the isolation between southern flowing rivers has led to little or no gene flow between the river systems, especially the Nahoon, Gamtoos and Gouritz River systems which had endemic (private) alleles (Fig 5.1). There was sharing of alleles between the Great Fish, Sundays and Bushmans river systems. This shows that there was probably recent gene flow between these river systems, with the Sundays River system being an important link between the southern flowing river systems. River systems that are not part of the IBTs and area above dams that were built before IBTs in affected systems are therefore important for the conservation of the genetic integrity of southern populations of *L. umbratus*.

The null hypotheses that 1) *L. capensis* and *L. umbratus* can not be distinguished from each other using mitochondrial DNA and 2) *L. umbratus* from different river systems have not been historically isolated, can be rejected. The null hypothesis that alleles from both species are not present in hybrid areas can be rejected for Darlington Dam, but not for Hardap Dam where no *L. umbratus* mitochondrial DNA was found.

Chapter 6: Comparison between methods and general discussion

6.1 Introduction

Different data (morphometrics characters, meristics characters, nuclear DNA and mitochondrial DNA) have been used to test for hybridisation separately in Chapters 3-5. Morphology and nuclear DNA techniques can identify hybrids on their own (Chapters 3 and 4). In contrast, mitochondrial DNA (Chapter 5) evidence cannot identify hybrids alone as it is maternally inherited and hybrids will only have one of the parental species' genomes (Rhymer and Simberloff, 1991). It can, however, be useful in identifying hybrids in combination with other methods. Introgression is not always expressed morphologically as some hybrid individuals can express the morphological characters of one of the parental species and some individuals that were identified as hybrids morphologically could appear to be pure genetically depending on how hybridisation and back-crossing occurred (Rhymer and Simberloff, 1996).

For nuclear DNA, two different species that have fixed differences between them will be heterozygous for those fixed differences in hybrids (May, 2003). F1 hybrids will have both parental species genomes represented, but F2 hybrids or backcrosses will have one of the parental genomes or a mixture (Rhymer and Simberloff, 1996). F2 hybrids or backcrosses can therefore not be reliably identified using morphology or a single nuclear marker alone. A combination of methods is therefore better in identifying them.

None of the individuals from the suspected pure populations tested as hybrids with each method. Darlington and Hardap dams therefore remain as the only areas where hybridisation has been detected, but a combination of techniques could identify further hybrid areas.

In this chapter, a final effort to identify hybrid individuals was implemented, using the techniques employed in the present study, to shed more light on how hybridisation has possibly occurred between *Labeo capensis* and *L. umbratus* and to assess the following null hypotheses based on thesis hypotheses outlined in Chapter 1:

1) *L. capensis* and *L. umbratus* can not be distinguished from each other using morphology and genetics.
2) *L. umbratus* from different river systems have not been historically isolated.
3) *L. capensis* and *L. umbratus* do not hybridise in Hardap and Darlington dams.
4) Morphometrics, meristics, nuclear DNA and mitochondrial DNA can not be used to identify hybrids successfully.

6.2 Materials and methods

Results from chapters 3, 4 and 5 were used in this chapter for comparative purposes. As the lowest sample size used were reported in Chapter 4, all the data was reduced to its sample size. Twenty six individuals from suspected pure populations and 51 from suspected hybrid areas (26 from Hardap Dam and 25 from Darlington Dam) were assessed across all methods to determine whether individuals were suspected pure or hybrids (Table 6.1).

Table 6.1 List of the number of specimens analysed and localities where they were collected. Suspected hybrid areas are shown with asterisks.

Locality	River system	No of specimens
Gariep Dam	Orange	4
Vaal River	Orange	4
Brak River	Orange	2
Middle Orange River	Orange	2
Lower Orange River	Orange	2
*Hardap Dam	Orange	26
Kat River Dam	Great Fish	2
Slagboom Dam	Sundays	2
*Darlington Dam	Sundays	25
Bushmans River	Bushmans	2
Nahoon River	Nahoon	2
Gamtoos River	Gamtoos	2
Gouritz River	Gouritz	2

6.2 Results

All individuals from suspected pure populations were identified as either pure *L. capensis* or *L. umbratus* using a combination of all techniques (morphometric characters, meristic characters, nuclear DNA and mitochondrial DNA). The two *L. capensis* populations from the lower (MR08F087 and 088) and middle Orange (MR08F098 and 099) were pure. For *L. umbratus*, the one allopatric population from the Orange River system (Brak River: MR08J021 and 022) and the six allopatric populations from southern flowing river systems (Kat River Dam in the Great Fish River system [MR08D006 and 011]; Slagboom Dam in the Sundays River system [MR08H001-2]; Bushmans River system [MR08G001-2]; Nahoon River system [OW08A025 and OW08A033]; Gamtoos River system [MR08J001-2]; Gouritz River system [AC08A020-21]) were also pure. In addition, no hybridisation were detected in the Vaal and upper Orange, despite the species occurring in sympatry, with pure *L. capensis* from the Gariep Dam (MR08A022 and MR08A054) and Vaal River (MR09A003 and MR09A073) and pure *L. umbratus* from the Gariep Dam (MR08A043 and MR08A057) and Vaal Dam (MR09A080-81) in the Orange River system.

Ten categories of hybrids based on different classification with the different methods tested, were identified (Table 6.2). Hybridisation was detected differently by the different methods between the two dams. Eight categories of hybrids (1, 3-5 and 7-10) were found in Hardap Dam (Table 6.2). All the specimens from Hardap Dam have *L. capensis* mitochondrial DNA. Category 1 (MR08F060) is a possible F1 hybrid as morphology and nuclear DNA identifies it as a hybrid and mitochondrial DNA identifies it as *L. capensis*. Category 3 (MR08F061) is a suspected backcross as it is identified by genetics (nuclear and mitochondrial DNA) as *L. capensis*, by meristics as a hybrid and morphometrics as *L. umbratus*. Category 4 (MR08F044 and MR08F074) is a backcross as it is identified by morphometrics as *L. umbratus*, mitochondrial DNA as *L. capensis* and as a hybrid by meristics and nuclear DNA. Category 5 (MR08F030 and MR08F053) is a suspected backcross as it is identified by morphometrics and nuclear DNA as *L. umbratus*, meristics as a hybrid and by mitochondrial DNA as *L. capensis*. Category 7 (MR08F001, MR08F011, MR08F021, MR08F051, MR08F057 and MR08F063-64) might be a pure *L. capensis* with an intermediate body form or a backcross, because it is classified by only morphometrics as a hybrid. Category 8 (MR08F029) is a backcross as meristics and mitochondrial DNA identifies it as *L. capensis*, with morphometrics and nuclear DNA identifying it as a hybrid. Category 9 (MR08F050) is a backcross as meristics and mitochondrial DNA identified it as *L. capensis*, morphometrics identifies it as a hybrid and nuclear DNA identifies it as *L. umbratus*. Category 10 (MR08F059 and MR08F062) is a backcross as it is identified by morphology (morphometrics and meristics) as a hybrid, nuclear DNA as *L. umbratus* and mitochondrial DNA as *L. capensis*. The rest of the specimens (MR08F002, MR08F004, MR08F010, MR08F012 and MR08F075) where identified as suspected pure *L. capensis* with all methods.

Table 6.2 Summary of the categories of pure (*L. capensis* = CAP and *L. umbratus* = UMB) and hybrid specimens that were identified with a combination of methods (morphology and genetics) in the suspected hybrids areas (Hardap and Darlington dams).

	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 F1 hybrid categories						
1	Hybrid	Hybrid	Hybrid	CAP	Hardap	1
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	UMB	Hybrid	UMB	UMB	Darlington	3
7	Hybrid	CAP	CAP	CAP	Hardap	7
8	Hybrid	CAP	Hybrid	CAP	Hardap	1
9	Hybrid	CAP	UMB	CAP	Hardap	1
10	Hybrid	Hybrid	UMB	CAP	Hardap	2

Only two categories of hybrids (2 and 6) were found in Darlington Dam (Table 6.2). Category 2 (DIFS07_133) is a possible F1 hybrid as morphology and nuclear DNA identifies it as a hybrid and mitochondrial DNA identifies it as *L. umbratus*. Category 6 (MR08F013, MR08F037 and MR08F039) might be a pure *L. umbratus* with an intermediate body form or a backcross as it is only identified as a hybrid with meristic analysis. All the hybrids found in Darlington Dam had *L. umbratus* mitochondrial DNA. Other specimens were classified as pure *L. umbratus* (MR08B003-4, MR08B008-009, MR08B011, MR08B014, MR08B017, MR08B019 and DIFS07_149) or pure *L. capensis* (MR08B022-024 and DIFS07_127, DIFS07_136, DIFS07_141, DIFS07_159) according to all the methods.

6.3 Discussion

A synthesis of different morphological and genetic methods is more effective in identifying pure and hybrid specimens than any of the methods separately. All individuals studied of the two species from suspected pure populations (lower and middle Orange, Brak, Vaal, Bushmans, Nahoon, Gamtoos and Gouritz rivers and Gariep, Kat River and Slagboom dams) were classified as pure according to all methods. The combination of methods was able to identify possible F1 hybrids and backcrosses in Hardap and Darlington dams.

Nuclear DNA and morphological characters identified some specimens as *L. umbratus* (Chapters 3 and 4), but the persistence of only *L. capensis* mitochondrial DNA alleles in Hardap dam is an indication that all the specimens identified as *L. umbratus* has a history of recent or past hybridisation (see also Bernatchez *et al*., 1995, Wilson *et al.*, 1998, Freyhof *et al.*, 2005). It seems as if the hybridisation occurred in a way that has led to the complete fixation of *L. capensis* mitochondrial DNA in Hardap Dam. Mitochondrial DNA replacement has been achieved at a population level in *Drosophila* species through experimental hybridisation in only a few generations (Aubert and Solignac, 1990). Mitochondrial DNA replacement has previously been reported in closely related salmonids with a history of hybridisation. Wilson *et al.* (1998), for example, reported mitochondrial DNA replacement where in the Arctic charr (*Salvelinus alpinus*) mitochondrial DNA became fixed in a southern Quebec population of lake trout (*S. namaycush*). A similar result was reported for brook trout (*S. fontinalis*) in southern Quebec, which is completely introgressed with Arctic charr mitochondrial DNA (Wilson *et al*., 1998). Both these studies found clear morphological differences between the different species, which is not the case in the present study, because of the presence of intermediates (hybrids) based on morphology and nuclear DNA. The presence of a combination of hybrid "types" suggests ongoing introgressive hybridisation (see Miller, 1963). This suggests that hybrids of *L. capensis* and *L. umbratus* are fertile.

Labeo capensis is more abundant than *L. umbratus* in Hardap Dam. The dam has mostly a rocky bottom (Gaigher and Bloemhof, 1975), which is probably why *L. capensis* thrive. It is therefore possible that they spawn everywhere, limiting options for *L. umbratus* to find suitable habitat to spawn on their own, as they prefer muddy habit with grass to spawn on (Mulder, 1973). It is further possible that *L. umbratus* struggled to find mates, since they occur in much lower densities than *L. capensis*. Another explanation may be that pure *L. umbratus* have been extirpated and that only hybrids with some *L. umbratus* characteristics now remain.

Hay (1991) recorded the presence of hybrids in the Fish River (tributary of Orange River) from the upper reaches of the river at Kub above Hardap Dam to below the dam at Tses. Most of the hybrids were caught below the dam. Although no hybrids were detected in the lower Orange River in the present study, securing the genetic integrity of the two *Labeo* species downstream of Hardap Dam and in the lower Orange may be difficult. Populations of the two species above the Augrabies Falls are probably secure, because the waterfall acts as a barrier that will prevent the spread of hybrids upstream.

Labeo umbratus has never been recorded from the lower Orange River (Hay, 1991; Skelton, 1986; Van Zyl, 1989), possibly because the river below the Augrabies Falls has reduced habitat diversity (Skelton, 1986). 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, 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 raises may suggest that *L. umbratus* was introduced into Hardap Dam. The introduction of relatively few fish could have sparked hybridisation with *L. capensis* due to a scarcity of mates from the same species. Dowling *et al.* (1989) stated that a rare species tends to possess more introgressed alleles that the more common ones, which is the situation found in Hardap Dam.

The introduction of *L. capensis* from the Orange River system into the Great Fish and Sundays river systems has led to intraspecific hybridisation with indigenous *L. umbratus* in Darlington Dam. Darlington Dam seems to have relatively new or limited introgression with one out of the four hybrids being potentially an F1 hybrid. This is an indication that there are potentially lower levels of hybridisation in this dam compared to Hardap Dam. Despite sampling effort being biased to potential hybrids and *L. capensis* specimens in Darlington Dam, the mitochondrial DNA analysis revealed Orange River *L. umbratus* alleles in only four individuals (8.5%) and *L. capensis* alleles in only 11 individuals (23%), compared to indigenous *L. umbratus* alleles occurring in 32 individuals (68%). This confirms, however, that fixation of one mitochondrial DNA genome has not occurred as is the case in Hardap Dam.

Darlington Dam hybrids had mitochondrial DNA 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 of *L. capensis* individuals in the dam, which makes it difficult for males of this species to find *L. capensis* females. The main stream rivers of the Great Fish and Sundays River systems have been changed by the water coming from the Orange River system to permanently flowing systems (Laurenson and Hocutt, 1985). This could favour *L. capensis* as it prefers fast flowing waters of the main stream (Skelton, 1986). In future, therefore, more *L. capensis* could invade the Great Fish and Sundays river systems and establish themselves in Darlington Dam.

Interbreeding between *L. umbratus* from the Orange River system and those from southern flowing populations were not found, but this could be due to the lack of a suitable method to identify such individuals, as there was not enough difference in morphology and nuclear DNA between pure populations of these two areas. The only method that could reliably identify the two lineages of *L. umbratus* was mitochondrial DNA, which cannot detect when interbreeding occurred between different lineages unless it can be combined with other methods.

Questions have been raised whether conservation authorities should concern themselves with intraspecific hybridisation, as populations of the same species share a similar genetic makeup. This could increase the fitness of populations by introducing new genetic variation and can have a positive effect on the adaptive potential of a population. This is, however, not true when it comes to populations with more genotypic differences. Local adaptation of such populations could be lost if hybridisation occurs (Allendorf *et al.*, 2001). In the case of *L. umbratus*, more genetic markers will have to be investigated to test the divergence between its two lineages. In the southern flowing 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 (Roux *et al.*, 2002; Laurenson and Hocutt, 1985). Possible interbreeding between *L. umbratus* from the

Orange River and *L. umbratus* from the two southern flowing river systems could negatively impact on local adaptations to respective habitats.

Based on the comparison and combined assessment among methods all of the thesis null hypotheses that 1) *L. capensis* and *L. umbratus* can not be distinguished from each other using morphology and genetics, 2) *L. umbratus* from different river systems have not been historically isolated, .3) *L. capensis* and *L. umbratus* do not hybridise in Hardap and Darlington dams and 4) morphometrics, meristics, nuclear DNA and mitochondrial DNA can not be used to identify hybrids successfully, can be rejected.

6.4 Conclusion and recommendations

Orange River *L. umbratus* have been confirmed to hybridise with *L. capensis* in Hardap Dam as shown in previous studies, but the present study provides the first evidence of hybridisation between *L. umbratus* from southern flowing rivers and *L. capensis*. The hybridisation in Hardap and Darlington dams seems to be at different stages and which may reflect production of different process. There is no overlap in hybrid “types” between the two dams. The introgression in Hardap Dam seems to be distinctly older and/or more extensive than in Darlington Dam, since it was easier to find hybrids during surveys and most of the hybrids are considered backcrosses.

The introduction of *L. capensis* and *L. umbratus* from the Orange River system into the southern flowing Great Fish and Sundays river systems, could lead to the extinction of pure populations of indigenous *L. umbratus* through introgressive hybridisation, similar to what appears to have happened in Hardap Dam. It is also possible that ecological niche competition (space and food) could lead to the decline of indigenous *L. umbratus* populations, due to these introductions. According to the criteria of Allendorf *et al.* (2001), the hybridisation in Hardap Dam rank as type five hybridisation (widespread introgression).

Large populations that seem to be pure representatives of the two *Labeo* species exist above the Augrabies Falls barrier and many dams support potential pure populations. Conserving pure populations of the two species in the Orange River system should therefore not be a major concern for conservation authorities. Darlington Dam seems to be undergoing type four hybridisation (hybridisation without introgression) or the type five hybridisation observed in Hardap Dam, according to the criteria of Allendorf *et al.* (2001). This is because at least one specimen was found in Darlington Dam that could be an F1 or a backcross hybrid. Pure populations of indigenous *L. umbratus* in the Great Fish and Sundays river systems have been identified in the Kat River and Slagboom dams respectively. Individuals from these dams are probably pure, because the dams were built before the IBTs were completed. These populations need to remain secure against invasion if genetically pure and indigenous *L. umbratus* are to survive in these two river systems. Translocations should also be avoided to protect the genetic integrity of other *Labeo* populations across the range of both species.

REFERENCES

Alam, M. A., Akanda, M. S., Khan, M. M. and Alam, M. S. (2002). Comparison of genetic variability between a hatchery and a river population of (*Labeo rohita*) by allozyme electrophoresis. *Pakistan Journal of Biological Sciences* **5**, 959-961.

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

Aubert, J and 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. and 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.

Bielawski, J. P. and Gold, J. R. (2001). Phylogenetic relationships of cyprinid fishes in subgenus *Notropis* inferred from nucleotide sequences of the mitochondrially encoded cytochrome *b* gene. *Copeia* **2001**, 656-667.

Bernatchez, L., Glemet, H. and 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.

Billington, N. (2003). Mitochondrial DNA. In *Population genetics: principles and applications for fisheries scientists* (Hallerman, E. M., ed), pp. 59-100. Bethesda: American Fisheries Society.

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

Cambray, J. A. and 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. (1991). Adaptive significance of a longitudinal migration by juvenile freshwater fish in the Gamtoos River system, South Africa. Suid Afrikaanse Tydskrif vir. Natuurnavorsing **20,** 149-156.

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

Clabaut, C., Salzburger, W. and Meyer, A. (2005). Comparative phylogenetic analyses of the adaptive radiation of Lake Tanganyika cichlid fish: nuclear sequences are less homoplasious but also less informative than mitochondrial DNA. *Journal of Molecular Evolution* **61**, 666-681.

Davies, B. R., Thoms, M. and Meador, M. R. (1992). Viewpoint: an assessment of the ecological impacts of inter-basin water transfer, and their threats to river basin integrity and conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems* **2**, 325-349.

Davies, B. R., O'Keeffe, J. H. and 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.**

Dowling, T. E., Smith, G. R. and 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.

Dowling, T. E. and Moore, W. S. (1984). Level of reproductive isolation between two cyprinid fishes, *Notropis cornutus* and *N. chrysocephalus*. *Copeia* **1984**, 617-628.

Doyle, J. J. and Gaut, B. S. (2000). Evolution of genes and taxa: a primer. *Plant Molecular Biology* **42**, 1-6.

Du Plessis, S. (1963). Revision of the genus *Labeo* in the Transvaal, with the description of a new species. *Annals of the Transvaal Museum* **24**, 327-333.

Ellender, R, E. (2009). The impact of angling on smallmouth and largemouth yellowfish, *Labeobarbus aeneus* and *Labeobarbus kimberleyensis* in Lake Gariep, South Africa. MSc thesis, Rhodes University, Grahamstown, South Africa.

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

Esposti, M. D., Vries, S. D., Criml, M. and Ghelh, A. (1993). Mitochondrial cytochrome b: evolution and structure of the protein. *Btochtmtca et Btophystca Acta (BBA)- Bioenergetics* **1143**, 243-271.

Funk, D. J. and Omland, K. E. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34**, 397-423.

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

Gaigher, I. G., Hamman, K. C. D. and Thorne, S. C. (1981). Relative fish population density changes in the P.K. Le Roux Dam during the first three years after inundation. *Journal of the Limnological Society of Southern Africa* **7**, 37-41

Gante, H. F., Collares-Pereira, M. J. and Coelho, M. M. (2004). Introgressive hybridisation between two Iberian *Chondrostoma* species (Teleostei, Cyprinidae) revisited: new evidence from morphology, mitochondrial DNA, allozymes and NOR-phenotype. *Folia Zoologica* **53**, 423-432.

Gilchrist, J. D. and Thompson, W. W. (1913). The freshwater fishes of South Africa. *Annals of the South African Museum* **11,** 321-463.

Gozlan, R. E., Britton, J. R., Cowx, I. (2010). Current knowledge on non-native freshwater fish introductions. *Journal of Fish Biology* **76**, 751-786.

Guo, X. and 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*, 139-156.

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.

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

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

Hayden, B., Pulcini, D., Kelly-Quinn, M., O'Grady, M., Caffrey, J, McGrath, A. and Mariani, S. (2010). Hybridisation between two cyprinid fishes in a novel habitat: genetics, morphology and life-history traits. *Evolutionary Biology* **10**: 169.

He, S., Mayden, R. L., Wang, X., Wang, W., Tang, K. L., Chen, W. and Chen, Y. (2008). Molecular phylogenetics of the family Cyprinidae (Actinopterygii : Cypriniformes) as evidenced by sequence variation in the first intron of S7 ribosomal protein-coding gene : Further evidence from a nuclear gene of the systematic chaos in the family. *Molecular Phylogenetics and Evolution* **46**, 818-829.

Hebert, P. D. N, Cywinska, A., Ball, S. L. and de Waard, J. R. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society* **B 270**, 313-321.

Hebert, P. D. N, Ratnasingham, S. and deWaard, J. R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society* **B 270**, S96-S99.

Hermida, M., Miguel, E. S., Bouza, C. and Castro, J. (2009). Morphological variation in a secondary contact between divergent lineages of Brown Trout (*Salmo trutta*) from the Iberian Peninsula. *Genetics and Molecular Biology* **32**, 42-50.

Hewitt, G. M. (1996). Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.

Hubbs, L. C. and Lagler, F. K. (1947). *Fishes of the great lakes region*. Michigan: Cranbook Institute of Science.

Hubbs, C. L. (1955). Hybridisation between fishes in nature. *Systematic Zoology* **4**, 1-20.

Jolliffe, I. (1986). *Principal Component Analysis*. New York: Springer-Verlag.

Jubb, R. A. (1967). *Fresh water fishes of Southern Africa*. Cape Town: A. A. Balkema.

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

Lavoué, S., Sullivan, J. P. and Hopkins, C. D. (2003). Phylogenetic utility of the first two introns of the S7 ribosomal protein gene in African electric fishes (Mormyroidea : Teleostei) and congruence with other molecular markers. *Biological Journal of the Linnean Society* **78**, 273-292.

Lydeard, C. and Roe, K. J. (1997). The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among Actinopterygian fishes. In *Molecular Systematics of Fishes* (Kocher, T. D. and Stepien, C. A., eds.), pp. 285–303. SanDiego, CA, Academic Press.

Koblmuller, S., Egger, B., Sturmbauer, C. and Sefc, K. M. (2010). Rapid radiation, ancient incomplete lineage sorting and hybridisation in the endemic Lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution* **55**, 318-334.

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

May, B. (2003). Nuclear DNA. In *Population genetics: principles and applications for fisheries scientists* (Hallerman, E. M., ed), pp. 101-123. Bethesda: American Fisheries Society.

McElroy, D. M. and Douglas, M. E. (1995). Patterns of morphological variation among endangered populations of *Gila robusta* and *Gila cypha* (Teleostei: Cyprinidae) in the upper Colorado River basin. *American Society of Ichthyologists and Herpetologists* **1,** 636-649.

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.

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

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

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

Polzin, T. and Daneschmand, S. V. (2003). On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* **31**, 12-20.

Posada, D. and Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

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

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

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

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

Rhymer, J. M. and Simberloff, D. (1996). Extinction by hybridisation and introgression. *Annual Review of Ecology, Evolution, and Systematics* **27**, 83-109.

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

Simonsen, V., Hansen, M. M., Sarder, M. R. and Alam, M. S. (2004). High level of hybridisation in three species of Indian major carps. *Worldfish Center Quarterly* **27**, 65-69.

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

Skelton, P. H., Tweddle, D. and Jackson, P. B. N. (1991). Cyprinids of Africa. In *Cyprinid Fishes: systematics, biology and exploitation*. (Winfield, I. J. and Nelson, J. S., eds.). pp. 211-239. London: Chapman and Hall.

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

Snaddon, C. D., Davies, B. R. and 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.

Swartz, E. R., Skelton, P. H. and 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. and Bloomer, P. (2009). 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**, 75-84.

Swofford, D. L. (2002). *PAUP* phylogenetic analysis using parsimony (*and other methods), version 4*. Sunderland: Sinauer Associates.

Templeton, A. R. (1998). Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7**, 381-397.

Templeton, A. R., Routman, E. and Phillips, C. A. (1995). Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinun*. *Genetics* **140**, 767-782.

Tomasson, T., Cambray, J. A. and 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.

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

Van Vuuren, N. G., Mulder, P. F., Ferreira, J. T. and 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. and 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.

Ward, R. D., Hanner, R. and Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology* **74,** 329-356.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* **26**, 375-400.

Wilson, C. C. and 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.

Winfield, I., and Nelson, J. (1991). *Cyprinid fishes: Systematic biology and exploitation*. NewYork: Chapman and Hall.

Xuzhen, W., Shunping, H. E. and Yiyu, C. (2002). Sequence variations of the S7 ribosomal protein gene in primitive cyprinid fishes: Implication on phylogenetic analysis. *Chinese Science Bulletin* **47**, 1638-1643.

Zhang, D. and Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations : practice, problems and prospects. *Molecular Ecology* **12**, 563 -584.

Websites

Posada, D. (2004). Collapse ver. 1.2. A tool for collapsing sequences to haplotypes. [Online] Available from http://darwin.uvigo.es.

South African National Committee on Large Dams (SANCOLD, 2009). South African register of large Dams. www.sancold.org.za

South African National Parks. www.sanparks.org/parks/augrabies