The genetic connectivity of *Diplodus capensis* (blacktail) and *Neoscorpis lithophilus* (stonebream) fish populations in the Southwest Indian Ocean

LINDILE VENENCIA CELE

Submitted in fulfilment of the academic requirements for the degree of Master of Science in the School of Life Sciences, College of Agriculture, Engineering and Science at University of KwaZulu-Natal, Westville Campus, South Africa

November, 2016

Supervisors:

Dr Angus H.H. Macdonald

Dr Gavin Gouws

Dr Sean T. Fennessy

ABSTRACT

The oceanography in the Southwest Indian Ocean is complex as the island of Madagascar interrupts the movement of water masses from the equator and results in the formation of eddy currents in the Southwest Indian Ocean. This oceanography could affect the distribution and connectivity of marine species in Mozambique, Madagascar and South Africa. Co-occurrence of many conspecifics has been described for Madagascar and South Africa, and includes the fishes *Diplodus capensis* and *Neoscorpis lithophilus.* The present study was aimed at examining the connectivity of populations of these two species in the Southwest Indian Ocean. This connectivity was investigated using both mitochondrial DNA (mtDNA) markers and microsatellite markers identified and sequenced using Next-Generation sequencing techniques. Three populations for both *D. capensis* and *N. lithophilus* were sampled in Fort Dauphin, Madagascar; Cape Vidal and Durban, South Africa. The mtDNA for both *D. capensis* (n = 24) and *N. lithophilus* (n = 13) illustrated no genetic structuring in the SWIO. This was illustrated by haplotype networks (Figures 3 and 8) and Analyses of Molecular Variance or AMOVA (Tables 2 and 6). Mismatch distributions for both these species showed evidence past demographic expansion (Figures 5 and 9). The microsatellite DNA contradicted the findings of the mtDNA, which present evidence for recent fine-scale structuring. The eighteen microsatellite loci used for *D. capensis* (n = 33) found fine-scale structuring between populations in the SWIO. This was evident in the AMOVA (Table 10) and pairwise F_{st} and G_{st} estimates (Table 11). The ten microsatellite loci used for *N. lithophilus* also illustrated fine-scale structuring between populations in the SWIO. This structuring was evident in the Principle Coordinate Analysis or PCA (Figure 18), the AMOVA (Table 13), the pairwise F_{st} and G_{st} (Table 14). However, recent dispersal between populations was evident for both *D. capensis* and *N. lithophilus.* However, the low percentage of migrants may not be sufficient to allow complete homogenization of allelic frequencies across the metapopulation. This suggests that enough historic dispersal occurred to allow the homogenizing of allele frequencies across the region, however, recent dispersal has become limited, thus resulting in the fine-scale structure between populations.

Keywords: population differentiation; *Diplodus capensis*; *Neoscorpis lithophilus;* mitochondrial DNA, microsatellites

PREFACE

The work described in this dissertation was carried out at the School of Life Sciences, University of KwaZulu-Natal, Westville Campus from January 2014 to November 2016, under the supervision of Drs Angus H.H. Macdonald, Gavin Gouws and Sean T. Fennessy. This dissertation represents the original work of the author and has not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

As the candidate's supervisor I have approved this dissertation for submission.

Signed:  Name: Angus Macdonald Date: 29 Nov 2016

Signed:  Name: Gavin Gouws Date: 11 Nov 2016

Signed: Name: Sean Fennessy Date: 11 November 2016

DECLARATION

PLAGIARISM

I, **Lindile Venencia Cele** declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a. Their words have been re-written but the general information attributed to them has been referenced
   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: 

TABLE OF CONTENTS

Acknowledgement

Firstly, I would like to give a giant thank-you to my three amazing supervisors, Drs A.H.H. Macdonald, G. Gouws and S.T. Fennessy, thank-you for your guidance and help throughout my MSc years. I have learnt a great deal from all three of you. I know that you have put a lot of time and energy into helping me and I am really grateful to have had you as my supervisors.

I would like to thank the members and honorary members of the Marine Evolution (MAREVOL) lab (J. Escobar-Porras, K. Etsebeth, B. Chiazzari, R. van Rooyen, J. Kara, K. Wiggill, C. Barausse and S. Pooran and B.L. Naidoo) for all their help and support in the lab and for pretty much putting up with me and my singing for three years. I have treasured working with you guys. I would like to thank my family and friends for all the emotional support and never letting me slack off.

I would like to thank ACEP III for funding this research and the South African National Department of Agriculture, Forestry and Fisheries (DAFF) for providing personal funding for the first two years of this project. I would also like to thank J. Featherston from the Agricultural Research Council (ARC) for all the help with the Next Generation Sequencing work. I would like to thank the Oceanographic Research Institute (ORI) and R. Jackson for the collection of the fish samples.

Lastly, I would like to thank Jesus Christ, my Lord and Saviour, without whom I would not have made it this far. May this work be worship unto You.

CHAPTER 1 – GENERAL INTRODUCTION

1.1 Connectivity

Population connectivity refers to the successful exchange of individuals (adults, juveniles and larvae), that influences the population demographics between geographically separated populations within a metapopulation (Söndgerath & Schröder 2002, Sale et al. 2005, Cowen & Sponaugle 2009). The rate, scale and spatial structure of this connectivity sustains local populations and thus the metapopulation (Cowen et al. 2006). This aspect of connectivity in the inshore marine fish, *Diplodus capensis* and *Neoscorpis lithophilus*, was the main focus of the present study. In terrestrial environments, the presence of corridors between suitable habitats is important for the connectivity between populations, as individuals are more likely to migrate from one patch to another through these pathways rather than through random migration (Bowne et al. 1999). The distance and the presence of alternate corridors between habitats plays a major role in the connectivity within the terrestrial environment (Bennett 1999). Connectivity in marine environments occurs mostly at a three-dimensional scale, which increases the scale at which ecological processes can occur (Carr et al. 2003). This results in a system that is more open than terrestrial habitats, which makes understanding the connectivity between marine populations more challenging (Carr et al. 2003). This is because a greater understanding of species life-history, oceanography and bathymetry is required, to describe the connectivity between marine populations.

The fragmentation of habitats in the terrestrial environment can be described as the process whereby habitats are altered to become smaller in size and are separated by habitats that are completely different from the original habitat (Wilcove et al. 1986). The areas between these patches may become unsuitable for organisms that previously occupied this region, and as a result the connectivity between these patches becomes limited. Habitat fragmentation within terrestrial environments is of great concern to conservationists as it can result in a decrease in population connectivity by influencing dispersal and immigration rates and thus increasing the risk of local extinction (Hanski et al. 1995, Lindenmayer & Possingham 1996, Söndgerath & Schröder 2002). Populations of the Glanville fritillary butterfly (*Melitaea cinxia*) in Finland declined to the point of local extinction due to habitat fragmentation in this region (Hanski et al. 1995, Fountain et al. 2016). An investigation into the evolutionary history of this species revealed that though selection favoured the genetic traits that allowed better colonization, the populations continued to decline to the point of extinction due to the rapid changes in habitat that resulted in fragmentation (Fountain et al. 2016). Habitat fragmentation in forest habitats has also been shown to affect the reproductive success of migratory birds, which would ultimately affect the population demographics in fragmented regions (Donovan et al. 1995). Reproduction rate has been suggested to increase the rate of connectivity; however high connectivity is suggested to result in low reproduction rates (Söndgerath & Schröder 2002).

Generally, the immigration or emigration that allows population connectivity can either be passive or active. Active migration occurs when individuals actively migrate from one habitat to another. This type of migration includes the seasonal movement of individuals from non-breeding areas to breeding grounds (Webster et al. 2002). Common examples of this type of migration include the seasonal migration of bird species such as the white stork (*Ciconia ciconia*) (Shamoun-Baranes et al. 2003, Berthold et al. 2004); and the willow warbler (*Phylloscopus trochilus*) (Hedenström & Petterson 1987). There are, of course, physiological

and mechanical requirements in order for these individuals to undertake these long migrations. These have been reviewed by Alerstam et al. (2003). This review discusses how individuals require instruction on the seasonality as well as the duration of the migration, the physiological requirements to fuel the migration, the navigation to the final destination, as well as the adaptations required to handle the conditions encountered during and after the migration. Genetic instruction has been suggested to play a role in these mechanisms required to undergo migration (Alerstam et al. 2003). Examples of organisms that undergo physiological changes to migrate, are the Pacific and Atlantic salmon (*Oncorhynchus* sp. and *Salmo salar*, respectively), which migrate from marine to freshwater habitats in order to reach their spawning grounds in North America (Hinch & Bratty 2000, Hodgson & Quinn 2002). These examples are of individuals migrating from one region to another region where locally established populations are absent. This type of connectivity can be referred to as migratory connectivity (Webster et al. 2002, Rubenstein & Hobson 2004, Miller et al. 2012). This differs from the active migration that results in population connectivity where individuals actively migrate from one established local population to another established local population. This means that populations are present in both regions all-year round. Examples of this type of active migration include the migration of freshwater fish species such as *Poecilia reticulate* (guppies) from upriver regions to lower regions of the drainage system in Northern Trinidad, South America (Barson et al. 2009). The movement of sub-adults and juveniles between life-stage dependant habitats is another type of connectivity. Although this connectivity is important in understanding the life-history of species, it was not be the focus of the present study.

Active migration is not exclusive to adult individuals. Many fish species have larvae that are able to control their vertical and horizontal movement in the water column during their late-larval stage (Stobutzki & Bellwood 1998). This adaptation may have developed to assist the fish larvae in recruitment, as they are able to avoid recruiting to unsuitable habitats (Stobutzki & Bellwood 1998, Armsworth 2000). Thus, this ability to actively select recruitment habitats would influence the population connectivity of fish species (Stobutzki & Bellwood 1998, Fisher et al. 2000). This introduces a new aspect to marine larval dispersal that is absent when examining the dispersal of marine invertebrates and may result in evolutionary and ecological differences between them (Armsworth 2000).

Passive migration occurs when individuals are physically transported from one region to another by external forces, such as ocean currents (Shulman & Bermingham 1995, de Queiroz 2005). This passive migration normally occurs during the larval stage of most marine organisms, where the size of the larvae allows their transportation via prevailing oceanic currents. For marine organisms, the pelagic larvae plays an important role in species distribution and connectivity within a metapopulation (Cowen et al. 2006, Cowen & Sponaugle 2009). Local ocean currents have been also been suggested to influence the connectivity of organisms which lack a pelagic larval stage, such as the marine snail (*Solenosteira macrospira*) (Kamel et al. 2014). The passive migration of individuals can also result in the colonization of previously unoccupied regions through founder effects, where areas are colonized by a small number of organisms which become isolated from remnant populations (Barton & Charlesworth 1984). The disconnection from the main metapopulation results in the populations becoming isolated. In terms of population demographics, the population will become self-sustaining with no emigration or immigration of individuals. Genetically, such populations develop fixed allelic differences through genetic drift and/or natural selection, and ultimately develop private alleles (Hellberg et al. 2002). The self-recruitment of local populations due to the lack of gene flow will result in local populations being unique compared to other populations within the metapopulation (Hellberg et al. 2002). For conservation, such populations need to be managed separately from the metapopulation especially if they are located in an economically or geographically important region. For well-connected populations, the dispersal of individuals

between populations over many generations is important for local population persistence (Botsford et al. 2009).

As mentioned previously, most marine organisms have a pelagic larval stage in which the larvae can be transported by currents or other physical features to separate regions (Gilg & Hilbish 2003). This dispersal influences the demographics of local populations (Sale et al. 2005). Pelagic dispersal within a species also affects the gene flow between populations, population sizes, potential barriers to gene flow, the adaptability of species to local conditions, the spatial scale of ecological processes, and the physical distribution of species (Shulman & Bermingham 1995). Thus, describing larval dispersal is important in understanding population connectivity within the marine environment. The duration of the pelagic larval stage was initially theorised to play an important role in the connectivity of marine species over vast geographical regions (Shanks et al. 2003, Cowen & Sponaugle 2009). The notion was that longer pelagic larval durations allowed the ocean currents to transport passive larvae further afield, thus sustaining connectivity at larger (regional) scales resulting in genetically homogenous populations (Bohonak 1999). Authors such as Waples (1987) would use a combination of high fecundity, pelagic larval stage duration and the observations of larvae offshore as indicators of high dispersal ability in marine species. This meant that the connectivity of marine species could be indirectly described by quantifying the pelagic larval duration. This has been shown in *Ctenochaetus* sp. and *Pterocaesio* sp. reef fish in the Great Barrier Reef (Doherty et al. 1995). However, some inconsistencies between the dispersal potential and the actual dispersal of many marine species has brought into question whether pelagic larval duration is solely responsible for the connectivity of populations (Banks et al. 2007). Ocean currents and local oceanography has been shown to influence the dispersal of larvae (Shulman & Bermingham 1995).

Generally, dispersal and movement patterns within a species are measured by tagging and tracking individuals within a population (Mäkinen et al. 2000, Mate et al. 2000, Weersing & Toonen 2007). However, due to the small size of propagules, using the same tracking methods used in larger organisms is physically impossible (Whitlock & McCauley 1999, Gilg & Hilbish 2003, Weersing & Toonen 2007). Genetic traits can be used to indirectly measure population connectivity, as the genetic history of migrants can be investigated (Hellberg et al. 2002). These genetic traits are investigated using genetic markers such as mitochondrial DNA markers, nuclear DNA markers, microsatellites, and Single Nucleotide Polymorphisms (SNPs). Such information can be used to calculate the genetic differences between populations, to determine the haplotypes within each population and their frequency, and to calculate the demographic history of populations as well as current or recent connectivity between populations (Ellegren 2004, Selkoe & Toonen 2006, Galtier et al. 2009, Hellberg 2009). However, population genetics can only describe the effects of gene flow between populations on both an evolutionary and contemporary scale and not the effect of dispersal on population growth rates and other population dynamics (Lowe & Allendorf 2010).

1.2 Motivation of study

Along the south coast of Africa, the warm waters of the Indian Ocean and the cold water of the Atlantic Ocean meet. The occurrence of these two oceans along the coast of southern African has resulted in the formation of nine bioregions, including coastal and offshore regions, in South Africa which are characterized by differences in both biota and physical attributes such as temperature and depth (Lombard 2004, Griffiths et al. 2010). Biogeographic provinces based on temperature have also been defined along the South African coastline which consist of tropical; sub-tropical, warm temperate and cold temperate regions (Teske et al. 2011). The

marine organisms in South Africa show a high level of endemism (Griffiths et al. 2010). The unique oceanography within the Southwest Indian Ocean has resulted in an increased biological variation compared to other west boundary systems (Halo et al. 2014). This may in turn influence the dynamics of coastal marine populations.

The island of Madagascar is located west of Mozambique along the east coast of southern Africa. This island separated from Africa during the spilt of Gondwanaland during the Jurassic period over 170 million years ago (Yoder & Nowak 2006, Gaina et al. 2013, Gibbons et al. 2013). One might expect a marked difference in species composition of inshore marine species between Madagascar and its neighbouring countries: Mozambique and South Africa. This is hypothesised as adult inshore fish are not expected to cross the open Mozambican Channel to reach the island. This, however, has not been observed.

Many inshore fish species are common along both the South African and Madagascan coastal environments and are endemic to the southern African region, such as *Elops machnata* (springer); *Chirodactylus jessicalenorum* (natal fingerfin); *Dichistius multifasciatus* (banded galjeon); *Pomadasys furcatum* (grey grunter); *Johnius dorsalis* (mini-kob); *Diplodus capensis* (blacktail); *Rhabdosargus thorpei* (bigeye stumpnose) and *Neoscorpis lithophilus* (stonebream) (van der Elst 2012, Mann 2013). This would then bring to question the mechanisms through which these species migrated to or from Madagascar and whether this migration of individuals still occurs among present populations. This study was aimed at answering such questions.

1.3 Rationale for species choice

The movement of ocean currents within the Southwest Indian Ocean (SWIO) are complex and differ in trajectory compared to other major western-boundary currents around the world, which will be discussed further in Chapter Two. With the contemporary available information on the trajectory of ocean currents in the SWIO (Lutjeharms & Roberts 1988, de Ruijter et al. 1999_b, Lutjeharms & Machu 2000, Lutjeharms et al. 2000_a, Lutjeharms et al. 2000_b, Lutjeharms et al. 2001, de Ruijter et al. 2002, Machu et al. 2002, de Ruijter et al. 2004, Quartly & Srokosz 2004, de Ruijter et al. 2005, Penven et al. 2006, Morris et al. 2013, Hancke et al. 2014, Roberts et al. 2014), two inshore fish species (*Diplodus capensis* and *Neoscorpis lithophilus*) have been chosen to investigate the connectivity of inshore fish species between South Africa and Madagascar.

Diplodus capensis was chosen due to its wide distribution range which extends from Angola to Mozambique, crossing from the Atlantic Ocean into the Indian Ocean (Summerer et al. 2001). This illustrates its assumed dispersal ability. Also, this species is quite common in recreational catches (Mann 2013), which should increase sample availability. Though the occurrence of *D. capensis* in a wide range of coastal habitats is largely opportunistic, *N. lithophilus* is found mainly in turbulent coastal environments. This allows the examination of a species that is more selective in its distribution.

1.4 Purpose of the present study

The present study was aimed at examining the connectivity of populations of two species of fish (*Diplodus capensis* and *Neoscorpis lithophilus*) in the Southwest Indian Ocean. This connectivity was investigated using the mitochondrial DNA marker, Cytochrome *c* Oxidase subunit I (coxI) and microsatellite markers identified using Next Generation Sequencing technology. These two markers were used to examine the evolutionary history and the

contemporary population structure and connectivity of populations of *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean.

1.5 Thesis outline

The contents of the present thesis are outlined below:

- Chapter One introduces connectivity and how it applies in the marine system. The importance of connectivity and the ways in which connectivity is measured is also discussed within this chapter. The aims, objectives and rationale of the present study are stated.
- Chapter Two discusses the area of study, the Southwest Indian Ocean, in terms of its geography and oceanography.
- Chapter Three discusses reproduction in fish by bringing to light different life-history strategies used in fish species and how they affect connectivity. This chapter also introduces the study species *Diplodus capensis* and *Neoscorpis lithophilus*.
- Chapter Four introduces the use of genetic marker to investigate population connectivity. The materials and methods used in this study are also described in this Chapter.
- Chapter Five presents the results of the genetic analyses used to investigate the evolutionary history of *Diplodus capensis* and *Neoscorpis lithophilus* in the Southwest Indian Ocean using mitochondrial DNA. This includes a discussion of the historic population structure and the demographic history of these populations.
- Chapter Six presents the results of the genetic analyses used to investigate the contemporary population structure of *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean based on microsatellite data. This includes a comparison of genetic structure among species, the possible explanations of the observed population structures and a general conclusion.
- Chapter Seven presents the general conclusions and the recommendations for further research in this study field.

CHAPTER 2 – STUDY AREA: THE SOUTHWEST INDIAN OCEAN

2.1 Geography

The idea of continental drift was first postulated by Wegener (1915), who aimed to explain the current position of the continents of the world today. According to geological evidence, the continents present today were once one large land mass, called Pangea, which comprised of two subregions, named Laurasia and Gondwanaland or Gondwana (Torsvik & Cocks 2013). Gondwana was initially an independent supercontinent about 550 million years ago until about 320 million years ago (Torsvik & Cocks 2013). After this period, Gondwana merged with Laurasia and became the largest land mass within Pangea (Torsvik & Cocks 2013). Gondwana comprised the landmasses known today as South America, Africa, Madagascar, Australasia, India and Antarctica about 100 million km^2 which accounts for 64% of the global landmass present today (Torsvik & Cocks 2013).

Gondwana spilt into eastern Gondwana, consisting of Madagascar, Australia, Seychelles, Antarctica and India, and western Gondwana, consisting of South America and Africa, as a result of rifting of tectonic plates during the Jurassic period, about 170 million years ago (Yoder & Nowak 2006, Gaina et al. 2013, Gibbons et al. 2013). This first exposed the west coast of Madagascar to ocean currents and for 20 million years, coastal marine organisms could have migrated between east and west provinces of Gondwana, allowing for colonization of new regions as well as population connectivity (Yoder & Nowak 2006). The remaining land masses of east Gondwana then also separated from each other and Madagascar reached its current position relative to Africa prior to being separated from India, about 130 to 118 million years ago (Scotese et al. 1988, Yoder & Nowak 2006, Gibbons et al. 2013). The Madagascar-India complex was one of the last landmasses of east Gondwana to separate, about 88 million years ago (Storey et al. 1995, Yoder & Nowak 2006). Madagascar has remained in the same position relative to Africa for the last 118 million years (Yoder & Nowak 2006). Africa and Madagascar then both migrated about 10° north towards the equator between 84 and 66 million years ago (Scotese et al. 1988, Yoder & Nowak 2006, Ali & Aitchison 2008). Madagascar is positioned roughly 400 km east of Africa and has an area of 750 000 km^2 which accounts for about 0.4% of all global landmass (Yoder & Nowak 2006).

2.2 Oceanography

The oceanography of the Southwest Indian Ocean (SWIO) is an interaction of factors. The vast southern Indian Ocean gyre circulation moves warmer water from the equator and removes cool water from lower latitudes. The strongest flow of this gyre system is found east of South Africa where the Agulhas Current is formed (Ali & Huber 2010). However, the trajectory of this movement is highly variable compared to the gyre systems associated with other western-boundary currents of the world (de Ruijter et al. 2005). This is mainly due to the position of the island of Madagascar (Penven et al. 2006).

The South Equatorial Current flows westward towards Africa between 8° S and 22° S carrying warm water towards the coast of Mozambique (Lutjeharms et al. 2000_b). This current meets the north east coast of Madagascar and is forced to split, resulting in a northward (the North Madagascar Current) and southward (the East Madagascar Current) current flow in this

region, see Figure 1 (Lutjeharms et al. 2000_b, Quartly & Srokosz 2004). Water that is forced northwards rejoins the water masses of the South Equatorial Current and eventually meets the African coast (Lutjeharms et al. 2000_b). Part of this current is pushed into the Mozambique Channel while the remaining water is forced northwards, up the coast of Africa (Lutjeharms et al. 2000_b).

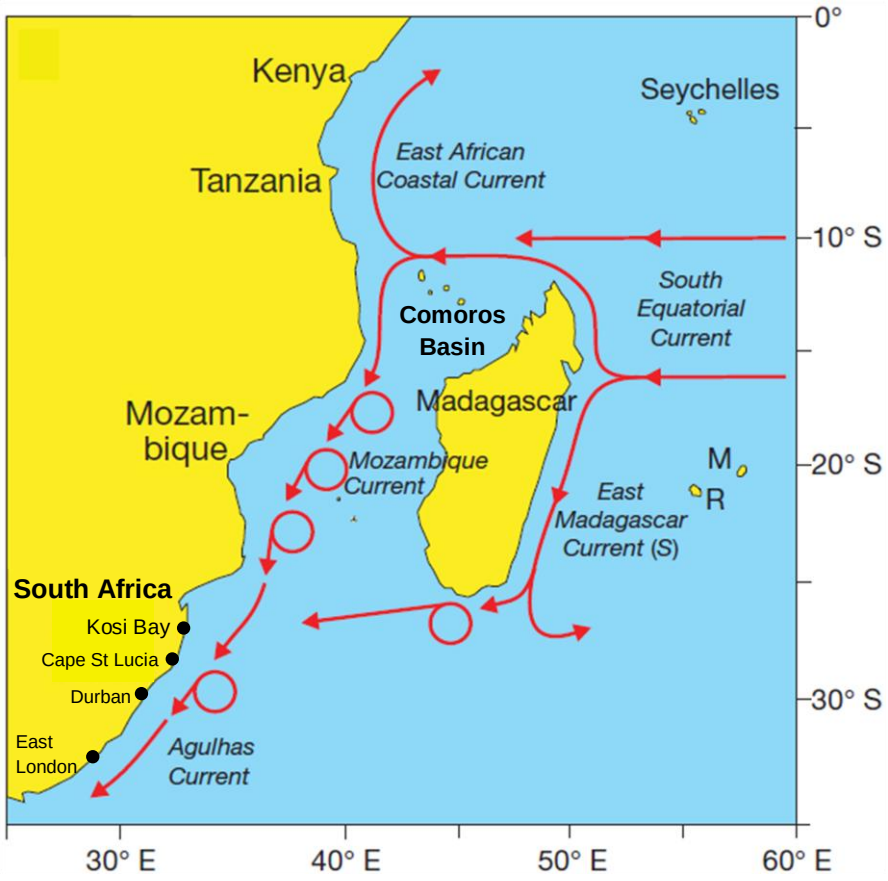

Figure 1: illustration of the oceanography within the Southwest Indian Ocean (adapted from Ali and Huber (2010)).

2.2.1 Mozambique Channel

Water that reaches the Mozambique Channel does not flow like most west boundary currents, instead it forms rings of eddies which travel through the channel (de Ruijter et al. 2002, Ridderinkhof & De Ruijter 2003, Quartly & Srokosz 2004). In the Mozambique Channel, eddies with a diameter of over 300 km and depth of 1500 m, have been shown to transport water southward (de Ruijter et al. 2002). The anti-cyclonic and cyclonic eddies are found moving through the channel on both the western and eastern sides of the Mozambique Channel (Hancke et al. 2014). Mozambique Channel eddies are generated about four times a year, travelling southward at an average speed of 0.052 $m \cdot s^{-1}$ with transport volumes of around 15 x 10^6 $m^3 \cdot s^{-1}$ (de Ruijter et al. 2002). Hancke et al. (2014) presented evidence that cyclonic eddies can retain surface water from the Comoros Basin as they travel southward for a few weeks to months, while surface water in anti-cyclonic eddies tends to move outwards from one eddy to another.

These eddies travel southward along the coast of Mozambique and eventually influence the flow of the Agulhas Current along the South African coast (Quartly & Srokosz 2004).

2.2.2 Eastern Madagascar

The East Madagascar Current is part of the western boundary current in the South Indian Ocean and has an average flow rate of $20.6 \times 10^6 \pm 6 \times 10^6$ $m^3 \cdot s^{-1}$ (de Ruijter et al. 2004). This current flows steadily southward along Madagascar's eastern coastline. Madagascar's continental shelf is narrow from 17°S where it allows the Southeast Madagascar Current to flow close to the coast (Lutjeharms & Machu 2000). The steep and narrow continental shelf of Madagascar has an average width of about 20 km (Lutjeharms & Machu 2000). This continental shelf increases in width at about 26°S and the continental slope becomes more gentle in slope (Lutjeharms & Machu 2000). This sudden change in continental shelf topography results in the frictional drag which creates eddies currents (de Ruijter et al. 2004). Rings of paired eddies (cyclone and anti-cyclone) have been shown to shed off the southern tip of Madagascar and cross the Mozambique Channel at a speed of 0.05 – 0.1 $m \cdot s^{-1}$ (de Ruijter et al. 2004, Morris et al. 2013). Nutrient-rich coastal water from Madagascar is carried in these eddies (de Ruijter et al. 2004). The dipoles or paired eddies are contra-rotating and can account for the irregular behaviour of the anti-cyclonic eddies travelling southwards in the Mozambique Channel (de Ruijter et al. 2004). The cyclonic component of the paired eddies initially decreases in speed as it reaches the African coast and then increases in speed as it moves south along the South African coast (de Ruijter et al. 2004, Morris et al. 2013). de Ruijter et al. (2004) estimated the volume of water added to the Agulhas system by one dipole per year to be about 8×10^6 $m^3 \cdot s^{-1}$. Some of the cyclonic eddies that reach KwaZulu-Natal result in the formation of upwelling cells off Sodwana Bay and south of Durban where prolonged cooling events have been observed (Morris et al. 2013).

2.2.3 Eastern South Africa

The Agulhas Current is formed by sub-gyral circulations in the Southwest Indian Ocean, with influence from both the Mozambique Channel eddies and the southeast Madagascar eddies (Lutjeharms & de Ruijter 1996, Schouten et al. 2002). The Agulhas Current becomes well-developed at 28°S latitude with an average speed of 2 $m \cdot s^{-1}$ and runs to depths greater than 2000 m (Lutjeharms & de Ruijter 1996, Lutjeharms et al. 2001). Continental shelf structure along the South African coast also differs from east to west. The continental shelf along the west coast is wide and has a gradual slope, but the shelf is mostly narrow and steep along the east coast of South Africa (Lutjeharms & de Ruijter 1996). The continental shelf from Kosi Bay to East London has a width of 12 – 25 km, except in the area between Cape St Lucia to Durban where it increases in width to reach a maximum of 50km off the Thukela River region (Lutjeharms & de Ruijter 1996, Lutjeharms et al. 2000_a, Meyer et al. 2002). This 160 km long region along the coast of South Africa is referred to as the Natal Bight (Lutjeharms et al. 2000_a, Meyer et al. 2002). The steep slope of the continental shelf along the east coast of South Africa is responsible for the rapid and relatively stable flow of the Agulhas Current (Lutjeharms & de Ruijter 1996). Instability in the Agulhas Current is observed when the intensity the current's flow exceeds a particular threshold and results in the formation of the Natal Pulse (de Ruijter et al. 1999_b, Lutjeharms et al. 2000_a). The Natal Pulse is a single meander in the Agulhas Current of about 15 km from its normal trajectory which is influenced by the Mozambique Channel eddies as well as the southeast Madagascar eddies (Lutjeharms & de Ruijter 1996, Penven et al. 2006). The Natal Pulse originates within the Natal Bight region and moves down the east coast

at a speed of about 0.21 - 0.23 $m \cdot s^{-1}$ (Lutjeharms & Roberts 1988, de Ruijter et al. 1999_a, Lutjeharms et al. 2001, Schouten et al. 2002).

2.2.4 Transportation of water between regions

Hancke et al. (2014) demonstrated that surface water can be transported between eddies, along their peripheries. These pathways were observed between convergent cyclonic and divergent anti-cyclonic eddies, with the net directional flow of surface water in one direction (Hancke et al. 2014). Such directional pathways were observed in the Mozambique Channel from north to south and from south to north (Hancke et al. 2014). The average time taken to transport surface water through the Mozambique Channel within eddies has been estimated to be 278 days; however, potential transport of surface water via the inter-eddy pathways was estimated to be 51 days for north to south transport and 62 days for south to north transport (Hancke et al. 2014). These authors also presented evidence of the bi-directional movement of inter-eddy surface water across the Mozambique Channel from Madagascar to Mozambique (range of 22 – 210 days) and from Mozambique to Madagascar (range of 15 – 85 days). These inter-eddy pathways could be potential pathways for relatively rapid passive larval transportation. These pathways would allow for the bi-directional movement of larvae across the Mozambique Channel. If this movement were possible, it could also explain the similar coastal species observed in KwaZulu-Natal, South Africa and southern Madagascar. Hancke et al. (2014) used inter-eddy transport to illustrate the connectivity between the surface water from east Madagascar and the Mozambique shelf over 58 – 228 days. Understanding the oceanography in the Southwest Indian Ocean is critical to the understanding of the distribution and connectivity of marine populations in the region (Hancke et al. 2014).

CHAPTER 3 – FISH REPRODUCTION AND BIOLOGY

3.1 Reproductive styles

Among fish species there are two types of reproductive styles; viviparous and oviparous (Hubbs & Lagler 1947, Budker 1971). Research on these types of reproductive styles have been conducted in the last few decades (Hubbs & Lagler 1947, Hoar 1957, Amoroso 1960, Budker 1971, Hogarth et al. 1976, Wourms 1981, Compagno 1988) and form the basis of most of the research on fish reproductive styles. The viviparous reproductive style involves internal fertilization, the development of the embryo within either ovary or the uterus of females, followed by the subsequent birth of live juveniles (Coward et al. 2002, Heemstra & Heemstra 2004). In most viviparous fishes, nourishment is provided to the developing embryo through, first, the egg yolk then a yolk-sac placenta that provides nourishment from the adult female once the egg yolk is depleted (Coward et al. 2002, Heemstra & Heemstra 2004). Most carcharhinoids use this reproductive style (Compagno 1988). Other viviparous fishes do not have a yolk-sac placenta and thus cannot provide extra nourishment to growing juveniles, fishes that use this strategy are referred to as ovoviviparous (Coward et al. 2002, Heemstra & Heemstra 2004). Some teleost fishes use this reproductive strategy (Wootton 1998). Another group of viviparous fishes use an oophagous reproductive style, in which juveniles feed on eggs or other smaller juveniles in the oviduct (Heemstra & Heemstra 2004). However, most teleost fishes use the oviparous reproductive style, which refers to the laying of eggs (Coward et al. 2002, Heemstra & Heemstra 2004) In this strategy, eggs are either deposited on the substrate or released into the water column and the zygote is formed outside of the parents body (Coward et al. 2002, Heemstra & Heemstra 2004). This group of fish can be divided further into two sub-groups: ovuliparous and zygoparous (Coward et al. 2002). Fish that are described as ovuliparous are those that release eggs into the water column, fertilization occurs in the water column and the egg envelope is broken by the juveniles hatching (Coward et al. 2002). Zygoparous fish refer to those that fertilize their eggs internally after which there are kept in the female body for a short period before the fertilized eggs are released into the water column (Coward et al. 2002). In terms of dispersal potential across large geographic regions, larvae of fishes with the oviparous reproductive style are more likely to be dispersed widely, either passively or actively.

In teleost fish, the most common mode of fertilization is external where the gametes are released simultaneously into the water column (Coward et al. 2002). In freshwater systems, some fishes produce demersal eggs with a large yolk sack and sink to the bottom of the water column, so they are generally not carried away by currents (Jones 1968). Other freshwater fishes produce pelagic eggs that have a large oil globule that increases their buoyancy (Jones 1968). This allows the eggs to be dispersed by currents within the environment. The larvae will use this oil globule for buoyancy until the swim-bladder develops (Jones 1968). In marine environments, eggs produced by pelagic fishes are small in size when produced but absorb water from the surrounding environment through osmosis, which allows them to float in the water column which increases their dispersal ability (Jones 1968). The size of demersal eggs produced in the marine environment tends to change with changing latitudes, increasing in size as latitude increases; however, pelagic eggs do not show this pattern (Wootton 1998). Demersal eggs are often adhesive when produced and form a cluster to decrease their changes of being carried away by ocean currents (Lobel & Johannes 1980). Some fishes also create nests, using rubble and sand to prevent the eggs from being taken by the currents (Lobel & Johannes 1980,

Kuwamura 1997). The passive transport of these demersal eggs is, thus, less likely compared to pelagic eggs.

3.2 Hermaphroditism

Hermaphroditism describes the occurrence of both male and female gonads within a single individual (Atz 1965, Sadovy & Shapiro 1987, Puurtinen 2004). In most vertebrates, individuals with this condition are considered abnormal (Buxton & Garratt 1990, Sadovy de Mitcheson & Liu 2008). However, the frequency of this occurrence in fishes has led to the conclusion that hermaphroditism plays a functional role in population dynamics and has thus been referred to as functional hermaphroditism (Buxton & Garratt 1990). Functional hermaphroditism in teleost fish has evolved independently in 27 families (Muraenidae; Gonostomatidae; Cholorophthalmidae; Ipnopidae; Scorpelarchidae; Alepisauridae; Bathysauridae; Rivulidae; Synbranchidae; Caracanthidae; Platycephalidae; Centropomidae; Latidae; Serranidae; Psuedochromidae; Nemipteridae; Lethrinidae; Sparidae; Centracanthidae; Pomacanthidae; Cirrhitidae; Pomacentridae; Labridae; Scaridae; Pinguipedidae; Trichonotidae; Gobiidae) out of a total of 448 families (Sadovy de Mitcheson & Liu 2008, Chopelet et al. 2009_b). In these cases, fish may reproduce as males or females either simultaneously or sequentially (Sadovy & Shapiro 1987, Buxton & Garratt 1990, Sadovy de Mitcheson & Liu 2008).

The question of the evolutionary advantage of hermaphroditism over the gonochoric strategy, where individuals develop into either males or females, is one that has interested biologist since it was first described in the Animal Kingdom. The evolutionary advantage of hermaphroditism has been discussed by authors such as Ghiselin (1969), Nikolsky (1963), Moe (1969), Heath (1977) and Warner (1975) to mention a few. Reproductive styles of many animals are adapted to maximize the probability of contributing to the gene pool of the next generation (Ghiselin 1969). Hermaphroditism has been shown to allow for higher reproductive success per generation compared to that in separate sexes (Ghiselin 1969, Munday & Molony 2002). Ghiselin (1969) reviewed three models which may explain the occurrence of hermaphroditism within species. The first model Ghiselin (1969) describes is called the ‘low density model’ in which hermaphroditism is advantageous because the chances of encountering a suitable individual of the opposite sex within a species are low, such as within sessile organisms and deep-sea fish species. The second model Ghiselin (1969) describes is called the ‘size advantage model’. In this model, a hermaphroditic strategy is favoured within a population in which the early development of individuals as one sex (when individuals are smaller in size) is advantageous but the later development of individuals into another sex (when they are larger in size) is most efficient to contribute to the gene pool of the next generation. The third model Ghiselin (1969) describes is called the ‘gene dispersal model’, which refers to the advantageous development of hermaphroditism under conditions in which gene flow is hindered or where populations are structured. Nevertheless, the evolutionary significance of hermaphroditism is still largely unknown (Puurtinen 2004).

Truly hermaphroditic fishes are those where a significant proportion of the population functions as both male and female individuals either simultaneously or at different stages of their lives (Sadovy & Shapiro 1987). There are several types of hermaphroditic strategies which have been defined and reviewed in previous literature (Sadovy & Shapiro 1987, Buxton & Garratt 1990, Sadovy de Mitcheson & Liu 2008). According to the definition by Sadovy de Mitcheson and Liu (2008), fish can adopt a gonochoristic reproductive strategy, in which

individuals function exclusively as either males or females. The rudimentary hermaphroditic strategy describes fishes whose juveniles have a dormant ovotestis that can develop into either a male or female organ when the fish matures (Richardson 2010, Richardson et al. 2011_b). Fishes that use the protandrous reproductive style include those that have juveniles that first mature as males then they undergo a sex change and develop into functional females (Buxton & Garratt 1990, Sadovy de Mitcheson & Liu 2008). Fishes that mature as functional females first then undergo sex change and become functional males are referred to as using a protogynous reproductive style (Sadovy de Mitcheson & Liu 2008). These are only a few of the reproductive styles defined in literature. Size-at-sex-change has been shown to be highly variable among local populations and over time in fish species such as *Lithognathus mormyrus* and *Chrysoblephus puniceus* (Mariani et al. 2013). This may illustrate that individual size-at-sex-change is highly dependent on the surrounding environment.

Sequential hermaphroditism has been suggested as one of the main factors influencing the genetic structure between geographically separated populations in the marine environment (Chopelet et al. 2009_b). Sequential hermaphroditism results in skewed sex ratios within populations where the sex that individuals first mature as, is favoured (Allsop & West 2004). The gender which individuals mature as first will have lower reproductive fitness than the one the individuals change into (Allsop & West 2004). In order for both males and females to contribute equally to the gene pool of the next generation, the number of individuals of the gender at first maturity must be greater than that of the gender after sex change (Allsop & West 2004). Although the older, larger and more fecund fish that have undergone sex change are generally fewer in number, they contribute to half of the genes inherited in the next generation due to their higher reproductive fitness (Allsop & West 2004, Chopelet et al. 2009_b). Thus, populations of sequentially hermaphroditic fish may become genetically structured over time as a result (Chopelet et al. 2009_b). Though individuals may migrate from one population to another, the ability to contribute to the gene pool of the next generation depends more on the individuals' reproductive success. If these migrant individuals are the individuals that have not yet undergone sex change, then populations may become structured. However, Chopelet et al. (2009_b), illustrated that once dispersal ability is factored in, the remaining genetic variation observed between populations of sequentially hermaphroditic fishes is not explained by sex change. These authors demonstrated that factors affecting the dispersal ability of species have a greater impact on population structure in marine fishes.

3.3 Larval dispersal and behaviour

Larval dispersal is complex and is based on emigration, immigration and the transfer of individuals through a matrix of habitats (Baguette & Van Dyck 2007). Dispersal ability and the factors that influence it, greatly impact the genetic structure of populations in marine systems (Chopelet et al. 2009_b). Understanding larval dispersal of marine organisms requires an understanding of the biological, physical and biophysical factors affecting dispersal (Cowen & Sponaugle 2009). The biological factors include the factors affecting offspring production; growth and survival, the physical factors include properties of water circulation, including advection and diffusion; and the biophysical factors include the interaction of physical factors and larval traits (Cowen & Sponaugle 2009).

Larval fish behaviour has been demonstrated to play an important role in local larval dispersal; however, the extent of this influence on dispersal is not well understood (Leis & Carson-Ewart 1997, Armsworth 2000, Pattrick & Strydom 2009). The larval stage is an important factor in the connectivity of fish populations which influences dispersal over

geographic distance (Pattrick & Strydom 2009). Fish larvae in their early life stages are passive like most marine invertebrate larvae. However, as the fish larvae develop, they gain the ability to control their vertical and horizontal movement in the water column, and are thus able to control their directional dispersal (Pattrick & Strydom 2009). This ability to control their directional dispersal is important in facilitating their ability to recruit to suitable local habitats (Stobutzki & Bellwood 1998, Armsworth 2000).

3.4 Study Species

3.4.1 *Diplodus capensis* (blacktail seabream)

Diplodus capensis (Smith 1844) is a marine fish species belonging to the family Sparidae. *Diplodus capensis* or the blacktail seabream was previously considered a subspecies of *D. sargus* and thus was known as *D. s. capensis* (Domingues et al. 2007). *Diplodus sargus* is a Mediterranean and Atlantic sparid species with six subspecies (Domingues et al. 2007). The name *D. s. capensis* is no longer valid and this species is now considered a separate species under the original name *D. capensis* (www.marinespecies.org accessed 03/02/15; (Heemstra & Heemstra 2004). In this dissertation, this species will be referred to as *D. capensis.*

Research by Summerer et al. (2001) illustrates the most recent and comprehensive phylogeny of the *Diplodus* genus. The genus *Diplodus* originates in the Atlantic Ocean and Mediterranean Sea, from which all other regions were colonized through stepwise dispersal (Summerer et al. 2001). The *D. sargus* lineage was the last to diverge from this group, originating from the East Atlantic (Summerer et al. 2001). A species endemic to Cape Verde, *D. sargus lineatus,* has been suggested to be the most ancestral species within the *D. sargus* lineage (Summerer et al. 2001, Domingues et al. 2007). *Diplodus cervinus hottentotus* or *D. hottentotus* is another species of the *Diplodus* genus that is endemic to the east coast of South Africa (van der Elst 2012). This species diverged from the *D. cervinus* lineage and is more ancestral than *D. capensis,* based on the findings of Summerer et al. (2001). However, it is interesting to see that its most closely-related species is another Atlantic-Mediterranean species *D. c. cervinus* (Summerer et al. 2001). Thus, there are two species from the same genus but of different lineages that are endemic to the same region. This shows the dispersal ability of this genus. One might suggest that this region was colonised by two separate ancestral species at different times or at the same time. Summerer et al. (2001) illustrated that *D. fasciatus* is the ancestral species to the *D. cervinus* lineage and like *D. sargus lineatus,* is an endemic species to Cape Verde. Thus, two ancestral species to two species endemic to southern Africa are found around Cape Verde.

3.4.1.1 Distribution

Diplodus capensis is distributed from Angola to Mozambique (approximately 12°S 13°E - 22°S 38°E), including southern Madagascar (Mann & Buxton 1992, Whitfield 1998, Summerer et al. 2001, Henriques 2012). Though this species has been documented on the southern tip of Madagascar, not much is known about the life history of populations in this region. Literature on *D. capensis* in the Northwest Indian Ocean is relatively lacking, however, this species has been recorded in Pakistan (Siddiqui et al. 2014). Adult *D. capensis* (larger than 150mm Standard Length or SL) are found in the inshore and rocky shore systems, while

juveniles (less than 150mm SL) are found in or near estuaries and in tidal pools (Wallace & van der Elst 1983, Coetzee 1986, Mann & Buxton 1992, Whitfield 1998). *Diplodus capensis* adults have been shown to be highly resident with small home-ranges (Joubert 1981, Attwood & Bennett 1995, Mann 2013).

3.4.1.2 Diet

Diplodus capensis adults are benthic omnivores, with a diet consisting of inshore reef-associated organisms, such as algae and gastropods (Coetzee 1986, Mann & Buxton 1992, Whitfield 1998). The ingestion of algae and their associated epiphytic organisms by *D. capensis* has been shown to be a reliable source of nutrients in the absence of preferred prey, though this consumption has been suggested to be opportunistic (Coetzee 1986, Mann & Buxton 1992). This wide dietary flexibility of *D. capensis* adults has enabled this species to survive in different marine habitats (Mann & Buxton 1992). The diet of juvenile *D. capensis* (less than 150 mm SL) consists mainly of polychaetes, copepods and amphipods (Whitfield 1998). The change in dentition and the elongation of the intestine of *D. capensis* as it develops, allows adults to feed more on the benthos (Mann & Buxton 1992, Whitfield 1998). Similarities between the diets of the *D. capensis* populations along the east and south coast of South Africa have been suggested, though the main species consumed is highly influenced by the geographic area (Coetzee 1986).

3.4.1.3 Reproduction

Diplodus capensis is a slow-growing and long-living species that becomes sexually mature at 4 years and has been found to live for 21 years (Mann & Buxton 1997, 1998). In southern Angola, where this species is heavily exploited, populations undergo an early maturation at 1.6 – 1.8 years, when individuals have a fork length of approximately 150 mm (Richardson 2010, Richardson et al. 2011_a). Early maturity allows the exploited size classes to reproduce before they are removed from the population and thus increases the populations resilience to exploitation (Richardson et al. 2011_a). *Diplodus capensis* populations in Angola were shown to have sexual dimorphism whereby females attain larger body sizes and older ages than males (Richardson et al. 2011_a). The adult sex ratio (males: females) of *D. capensis* populations in southern Angola ranges between 1: 2.2, in exploited areas, and 1: 4.7, in unexploited areas (Richardson 2010). Populations of *D. capensis* in South Africa have shown no evidence of sexual dimorphism, although the larger size classes are dominated by female individuals (Mann & Buxton 1998). The sex ratio of the population of *D. capensis* in the Tsitsikamma National Park in South Africa was found to be 1: 1.98 (Mann & Buxton 1998). In *D. capensis* populations on the coast of KwaZulu-Natal, South Africa, an adult sex ratio of 1.3: 1 has been described (Joubert 1981), though this could have changed in the last 35 years. For both male and female individuals, complete sexual maturity of the entire population occurs when individuals are between 160 – 170 mm in length (Joubert 1981).

The definition of hermaphrodites used by Sadovy and Shapiro (1987) is applied to entire species. However, there are cases where one species has several different hermaphroditic strategies such as in *Boops boops* and *Rhabdosargus globiceps* (Buxton & Garratt 1990). *Diplodus capensis* is also one such case. The Sparidae family is well-known for its wide range of hermaphroditic strategies (Buxton & Garratt 1990). Populations of *D. capensis* in Angola, which are heavily exploited, have been shown to have a rudimentary hermaphroditic reproduction style (Richardson et al. 2011_b), where juvenile fish have a dormant ovotestis that

can develop into either a male or female organ once the fish matures (Buxton & Garratt 1990). However, according to Sadovy and Shapiro (1987) and Sadovy de Mitcheson and Liu (2008), rudimentary hermaphrodites are not true hermaphrodites by their definition. The delayed sex determination displayed by *D. capensis*, is a more advantageous strategy than that of true hermaphroditism, which is associated with high energy costs as well as the loss of mating opportunities during sex changes (Hoffman et al. 1985, Munday & Molony 2002, Richardson 2010). Though the reproduction style of *D. capensis* provides some resilience to the effects of exploitation on the Angolan populations, their slow growth rate and long life span still make this species vulnerable to the effects of fishing (Richardson et al. 2011_a). *Diplodus capensis* populations in South Africa demonstrate a protandrous reproductive style (Coetzee 1986). However, Mann and Buxton (1998) found that a large proportion of the Tsitsikamma National Park population consisted of primary females, i.e. individuals matured as females without passing through a male phase and, thus this population cannot be considered to be truly protandrous. These authors suggested that *D. capensis* shows a digynic reproductive style in which individuals develop into either functional males and females; however, some males have the ability be undergo sex change to develop into functional females. Thus both primary and secondary females are present in these populations (Mann & Buxton 1998, Sadovy de Mitcheson & Liu 2008).

Spawning activity of *D. capensis* adults in South Africa occurs from May to December, but peaks in mid-winter and early spring (July to September) in the subtropical waters of KwaZulu-Natal and in spring to summer (October to December) in warm-temperate waters of the Eastern Cape (Joubert 1981, Coetzee 1986, Mann & Buxton 1998, Whitfield 1998, Beckley 2000, Pattrick & Strydom 2008, Strydom 2008, van der Elst 2012). Research suggests that this spawning pattern still occurs in current populations in KwaZulu-Natal (Dr Allan Connell, unpubl. data http://fisheggs-and-larvae.saiab.ac.za/Default.htm; Accessed 02/11/15). Research regarding the spawning activity of *D. capensis* in Madagascar has not been conducted; thus this information is unavailable.

3.4.1.4 Larval and juvenile development

Diplodus capensis embryos hatch after 50 to 90 hours (Divanach et al. 1982, Potts et al. 2014) and have a planktonic development stage that lasts for 45 days (Attwood & Bennett 1995). *Diplodus capensis* juveniles have been shown to recruit into sub-tidal gullies, intertidal pools, surf zones or estuaries (Whitfield 1998). Although *D. capensis* juveniles are able to survive salinities ranging from 6‰ to 42‰ and are found near estuary mouths (Whitfield 1998), they are not a strongly estuarine-associated species. The use of tidal pools as a nursery area for juvenile *D. capensis* has been shown to be opportunistic, as this species shows increased metabolic rates when exposed to high temperatures and fluctuating salinities (Kemp 2009). Along the Western Cape and Eastern Cape, peak recruitment into estuaries, intertidal pools and surf zones have been documented around summer from October to December (Whitfield 1998). The recruitment of juveniles into tidal pools on Treasure Beach, KwaZulu-Natal occurs from October to April (Beckley 2000). *Diplodus capensis* juveniles have been found to associate with *Zostera capensis* turf in warm-temperate estuaries (Beckley 1983, Whitfield 1998). Unfortunately, such biological information is unavailable for Madagascan populations of *D. capensis*.

Larval swimming speeds of Sparidae have been shown to be similar to those of typical reef fish species (Leis & Carson-Ewart 1997, Fisher et al. 2005, Pattrick & Strydom 2009). Pattrick and Strydom (2009) conducted a study aimed at determining the critical speed and endurance swimming ability of late-stage larvae of two fish species in the family Sparidae (*Diplodus capensis* and *Sarpa salpa*) in South Africa. They showed that these species are strong

swimmers, with an average swimming speed of 0.186 ± 0.077 m. s^{-1}. These Sparidae species, thus, have the ability to swim faster than local currents for short periods of time (Pattrick & Strydom 2009).

3.4.2 *Neoscorpis lithophilus* (stonebream)

Neoscorpis lithophilus (Gilchrist & Thompson 1908) is of the Kyphosidae family. The Kyphosidae family consists of two genera; *Kyphosus* and *Neoscorpis* (Knudsen & Clements 2013). The genus *Kyphosus* consists of 11 species, while *Neoscorpis* is monotypic (Knudsen & Clements 2013, Knudsen & Clements 2016). The physical appearance of *N. lithophilus* is quite similar to species of the Scorpididae family and was originally placed in the *Scorpis* genus by Gilchrist and Thompson (1908). However, Smith (1931) found marked differences in the morphology of this species and placed it in a new genus, *Neoscorpis* (Knudsen & Clements 2016). Other authors such as van der Elst (2012), have classified *N. lithophilus* in the Scorpididae family. However, Nelson (2006) grouped this species with the Kyphosidae family and morphologic and phylogenetic work by Knudsen and Clements (2013), Knudsen and Clements (2016) has since supported this notion.

3.4.2.1 Distribution

This species is an inshore shallow water species, often found in turbulent marine habitats (van der Elst 2012). *Neoscorpis lithophilus* is endemic to the southeast African coast but has also been found along the south east coast of Madagascar (approximately 34°S 20°E - 20°S 35°E) (van der Elst 2012, Mann 2013). In 1975, the fished KwaZulu-Natal populations of *N. lithophilus* were comprised largely of smaller individuals (100 – 160 mm) in the greater part of the year (Joubert 1981). Larger individuals were found along the coast for several months through the year but at a lower frequency compared to the smaller individuals (Joubert 1981). The abundance of the *N. lithophilus* population in KwaZulu-Natal has not been studied; thus, current population dynamics are unknown. *Neoscorpis lithophilus* is a herbivorous fish, with a diet containing mostly red algae and on which they feed constantly (Joubert 1981, van der Elst 2012).

3.4.2.2 Reproduction

An adult male to female sex ratio of 1.2:1 has been described for *N. lithophilus* populations in KwaZulu-Natal, South Africa (Joubert 1981); however, this ratio may have changed in the last 35 years. The size at first maturity for this species is 260 mm fork length in males and 290 mm fork length in females, at about 3 to 4 years of age (Joubert 1981, Mann et al. 2002). No differences in age at sexual maturity between populations have been observed in this species. The reproductive style of *N. lithophilus* has been described as gonochoristic, which is the development of individuals into exclusively mature males or females, regardless of their gonadal morphology; thus this species does not undergo sex change (Sadovy de Mitcheson & Liu 2008, Mann 2013). *Neoscorpis lithophilus* has been suggested to spawn in mid-winter to late summer (from July to January) in southern Mozambique and northern KwaZulu-Natal (Joubert 1981, Mann et al. 2002). The larval development, including late-larval swimming ability of *N. lithophilus* is largely unknown. The biological aspects of a species are important in understanding their connectivity. These factors such as distribution, reproduction, larval dispersal and behaviour influence the species ability to successfully disperse from one population to the next.

CHAPTER 4 – MATERIALS AND METHODS

4.1 Molecular Markers

4.1.1 Mitochondrial DNA

The mitochondrial genome is a circular molecule about 17 000 base pairs long and is found in both animals and plants (Avise 2009, Lowe et al. 2009). This molecule contains 37 genes in animal species and lacks introns and DNA repeats (Avise 2009, Galtier et al. 2009). Mitochondrial DNA (mtDNA) is maternally inherited and is fast evolving in most animals (Hurst & Jiggins 2005, Ward et al. 2005, Galtier et al. 2009). When mtDNA and nuclear DNA (nDNA) are compared in animals, mtDNA is found to have a similar or greater nucleotide substitution rate (1-10 times) than nDNA (Avise et al. 1987, Lowe et al. 2009). Certain nucleotide positions within the mitochondrial gene are more evolutionarily important than others and these positions reveal the rapid mutation rate of parts of the mtDNA (Avise et al. 1987). Other mutations within the mtDNA occur at a slower rate and thus can be used to estimate phylogenetic relationships between closely-related species (Avise et al. 1987).

Mitochondrial DNA markers are the most used genetic markers in animal species for examining the evolutionary history of species over millions of years and the phylogenetic relationships between species (Hewitt 2004, Hurst & Jiggins 2005). The maternal inheritance of mitochondrial DNA can result in uncertainties in species identification when dealing with hybrid species, where only the DNA of the maternal species can be examined (Ward et al. 2005). This maternal inheritance of mtDNA also means that only the evolutionary history of the female can be examined and not of the species as a whole (Hurst & Jiggins 2005). Another issue with mtDNA is its integration into the nuclear genome in some species, which can result in mtDNA markers amplifying unwanted regions of the nuclear genome (Bensasson et al. 2001, Galtier et al. 2009). The cytochrome oxidase subunit I gene within the mitochondrial DNA is used within the global bioidentification system to identify, delineate and barcode animal species (Hebert et al. 2003, Ward et al. 2005). The mtDNA marker that was used in this study is cytochrome *c* oxidase subunit I (coxI) using primers designed by Ward et al. (2005) for use as universal primers that amplify DNA from a wide range of fish species.

4.1.2 Microsatellites

Microsatellites or Single Sequence Repeats (SSRs) are short tandem repeats of DNA sequences found in both coding and non-coding regions of the whole genome (Tóth et al. 2000, Avise 2004, Selkoe & Toonen 2006). These repeats include single- (mono-), double- (di-), triple- (tri-), quadruple- (tetra-) nucleotide repeats, to name a few (Ellegren 2004). These DNA regions are highly variable or polymorphic, occur in high abundance in the genome and are inherited from both parents (Chistiakov et al. 2006, Plough & Marko 2013). This makes microsatellites suitable genetic markers to use in both population and evolutionary studies (Plough & Marko 2013). Microsatellites in the non-coding regions of DNA are assumed to evolve neutrally; thus mutation processes should be reflected by their allelic frequency and distribution (Ellegren 2004). The short sequence length of these regions also means they can be amplified relatively easily (Chistiakov et al. 2006). These short sequences also have high mutation rates (about 10^{-2} to 10^{-6} per locus per generation), and thus can be used determine

present or recent fine-scale population structure (Ellegren 2000, Selkoe & Toonen 2006, Sala-Bozano et al. 2009, White et al. 2010, Henriques et al. 2014). This means recent fine-scale population structure refers to the population structure of present day populations. This high mutation rate also makes microsatellites important in genome evolution and they can thus be used to investigate this process (Tóth et al. 2000, Plough & Marko 2013). These qualities have all contributed to popular use of microsatellite loci in a wide range of medical and biological fields (Chistiakov et al. 2006). These were also some of the reasons why microsatellites were used in this study. In the coding regions of DNA, frame shift mutations are selected against and this inhibits the formation of repeats other than trinucleotide repeats (Metzgar & Wills 2000, Ellegren 2004). The use of multilocus microsatellites allows for higher resolution compared to other markers such as Randomly Amplified Polymorphic DNA (RAPD) techniques (Sunnucks 2000, Selkoe & Toonen 2006). Previously used, size-based allele identification of microsatellites can result in ambiguous species genealogies or homoplasy (Hewitt 2004, Selkoe & Toonen 2006). This made microsatellites discovered using this method unsuitable for the identification of conservation units (Chistiakov et al. 2006).

Previously, the use of microsatellites was based on pre-determined microsatellite primers which limited the diversity of the microsatellites under investigation, examining a small subset of microsatellite regions in the whole genome (Castoe et al. 2010). Next Generation Sequencing technology allows for the development of randomly sampled microsatellite regions throughout the entire genome (Castoe et al. 2010). This technique is useful for non-model organisms, which are those that have not yet been investigated using genetics. Next Generation Sequencing allows the rapid recognition of a large number of microsatellite regions at a lower cost than with traditional sequencing (Mardis 2008, Castoe et al. 2010, Ekblom & Galindo 2010, Davey et al. 2011, Plough & Marko 2013). Thus, this technology was used in the present study.

4.2 Sampling and DNA extraction

In this study, fin clip samples of *D. capensis* and *N. lithophilus* were collected from two sites in KwaZulu-Natal, South Africa (Cape Vidal and Durban) and one site in Fort Dauphin, Madagascar. A standard salting-out protocol (Sunnucks & Hales 1996) was used to extract genomic DNA. This protocol involved the digestion of tissue samples in 600µl of TNES buffer (50mM Tris pH 7.5; 400mM NaCl; 20mM EDTA; 0.5% SDS) and 15µl of proteinase k with overnight incubation at 55°C. Thirty minute before continuing with the extraction, 10µl of RNase A was added to the samples. The extraction was then conducted, where 240 µl of 3M NaAc was added to each sample, then samples were inverted for 15 seconds before they were centrifuged at 14 000 rotations per minute (rpm) for 40 minutes. The supernatant was retained and 800µl of cold 100% ethanol was added to it. This solution was then incubated at around -10°C for an hour. The solution was then centrifuged at 14 000rpm for 30 minutes. The supernatant was then discarded and 800µl of 70% ethanol was added to the remaining solution. This was then centrifuged at 14 000rpm for 20 minutes. The supernatant was discarded and the DNA pellet that formed was then dried. Once dry the DNA was eluted in 55µl of TE Buffer (10 mM Tris; 1mM EDTA) or AE buffer from the Qiagen DNeasy Blood & Tissue Kit.

4.2.1 Mitochondrial DNA

4.2.1.1 MtDNA amplification

Part of the cytochrome *c* oxidase subunit I (cox1) gene within the mitochondrial DNA was amplified using the universal fish primers Fish F2t1 and Fish F2t1 (see Appendix 1) through Polymerase Chain Reactions (PCR). Each 25 µl PCR reaction contained 12.5µl of Econotaq (Fermentas TM) Green 2X Mastermix, 0.84 µl of each primer (10µM), 1 µl of Borive Serum Albumin or BSA (10mM), 7.82 µl of PCR water and 2 µl of template DNA. These amplifications were conducted using a Bio-Rad T100 Thermal Cycler using the following cycling conditions: 95ºC for 3 minutes for initial denaturation, followed by 35 cycles of denaturation (94ºC for 30 seconds), annealing (45ºC for *D. capensis* and 48 ºC for *N. lithophilus* for 45 seconds) and extension (72ºC for 1 minute) and the final stage of the PCR cycle included a final extension stage where the reactions were held at 72ºC for 10 minutes. The products of these reactions were verified using agarose gel electrophoresis of a 1.2% agarose gel containing 0.05 mg/mL ethidium bromide and were visualized using the Biorad Molecular Imager, Gel Doc™ XR+. All PCR products were sent to Inqaba Biotechnical Industries where the Sanger BigEnd Dye Terminator sequencing method was applied using an ABI 3130XL sequencing platform.

4.2.1.2 mtDNA Analysis

Mitochondrial DNA sequences were then manually edited and aligned with the multiple-alignment algorithm CLUSTAL W within the BioEdit v7.0.9.0 program (Hall 1999). Genetic diversity indices (number of polymorphic sites, haplotype estimation, haplotype and nucleotide diversity) were calculated from these sequences for each population within each fish species using the DNA Sequence Polymorphism programme (DnaSP) v4.90.1 (Rozas & Rozas 1999). The software program, Genetic Analysis in Excel or GenAlEx v6.41 (Peakall & Smouse 2006), was used to estimate most of the remaining genetic diversity indices. These included number of alleles (N_a), number of effective alleles (N_e), Shannon's Index (I) and number of private alleles. ARLEQUIN v3.5 (Excoffier & Lischer 2010) was used to estimate the expected heterozygosity. Parsimony haplotype networks were constructed in R v3.0.2 using the COI data for each fish species. The population structure of each species was calculated using an Analysis of Molecular Variance (AMOVAs) which was calculated from genetic distance matrices in GenAlEx v6.41 (Peakall & Smouse 2006). This analysis was used to determine whether there was a significant genetic difference between all sampled populations within the Southwest Indian Ocean. The pairwise differences between populations for each species was estimated using both the F_{st} statistic (Wright 1946, 1949, 1965) in ARLEQUIN v3.5 (Excoffier & Lischer 2010) and pairwise Φ_{PT} values (Peakall et al. 1995), which are analogous to the F_{st} statistic, calculated in GenAlEx v6.41 (Peakall & Smouse 2006).

The evolutionary model used to reconstruct the phylogenetic relationships between all sampled individuals was determined in MrModeltest 2.3 (Nylander 2004) based on the Akaike Information Criterion (AIC) scores. The results from MrModeltest were used to create a script for MrBayes (lset nst=6 rates=equal). The HKY model (Hasegawa et al. 1985) was selected as the best model of evolution to fit the phylogenetic relationship between samples. A phylogenetic tree based on Bayesian inference was constructed in MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003) with Markov chain Monte Carlo (MCMC) chains run for 1 000 000

generations, trees sampled every 10 generations and a burnin of 10 000 trees. The remaining samples were used to construct a consensus tree based on the 50 % majority rule. Samples from both *D. capensis* and *N. lithophilus* were included in a single phylogenetic tree. *Cyprinus carpio* (JX983284.1) was used to root all trees, while *Lethrinus ornatus* (KM079313.1), *Nemipterus marginatus* (KF009635.1), and *Satyrichthys amiscus* (FJ237914.1) were used as outgroups as these were both distantly related to both *D. capensis* and *N. lithophilus*. Ten sequences from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and BOLD (www.bins.boldsystems.org/) databases were used as ingroups for both *D. capensis* and *N. lithophilus.* The ingroups for *D. capensis* were: JX192306.1; JX192287.1; and JX192306.1 from South Africa (Henriques 2012) ; JF493380.1 (Specimen voucher Smith 183.17 from South Africa). The ingroups for *N. lithophilus* were: JF493986.1 from Cape Vidal, South Africa; JF493987.1 from Durban, South Africa and TZSAN189-06; TZSAN190-06; DSLAR452-09; and DSLAR383-08 from KwaZulu-Natal, South Africa.

Neutrality test indices, Tajima's D (Tajima 1989) and Fu's Fs (Fu 1997) generated in ARLEQUIN v3.5 (Excoffier & Lischer 2010), would normally be used to test whether populations of both *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean are under neutral or non-selective evolution. However, here these indices were used to test for demographic expansion. Tajima's *D* compares two estimators of molecular evolution within populations, which should estimate the same quantity under the infinite-site model in population genetics (Tajima 1989). These estimators are the number of segregating sites or sites showing polymorphism, and the mean number of pairwise difference between haplotypes (Tajima 1989). Like Tajima's D, Fu's Fs test the neutrality of genetic mutations however, Fu's Fs is a more powerful test and can also be used to indicate demographic changes within populations (Fu 1997). Mismatch distributions illustrate the pairwise differences among all DNA sequences (Rogers & Harpending 1992, Harpending et al. 1993), where the slope of the graph represents the genealogical history of the population (Aris-Brosou & Excoffier 1996). A unimodel or smooth mismatch distribution reflects an expanding population where the peak represents the time of population growth, when genetic information has been conserved instead of lost as observed during population declines (Harpending 1994). A ragged or erratic mismatch distribution reflects a population that has been stationary for a long time (Harpending 1994). The raggedness index quantifies the smoothness of the observed mismatch distribution (Harpending 1994). The observed and expected mismatch distributions estimated under the model of demographic expansion and their raggedness indices were calculated for each population for *D. capensis* and *N. lithophilus* using ARLEQUIN v3.5 (Excoffier & Lischer 2010). The Sum of Squared Differences (SSDs) was used to test for sudden expansion (Schneider & Excoffier 1999) and was estimated using ARLEQUIN v3.5 (Excoffier & Lischer 2010).

Bayesian Skyline Plots were used to depict the fluctuations in effective population size (N_e) over time, and were constructed using mtDNA data for both *D. capensis* and *N. lithophilus* populations in Tracer v1.5 (Rambaut & Drummond 2010). Input files for BEAST v1.8.1 (Drummond & Rambaut 2007) were first generated in BEAuti v1.8.1 using the following parameters: an HYK substitution model, a 4 category gamma site heterogeneity model and lognormal relaxed molecular clock model. The MCMC chains were runs for 30 000 000 iterations with a burnin of 300 000 runs. Two independent runs were conducted and combined before the data were analysed in Tracer v1.5 (Rambaut & Drummond 2010), where Effective Sample Size or (ESS) values greater than 200 were used a indicators of accuracy. Tables of the posteriors and the continuous parameters (mean and ESS values) are presented in the Appendix (Appendix 3 and 4) for each species. In Tracer v1.5, the Bayesian Skyline plots were constructed using the stepwise constant model.

4.2.2 Microsatellite analyses

4.2.2.1 Microsatellite discovery

Only pure, high quality DNA (absorption ratio $A_{260/280}$ of 1.8 – 2.0) was used for DNA library construction using shotgun sequencing for both *D. capensis* and *N. lithophilus*. These samples were then sent to the Agricultural Research Council in Pretoria for library construction using shotgun sequencing with the Illumina Miseq platform. Shotgun sequencing techniques can be used to identify microsatellite loci, by first producing a large number of random sequences and using the most repeated sequences for analysis (Hudson 2008, Mardis 2008). Shotgun sequencing involves the construction of contigs which are sequences constructed using sequenced reads (Hudson 2008, Ellegren 2014).The resultant paired end read data were processed on Illumina's online cloud system, BaseSpace and the online system The Galaxy Project (www.galaxyproject.org); a free online NGS bioinformatics system. The adaptors were removed from reads using the FASTQ toolkit v1.0 on BaseSpace. This raw data were then uploaded onto The Galaxy Project where the reads were concatenated. SOAPdenovo v2.01 (Luo et al. 2012) was then used to assemble the sequence reads without a reference genome or *de novo*. The aligned reads were saved as FASTA-formatted files and used in MSATCOMMANDER (Faircloth 2008) to detect di-, tri- and tetranucleotide microsatellite regions within the assembled reads, using default settings. Dinucleotide repeats are found largely in the noncoding regions of DNA, while trinucleotide repeats are found largely in the coding regions of DNA (Senan et al. 2014). The amplification of dinucleotide repeats has been shown to result in the occurrence of slippage or an increase in the primer-template complex, and thus may not be favourable (Edwards et al. 1991). However, dinucleotide repeats have been shown to be polymorphic in many fish species such as *Chrysoblephus puniceus, Pagellus bogaraveo; Diplodus sargus* and *Diplodus vulgaris* (Piñera et al. 2006, Roques et al. 2007, Chopelet et al. 2009a), which are all of the Sparidae family like *D. capensis* and thus were used in this study. Interrupted microsatellite repeats and repeats of less than eight were also avoided during microsatellite selection.

4.2.2.2 Primer design and selection

Primer3 (Rozen & Skaletsky 1999) was used to design primers for the microsatellite regions identified with MSATCOMMANDER. Thirty primer pairs were selected from these potential primers based on the following characteristics: microsatellites must be uninterrupted or perfect as defined by Oliveira et al. (2006), microsatellites must be repeated eight times or more, primer pair product size less than 280bp due to the read length of the Miseq platform, primer pairs that have a low chance of secondary structure formation, and repeated primer sequences must also avoided. Loci that have been previously found to be polymorphic in related species were included with their designed primers (Appendix 2). The primers labeled DCMS14 and 20 have been used to amplify microsatellites in *Pachymetopon blochii* (Reid et al. 2012). The primer labeled DCMS 26 has been used to amplify microsatellites in *Pagellus bogaraveo* (Piñera et al. 2006). The primers labeled DCMS 27 and 28 have been used to amplify microsatellites in *Acanthopagrus schlegeli* and *Pagrus auratus* (Adcodk et al. 2000, Liu et al. 2007). For the *N. lithophilus* samples, the use of the assembled sequence reads resulted in Primer3 producing less than 30 usable primer pairs. Thus, the unassembled sequence reads data was used in the microsatellite discovery and primer design. To aid in the identification and processing of amplicons during the genotyping phase, nextera tails with the following sequence were added to the 5' end of all forward and reverse primers:

5’-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3’
5’-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3’

These oligonucleotides were then synthesized at the Beijing Genomics Institute (BGI).

4.2.2.3 Microsatellite amplification

Thirty microsatellite loci from each fish species (*D. capensis* and *N. lithophilus*) were amplified using the QIAGEN® Multiplex PCR *Plus* kit following the standard kit protocol. Multiplex PCR technology allows the amplification of multiple DNA loci from a single DNA sample at the same time. The use of this type of PCR allows the investigation of several more DNA loci that with singleplex PCR while saving time and costs (Kircher et al. 2011). The thirty primer pairs were separated into two groups of fifteen primer pairs each. These primer pairs were then mixed together by mixing 10µl of each primer (2µM) to make the ‘primer mix’ solution. Each species had two primer mixes that amplified the selected thirty microsatellite loci. Each 50 µl reaction contained 25 µl of 2x Multiplex PCR Master Mix, 5 µl of Primer mix, 5 µl of 10x CoalLoad Dye, 14 µl of RNase free water and 1 µl of template DNA. These microsatellite regions were amplified using the following cycling conditions: 95ºC for 5 minutes for initial denaturation, followed by 35 cycles of denaturation (95ºC for 30 seconds), annealing (60ºC for 90 seconds) and extension (72ºC for 90 seconds), and the final stage of the PCR cycle included a final extension stage were the reactions were held at 68ºC for 10 minutes. Amplicons were cleaned following a standard PCR clean-up protocol and suspended in 20µl 0.1X TE Buffer. These amplicons where then sent to the Agricultural Research Council in Pretoria for sequencing.

4.2.2.4 Microsatellite data analyses

A matrix was constructed for each species containing the scores of each microsatellite locus in each sample. A total of 24 microsatellite loci were amplified in the *D. capensis* samples and 11 microsatellite loci were amplified in the *N. lithophilus* samples. These data were then imported into GenAlEx v6.41 (Peakall & Smouse 2006). Molecular diversity indices such as number of alleles, number of effective alleles, Shannon’s Information Index, number of private alleles, expected and observed heterozygosity as well as Hardy-Weinberg tests were estimated for each population in GenAlEx v6.41 (Peakall & Smouse 2006). Linkage disequilibrium test was also conducted in ARLEQUIN v3.5 (Excoffier & Lischer 2010) was used to determine whether pairs of loci are linked or associate with one another. An estimation of the Probability of Identity, which estimates the probability that any two unrelated individuals within a population would have the same multilocus genotype (Waits et al. 2001) was calculated for each *D. capensis* population in GenAlEx v6.41 (Peakall & Smouse 2006). This test could not be conducted with the *N. lithophilus* data, which could be a result of the low number of loci examined (only 10 loci). A population assignment test was conducted using GENECLASS2 (Piry et al. 2004). This analysis determines the likelihood that an individual sample originated from a certain population based on the genotypes of individuals and populations, thus allowing the estimation of recent dispersal or migration of individuals (Paetkau et al. 2004). Principal Coordinate Analyses (PCA) of allelic frequencies were calculated in GenAlEx v6.41 (Peakall & Smouse 2006) for both *D. capensis* and *N. lithophilus*. The Analysis of Molecular Variance (AMOVA) using R_{st} statistic (Slatkin 1995) in ARLEQUIN v3.5 (Excoffier & Lischer 2010) was used to determine whether there is a genetic difference between all populations of *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean. Though the R_{st} statistic uses the step-

wise mutation model, it may not be the best estimation of genetic variation (Neigel 2002). Therefore, an AMOVA using the F_{st} statistic was also conducted using ARLEQUIN v3.5 (Excoffier & Lischer 2010). A pairwise estimation of population differentiation using the F_{st} statistic was also conducted using ARLEQUIN v3.5 (Excoffier & Lischer 2010). The pairwise estimations of Nei's standardized G_{st} statistic (Nei 1968, 1987), which is analogous to F_{st}, was estimated in GenAlEx v6.41 (Peakall & Smouse 2006). This analysis determines the geographic distribution of genetic variation across populations (Palumbi 2003). A Mantel test (Mantel 1967) was conducted in GenAlEx v6.41 (Peakall & Smouse 2006) to test for isolation-by-distance between populations of *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean. This analysis tests the statistical significance between two matrices of fish populations; geographic distance matrix and genetic variation matrix (Smouse et al. 1986, Smouse & Long 1992).

The program STRUCTURE v.2.3 (Pritchard et al. 2000) uses Bayesian clustering methods to decipher population structure within a sample set. This analysis uses 'K' to represent a certain number of populations, with each population being characterised by specific frequencies at each locus. This means that each population can be identified by these frequencies. The analysis was first run without using the 'locprior' function which clusters the samples based on predefined populations, which were Madagascar, Cape Vidal and Durban for both *D. capensis* and *N. lithophilus*. The second analysis was run using the 'locprior' function for predefined populations. All analyses were run 10 times using a burnin length period of 10 000 and 20 000 MCMC reps and the numbers of simulated populations or 'K' values were from one to ten. Structure harvester (Earl & von Holdt 2012) was used to infer the best K, which used the Delta K method (Evanno et al. 2005).

CHAPTER 5 – The evolutionary history of *Diplodus capensis* and *Neoscorpis lithophilus* in the Southwest Indian Ocean based on mitochondrial DNA

5.1 Results

5.1.1 *Diplodus capensis*

A total of eight haplotypes were recovered from 24 *D. capensis* DNA sequences, which were 524bp in length. Three of the haplotypes recovered were shared and represented by more than one individual, while the remaining five haplotypes were unique in that they were represented by a single individual. Two haplotypes were represented by samples in all three populations (Figure 2). The haplotype frequencies in each *D. capensis* population sampled are summarised in the pie charts shown in Figure 2. Haplotype 2, illustrated in orange, has a high frequency across all populations and consists of 10 individuals which make up about 40% of all sampled individuals (Figure 1). In the Madagascar population, Haplotype 1 has the greatest frequency. The overall haplotype diversity was estimated to be 0.757 (±0.066). The molecular indices for each population are summarized in Table 1 and Figure 2. The Cape Vidal population was found to be the most genetically diverse of the populations sampled, as this population illustrated the highest values for all the genetic indices estimated (Table 1; Figure 2). *Diplodus capensis* populations in Madagascar and Durban were found to have similar levels of molecular diversity, in terms of haplotype (0.75 and 0.70, respectively) and nucleotide diversity (2.12×10^{-3} and 2.29×10^{-3}, respectively), mean number of alleles (1.33 for both populations), polymorphism within populations (3 polymorphic sites in both populations) and expected heterozygosity (0.12 and 0.13, respectively). The samples sizes between Cape Vidal and Madagascar differ only by one individual, however the polymorphism observed in the Cape Vidal population is about twice that observed in Madagascar (Table 1). The only index that was found to be similar across all sampled populations of *D. capensis* was the number of effective alleles, with a range of 1.16 – 1.28 (Figure 2).

Table 1: Molecular diversity indices of the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *D. capensis* populations in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban).

Population	n	h	Hd	π	No. of alleles	No. of polymorphic sites	Expected heterozygosity
Madagascar	9	4	0.750	2.12×10^{-3}	1.333 (±0.50)	3	0.123 (±0.20)
Cape Vidal	10	5	0.822	3.77×10^{-3}	1.778 (±0.44)	7	0.218 (±0.16)
Durban	5	3	0.700	2.29×10^{-3}	1.333 (0.50)	3	0.133 (±0.20)

N = sample size; h = number of haplotypes; Hd = haplotype diversity; π = nucleotide diversity

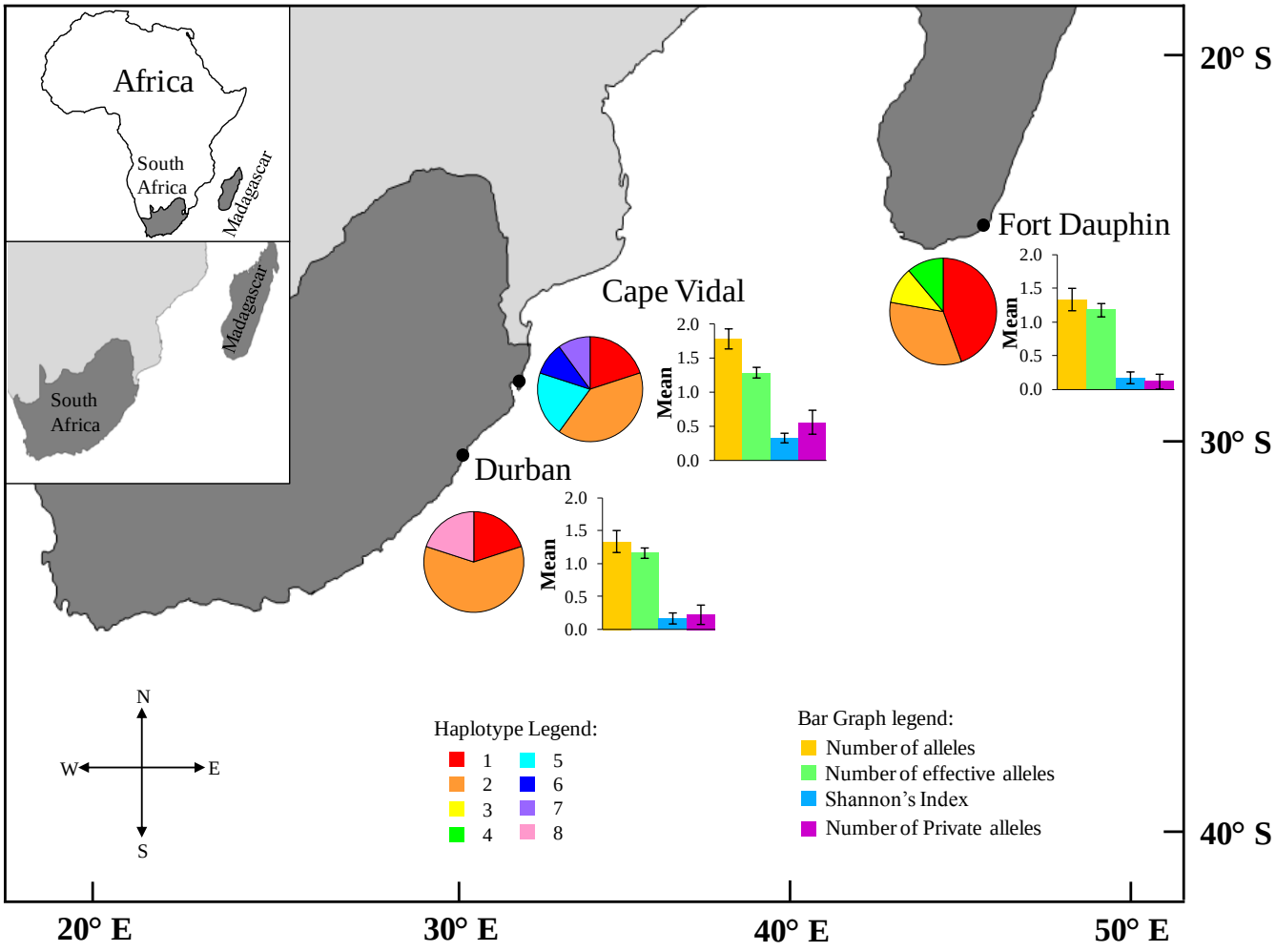

Figure 2: Haplotype frequencies of cytochrome oxidase subunit I (cox1) gene in the mtDNA of the *D. capensis* populations sampled in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban) are illustrated using pie charts. The colours represent the haplotypes found within each population. The adjacent bar graphs illustrate the genetic diversity indices (number of alleles; number of effective alleles; Shannon's Index; and Number of Private alleles) calculated for each population.

The haplotype network illustrated in Figure 3 shows no significant structuring between the sampled populations of *D. capensis* in the SWIO. The central haplotypes in this figure, HI and HII, are surrounded by the remaining haplotypes, separated by only one or two mutational steps. The two unique haplotypes in the Madagascar population are both found to cluster with HI, while the only unique haplotype in the Durban population clusters with HII. This suggests that the unique haplotypes found in the Madagascar population diverged from HI, while the unique haplotype in the Durban population diverged from HII. The unique haplotypes of the Cape Vidal population cluster with both haplotypes.

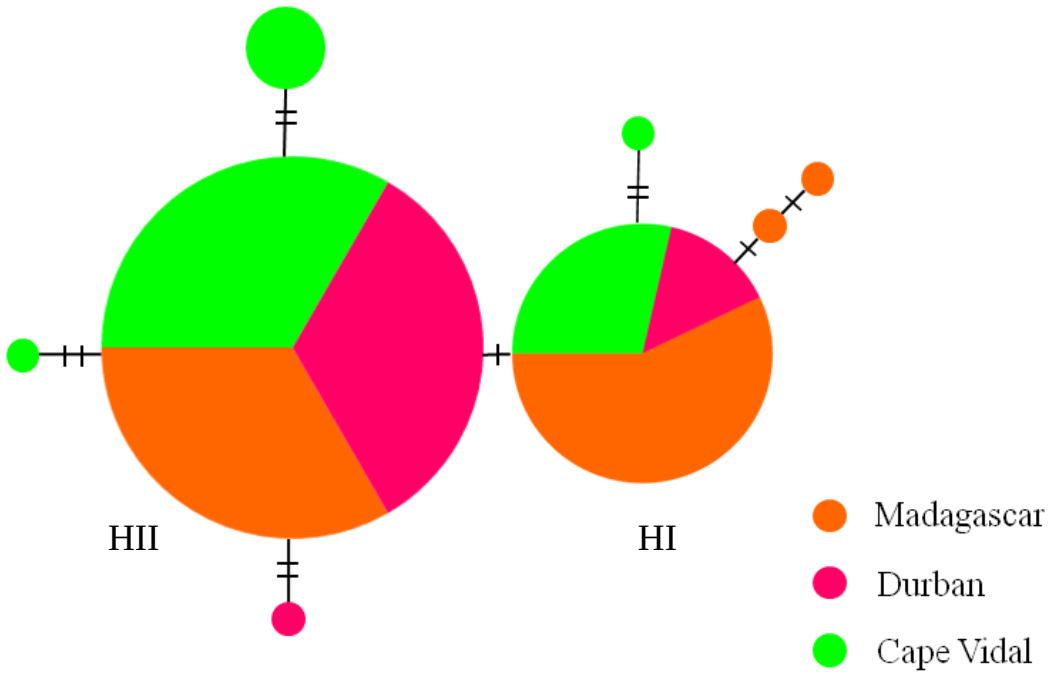

Figure 3: Haplotype network of the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *D. capensis* of three populations in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban). The circles represent the eight haplotypes discovered in this study. Populations are represented by the colours in the circles, with circle size drawn in proportion to the number of individuals within each haplotype. The dash lines represent the number of mutational steps between haplotypes.

The AMOVA conducted to estimate the genetic variation between all sampled populations of *D. capensis* in the SWIO showed no significant difference, with a p-value of 0.24, between populations (Table 2). The variation within populations accounted for 95% of the observed variation, while the variation among populations accounted for only 5% of the observed molecular variance. The Φ_{PT} value, calculated using 1000 permutations, was estimated to be 0.052; however this value was not significant (Table 2).The pairwise calculations of genetic variation between populations are summarized in Table 3. Estimates of genetic variation using the F_{st} statistic were found to be fairly low, ranging from -0.006 (between Durban and Cape Vidal) and 0.133 (between Madagascar and Durban). This is indicates that there is some evidence of historic gene flow between all sampled populations of *D. capensis* in the SWIO. However, these values were not found to be significant and may have been affected by the sample size used to make these estimations. These results coincide with the estimated pairwise

Φ_{PT} values, which ranged from -0.03 (between Durban and Cape Vidal) and 0.138 (between Durban and Madagascar), however, had no statistical significance.

Table 2: Results of the Analysis of Molecular Variance (AMOVA) showing the molecular variance of cytochrome oxidase subunit I (cox1) gene in the mtDNA of *D. capensis* among and within all populations (Madagascar, Cape Vidal and Durban) within the Southwest Indian Ocean.

Source of variation	df	Estimated variation	Variation %	Φ_{PT}
Among Populations	2	0.041	5%	0.052
Within Populations	21	0.750	95%	

Bold values illustrate statistically significant values

Table 3: Results of the pairwise estimation of population differentiation based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA using Φ_{PT} and F_{st} statistics between populations (Madagascar, Cape Vidal and Durban) of *D. capensis* in the Southwest Indian Ocean. The values below the diagonal represent the Φ_{PT} values and the values above the diagonal represent the F_{st} values.

	Madagascar	Cape Vidal	Durban
Madagascar		0.067	0.133
Cape Vidal	0.063		-0.006
Durban	0.138	-0.03	

Bold values illustrate statistically significant values

The phylogenetic relationship between *D. capensis* individuals from three populations in the SWIO (Madagascar, Cape Vidal and Durban) is illustrated in Figure 4. Although the sampled *D. capensis* individuals branch out together with other specimen sequences from Genbank sampled in South Africa, the samples were shown to be more similar to each other rather than the sequences from Genbank. The *D. capensis* samples in this tree showed a comb pattern of divergence, illustrating that no genetic structure was found between these samples.

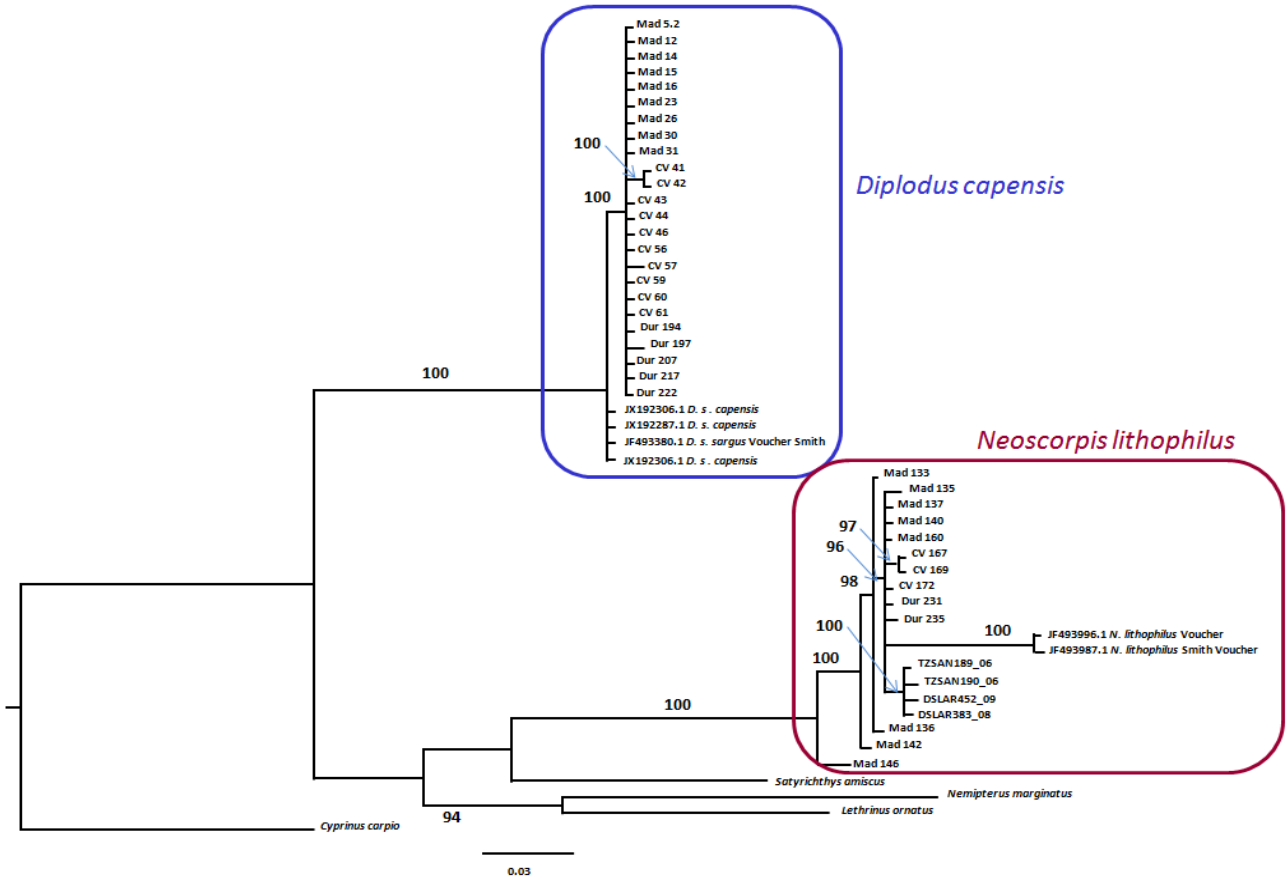

Figure 4: Phylogenetic tree constructed using Bayesian Inference based of mitochondrial COI sequences from both *D. capensis* and *N. lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban). Values represent the bootstrap support for the illustrated node.

The neutrality test using Tajima's *D* has failed to reject the hypothesis of neutral selection and shows no evidence of past demographic expansion. A positive value for the estimation of Tajima's *D* implies the selection of one allele over alternative alleles, while a negative value implies that heterozygotes have a selective advantage in the population (Stephens et al. 2001). The estimated value of Tajima's *D* for all populations was shown to be a negative value of -1.16 (*p*-value of 0.125). This suggests that heterozygosity is selected for in the *D. capensis* populations of the Southwest Indian Ocean and shows no evidence of past demographic expansion. Separate calculations of Tajima's *D* for each population, showed negative values for the Cape Vidal and Durban population (-0.85 and -1.05, respectively) and a positive value for the Madagascar population (0.03). Thus, these estimates reveal that one allele in the Cape Vidal and Durban populations is selected for over alternate alleles in these populations. On the other hand, heterozygosity is being favoured in the Madagascar population. However, these values for Tajima's *D* were found to be insignificant (Table 4). The estimated value of Fu's Fs for all populations was found to be a negative value of -2.67 (*p*-value of 0.045). This negative values illustrates that these populations have an excess of recent mutation or rare alleles (Fu 1997). Thus, populations of *D. capensis* in the Southwest Indian Ocean are not under neutral selection. However, Fu's Fs is sensitive to past demographic expansion and this sensitivity is manifested in the form of large negative values (Liu et al. 2006). This might be the case with this data, thus this test suggests that populations of *D. capensis* in the Southwest Indian Ocean have undergone

past demographic expansion. Estimations of Fu's Fs for each population were found to be negative for all sampled populations, ranging from -0.19 for the Durban population to -0.82 for the Madagascar population (Table 4).

The demographic history for the three sampled populations of *D. capensis* in the Southwest Indian Ocean was examined using mismatch distributions (Figure 5A, B, and C). All populations illustrated smooth and unimodal mismatch distributions, which are an indication of growing populations where a single population expansion event occurred in the past (Rogers & Harpending 1992, Harpending 1994). The raggedness index (RI), which tests the smoothness of these mismatch distribution graphs (Harpending 1994), also presented evidence of past demographic expansion. The highest RI value was estimated for the Madagascar population (0.15) and the lowest RI value was estimated for the Durban population (0.05). All of the estimated values for RI were found to be insignificant (Table 4), thus the analysis failed to reject the hypothesis of past demographic expansion in the *D. capensis* populations of the SWIO. The Sum of Square Differences (SSD), which tests for sudden population expansion (Schneider & Excoffier 1999) also illustrated that these populations have undergone past demographic expansion. These SSD values for Madagascar, Cape Vidal and Durban (0.015; 0.014 and 0.006, respectively) were also found to be insignificant (Table 4), and thus, the analysis failed to reject the hypothesis of past demographic expansion. The Bayesian Skyline plot shown in Figure 6 represents the changes in effective population size over time. This graph shows that the *D. capensis* population has had a fairly stable effective population size for female individuals in Madagascar and South Africa with no evidence indicating population expansion or bottleneck events. This observation contradicts the results of the mismatch distribution.

Table 4: The results of the test of neutrality (Tajima's D and Fu's Fs) and the population demographic indices (Raggedness Index and Sum of Squared Differences) based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of the *D. capensis* populations (Madagascar, Cape Vidal and Durban) in the Southwest Indian Ocean.

	Madagascar	**Cape Vidal**	**Durban**
Number of pairwise differences	4	6	4
Tajima's D	0.025	-0.850	-1.048
Fu's Fs	-0.822	-0.628	-1.86
Raggedness Index	0.146	0.062	0.050
SSD	0.015	0.014	0.006

Bold values represent statistically significant values

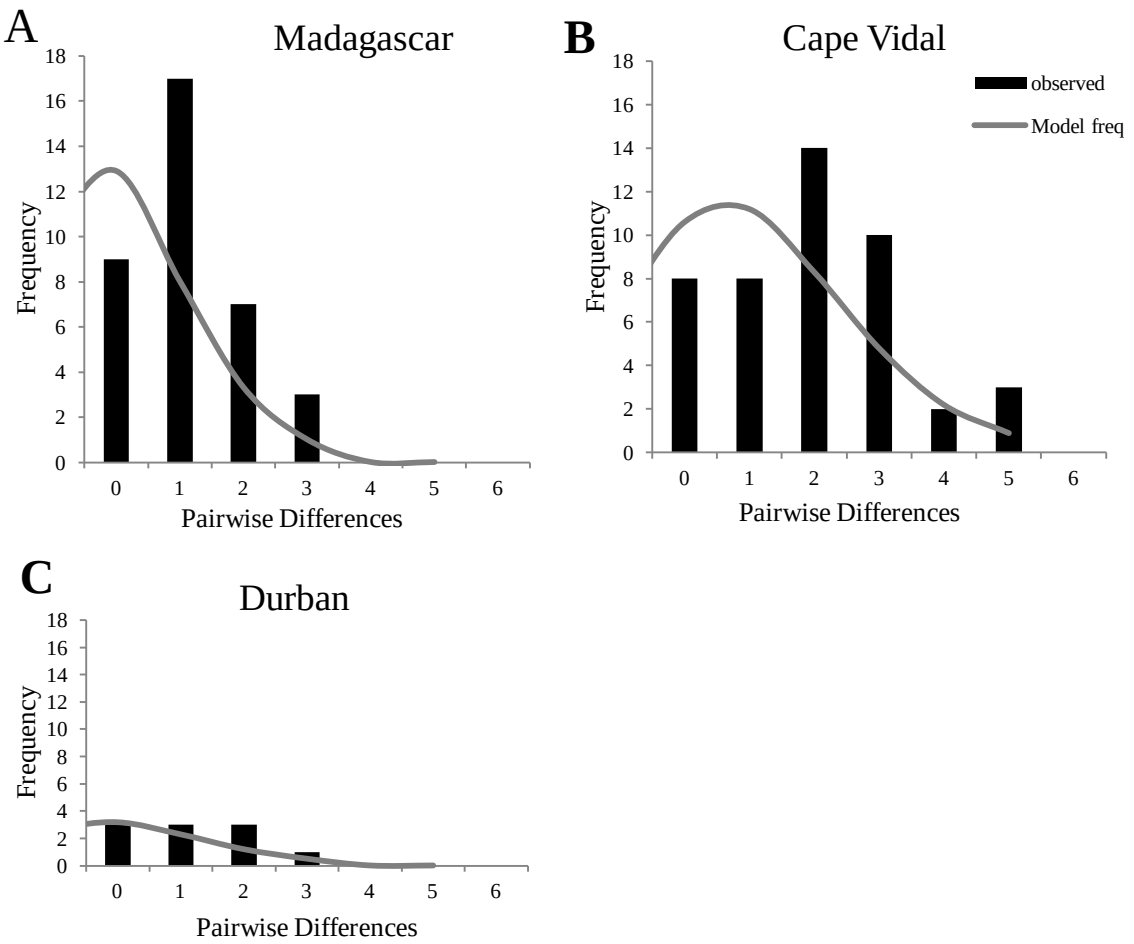

Figure 5: Mismatch distributions showing the observed pairwise differences of the cytochrome oxidase subunit I (cox1) gene in the mtDNA (illustrated by the black bars on the graphs) and the expected expansion models (illustrated by the grey lines on the graphs) for *D. capensis* populations in the Southwest Indian Ocean.

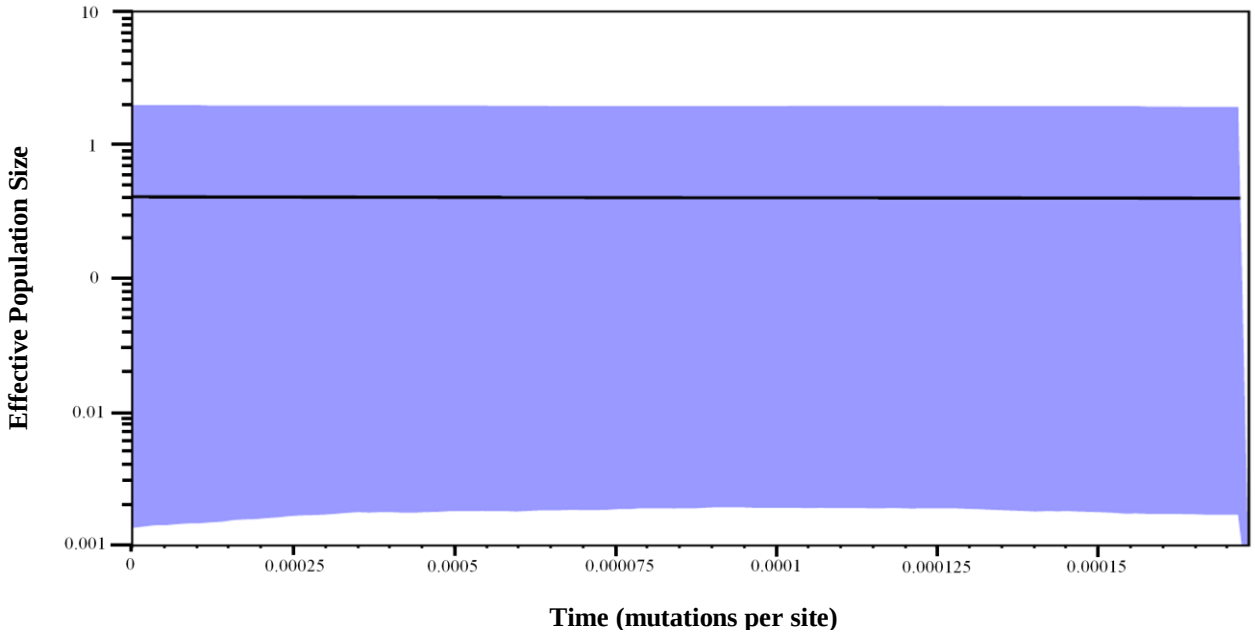

Figure 6: Bayesian Skyline Plot (BSP) showing the mean effective population size of female *D. capensis* in the Southwest Indian Ocean (Madagascar, Cape Vidal, Durban) over time based on cytochrome oxidase subunit I (cox1) gene in the mtDNA. The blue shading on the graph represents the 95% confidence intervals of the calculated mean effective population size.

5.1.2 *Neoscorpis lithophilus*

Seven haplotypes were recovered from thirteen *N. lithophilus* DNA sequences, each of which was 482bp in length. Three of the haplotypes recovered were shared and were represented by more than one individual, while the remaining four haplotypes were private haplotypes in that they were only represented by one individual. Only one haplotype was represented by samples in all three populations, illustrated in yellow in the pie charts (Figure 7). The Cape Vidal and Durban populations were shown to illustrate a different haplotype frequency compare to that of the Madagascar population (Figure 7). The overall haplotype diversity was estimated to be 0.85 (± 0.09). The Madagascar population was found to have the greatest genetic diversity, in terms of nucleotide diversity (8.82×10^{-3}); number of alleles (1.88); number of private alleles (15 polymorphic sites) and Shannon's Index (0.36) (Table 5; Figure 7). The Cape Vidal population was estimated to have the lowest genetic diversity (Table 5; Figure 7). The genetic diversity of the Durban population was estimated to be higher than that of the Cape Vidal population. Estimations of haplotype and nucleotide diversity and expected heterozygosity for the Durban population were found to be higher than expected. Considering the sample size (n = 2) of this population, this may have affect the accuracy of these estimation (Pruett & Winker 2008, Sinclair & Hobbs 2009, Fumagalli 2013). The Cape Vidal population also has a small sample size (n = 3), thus the genetic indices estimated for this population may also be exaggerated.

Table 5: Molecular diversity indices based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA in sampled populations of *N. lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban).

Population	n	h	Hd	π	No. of alleles	No. of polymorphic sites	Expected heterozygosity
Madagascar	8	5	0.857	8.82×10^{-3}	1.882 (±0.32)	15	0.250 (±0.13)
Cape Vidal	3	2	0.667	1.38×10^{-3}	1	1	0.667
Durban	2	2	1.000	2.07×10^{-3}	1	1	1.000

N = sample size; h = number of haplotypes; Hd = haplotype diversity; π = nucleotide diversity

The haplotype network based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *N. lithophilus* individuals in the SWIO is illustrated in Figure 8. This haplotype network forms a star-burst pattern that illustrates population expansion where unique haplotypes have diverged from a central or ancestral haplotype. The ancestral haplotype contains individuals from all three sampled populations of *N. lithophilus;* however this haplotype is represented in more individuals from Madagascar than any other population. This could be, however, related to the sample size of each of these populations. This haplotype is also the only shared haplotype among all populations. The remaining haplotypes seem to cluster closely to this central haplotype and do not differ by more than one or two mutational steps. This haplotype network shows that one haplotype, which is represented by one individual from Madagascar, differs from the main cluster of haplotypes by 11 mutational steps. This distinct haplotype corresponds with the sample labelled 'Mad 146' in Figure 4.

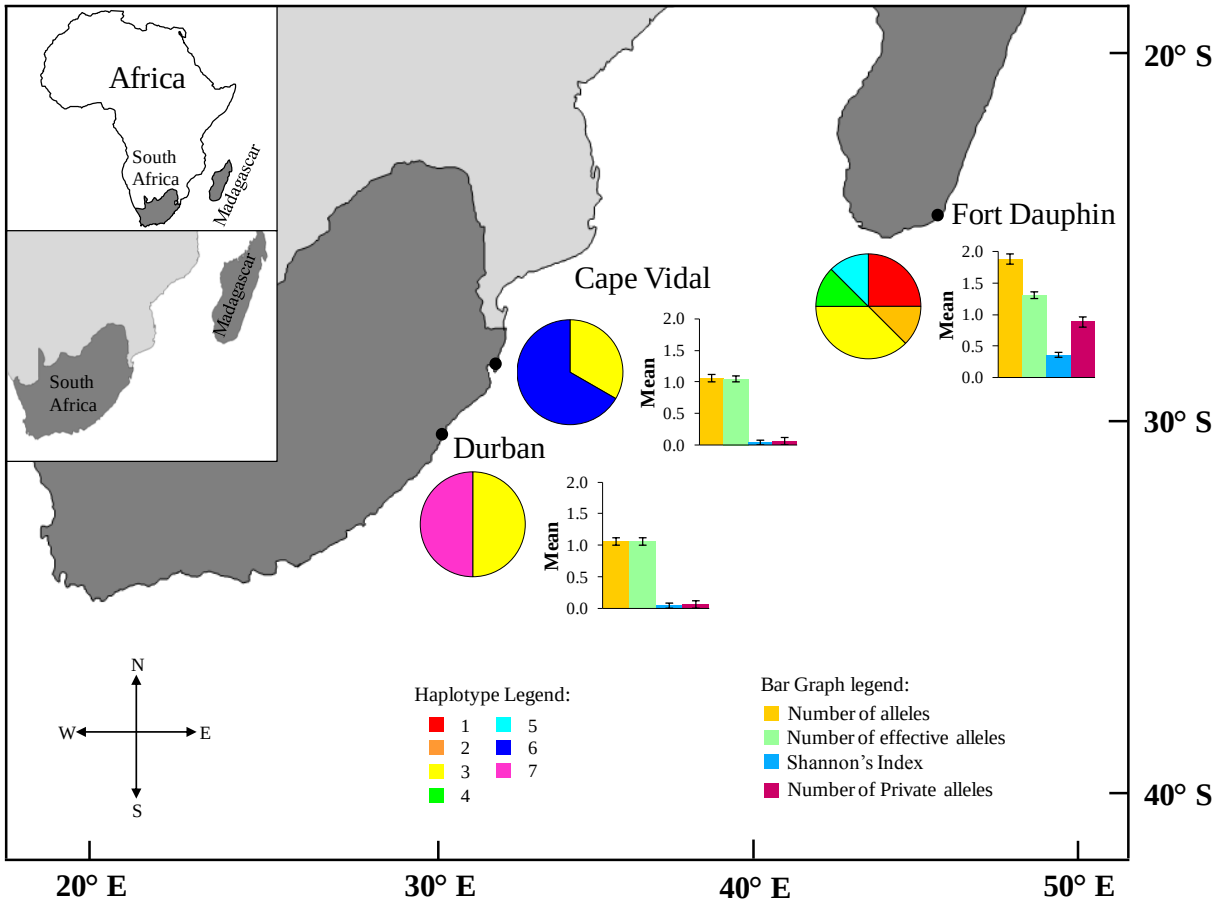

Figure 7: Haplotype frequencies based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *N. lithophilus* populations sampled in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban) are illustrated using pie charts. The colours represent the haplotypes found within each population. The adjacent bar graphs illustrate the genetic diversity indices (number of alleles; number of effective alleles; Shannon's Index; and Number of Private alleles) calculated for each population.

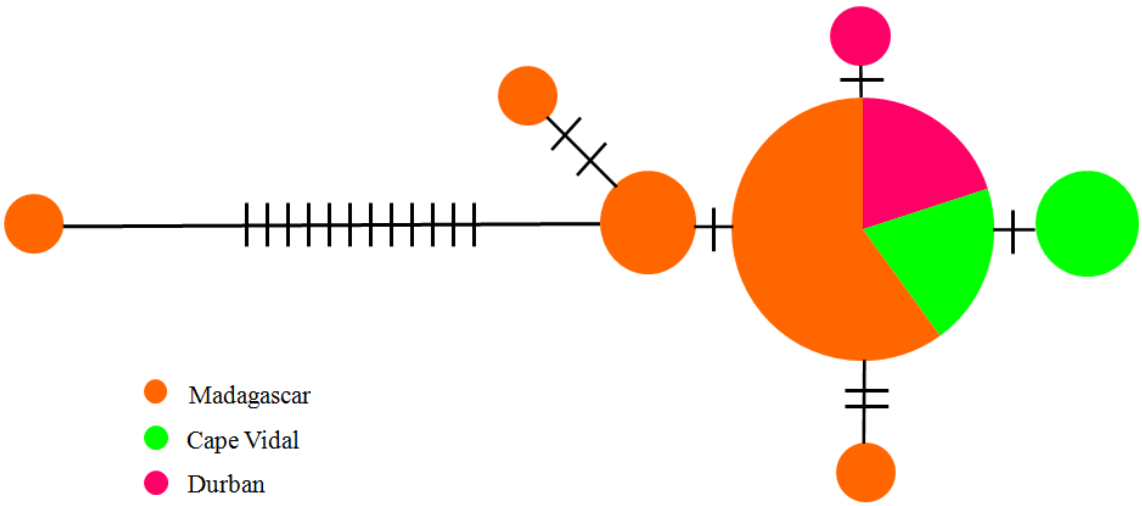

Figure 8: Haplotype network based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *N. lithophilus* in three populations in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban). The circles represent the seven haplotypes discovered in this study. Populations are represented by the colours in the circles, with circle size drawn in proportion to the number of individuals within each haplotype. The dash lines represent the number of mutational steps between haplotypes.

The AMOVA was conducted to determine the genetic differentiation between populations of *N. lithophilus* in the SWIO (Table 6). This analysis shows no significant difference between the populations of *N. lithophilus* in the SWIO, with an insignificant Φ_{PT} value of -0.015. The AMOVA illustrated that none of the observed variation within this data set was due to the differences between populations. However, all of the variation (100%) observed in the sample set was due to the difference within populations. This is not expected due to the difference in genetic diversity observed in Table 5 and Figure 7. The low sample size of the Cape Vidal and Durban populations (n = 3 and 2, respectively) could have affected the estimation of Φ_{PT} during this analysis. Individuals with high genetic diversity, like the sample 'Mad 146', have been known to hinder the accurate estimation of genetic structure between populations (Charlesworth 1998). Thus this AMOVA was re-run without the 'Mad 146' sample that was estimated to be highly divergent by the haplotype network (Table 6). This AMOVA, however, also did not find any evidence for genetic structure between populations. However, the variation among populations was shown to increase from 0% to 14% with the removal of this sample from the analysis.

The pairwise estimation of Φ_{PT} between populations of *N. lithophilus* found relatively low, insignificant values for comparisons involving Madagascar (0.05 and -0.13 for Cape Vidal and Durban, respectively) (Table 7). The estimated pairwise Φ_{PT} between Cape Vidal and Durban was found to be 0.32, which was not found to be significant. The estimation of the pairwise F_{st} was estimated to be highest for the comparison of Cape Vidal and Durban (0.29), which shows evidence of high genetic differentiation between these two populations. The lowest estimation of pairwise F_{st} values was found to be between Madagascar and Durban (0.087) which suggests moderate genetic differentiation between these populations. Once again, the sample size of Cape Vidal and Durban populations may have influenced the outcome of this analysis, resulting in an over estimation of the F_{st} statistic.

Table 6: Results of the Analysis of Molecular Variance (AMOVA) showing the molecular variance among and within all populations (Madagascar, Cape Vidal and Durban) of *N. lithophilus* within the Southwest Indian Ocean.

	Source of variation	df	Estimated variation	Variation (%)	Φ_{PT}
All samples	Among Populations	2	0.000	0	-0.015
	Within Populations	10	1.604	100	
Excluding 'Mad 146'	Among Populations	2	0.115	14%	0.141
	Within Populations	9	0.701	86%	

Bold values illustrate statistically significant values

Table 7: Results of the pairwise estimation of population differentiation using Φ_{PT} and F_{st} statistics (estimated in GenAlEx and DnaSP, respectively) between populations (Madagascar, Cape Vidal and Durban) based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of *N. lithophilus* in the Southwest Indian Ocean. The values below the diagonal represent the Φ_{PT} values and the values above the diagonal represent the F_{st} values.

	Madagascar	Cape Vidal	Durban
Madagascar	-	0.192	0.087
Cape Vidal	0.047	-	0.286
Durban	-0.128	0.323	-

Bold values illustrate statistically significant values

The phylogenetic tree that is depicted in Figure 4 also shows the phylogenetic relationship between *N. lithophilus* individuals in the SWIO (Madagascar, Cape Vidal and Durban). This tree shows slight structuring between some samples from Madagascar, Cape

Vidal and Durban. Four samples from Madagascar (Mad 146, Mad 142, Mad 133 and Mad 136) were found to be substantially dissimilar to the remaining samples. These samples formed their own private haplotypes within this data set (also seen in Figure 8). The sample labelled 'Mad 146' corresponds to the distinct haplotype in the haplotype network. The *N. lithophilus* clade of Figure 4 contains the both sequences from specimens from Genbank (sampled in KwaZulu-Natal, South Africa) and sampled individuals. The voucher specimens are shown to be more similar to the majority of the *N. lithophilus* samples. This main clade consists of all the individuals sampled from both Cape Vidal and Durban and the remaining Madagascan samples. These samples represent the central haplotype observed in Figure 8 (Mad 137; Mad 140; Mad 160; CV 172; and Dur 231).

The neutrality test conducted for all the populations using Tajima's *D* rejected the hypothesis of neutral or no selection in *N. lithophilus* populations in the SWIO, with an overall estimation of -1.77 (*p*-value of 0.02). This suggests that these populations are under selection, where one allele is selected for over alternative alleles (Stephens et al. 2001). These results also present evidence of past demographic expansion in the *N. lithophilus* populations of the SWIO. Calculations of Tajima's *D* for each population showed a negative value for the Madagascar population and estimations for Madagascar that exclude the sample 'Mad 146' that is refered to as *Madagascar (-1.36 and -0.793) (Table 8), however these values were found to be insignificant. The neutrality test conducted for all populations using Fu's Fs failed to reject the hypothesis of neutral selection in the *N. lithophilus* populations in the SWIO, with an overall estimation of -0.95 (*p*-value of 0.29). This analysis also shows evidence of past demographic expansion of *N. lithophilus* in the Madagascar population in the SWIO. These results agree with those suggested by the Tajima's *D* analysis. Calculations of Fu's Fs for each population found positive values for the Madagascar and Cape Vidal populations (0.45 and 0.20, respectively) and a negative value for *Madagascar of -0.428 (Table 8). However, these estimations were found to be insignificant. The Tajima's *D* value for the Durban population was found to be zero, which could have been influenced by the small sample size of this population.

Table 8: The results of the test of neutrality (Tajima's D and Fu's Fs) and the population demographic indices (Raggedness Index and Sum of Squared Differences) based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA of the *N. lithophilus* populations (Madagascar, Cape Vidal and Durban, as well as Madagascar without 'Mad 146' which is represented by an astrix) in the Southwest Indian Ocean.

	Madagascar	*Madagascar	Cape Vidal	Durban
No. of pairwise differences	15	5	2	-
Tajima's D	-1.358	-0.793	0	0
Fu's Fs	0.445	-0.428	0.201	0
Raggedness Index	0.066	0.726	0.556	-
SSD	**0.179**	0.0126	0.090	-

Bold values represent statistically significant values

The demographic histories for the Madagascar and Cape Vidal populations of *N. lithophilus* in the SWIO were examined using mismatch distribution curves (Figure 9). Mismatch distributions of the Durban population could not be calculated due to the small sample size of this population. The Cape Vidal populations showed erratic curves which show no evidence of past demographic expansion (Harpending 1994). The mismatch distribution for the Madagascar population shows evidence of past demographic expansion, which is more evident in the estimation where 'Mad 146' was removed (Figure 9c). The raggedness index (RI) suggests that these populations have undergone past demographic expansion. The RI values for the Madagascar and Cape Vidal populations were estimated to be insignificant values 0.07 and 0.56, respectively (Table 8). Thus, this analysis failed to reject the hypothesis of past

demographic expansion in the *N. lithophilus* populations in the SWIO. The calculated Sum of Square Deviations (SSD) suggests that only the Cape Vidal population has undergone past demographic expansion, while the Madagascar population shows no evidence of demographic expansion. The SSD for the Madagascar population was estimated to be a significant value f 0.18, which demonstrates that the analysis rejected the hypothesis of past demographic expansion (Table 8). However, the SSD for the Cape Vidal population was estimated to be an insignificant value of 0.09, which shows that the analysis failed to reject the hypothesis of past demographic expansion (Table 8). The Bayesian Skyline plot of the effective female population size over time of *N. lithophilus* in all three sampled populations in the SWIO is shown in Figure 9. This graph shows that there has been a gradual increase in the effective female population size over the last 105 948 years in this region, with no sudden expansion events detected.

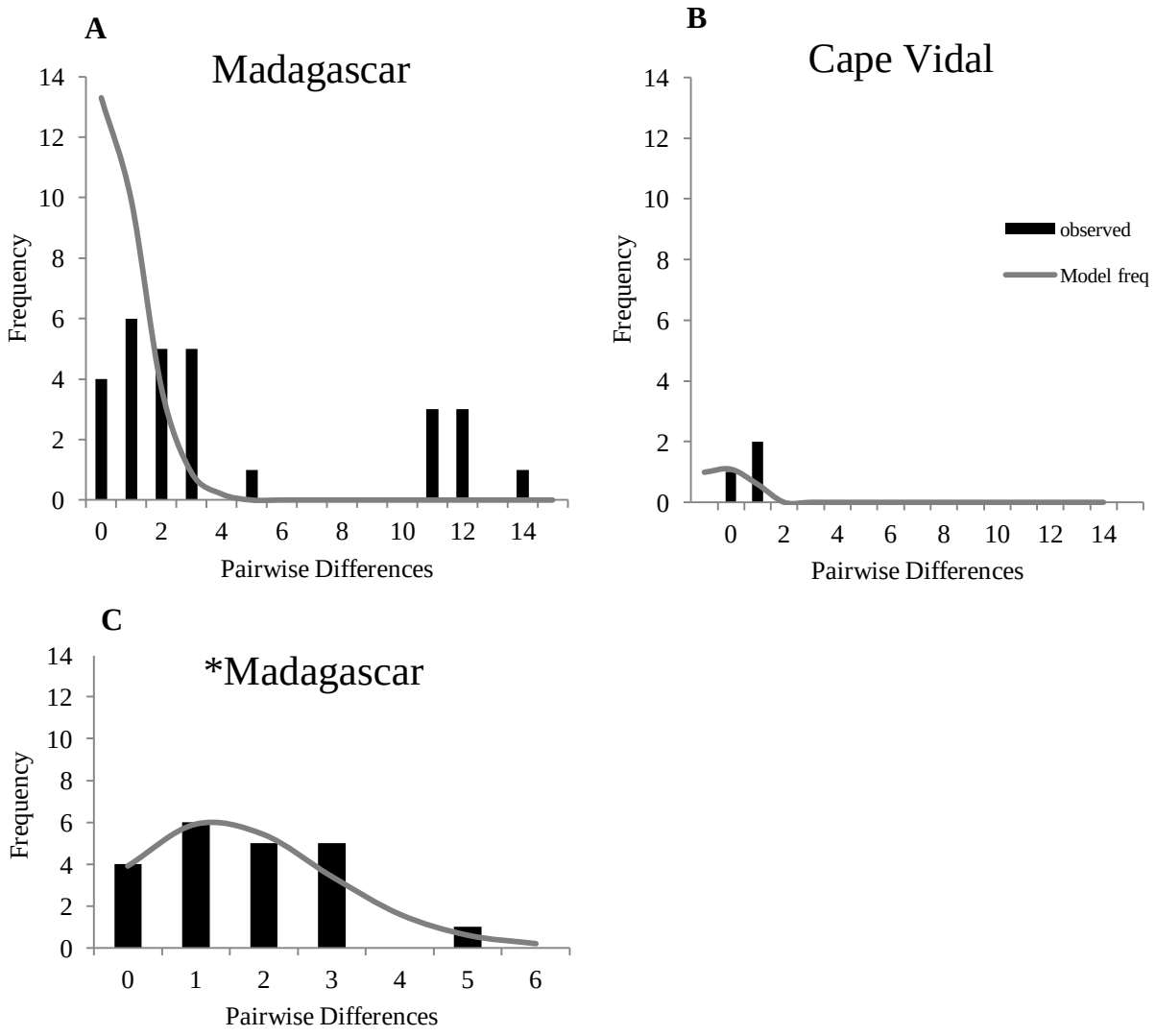

Figure 9: Mismatch distributions based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA showing the observed pairwise differences in the mtDNA (illustrated by the black bars on the graphs) and the expected expansion models (illustrated by the grey lines on the graphs) for *N. lithophilus* populations in the Southwest Indian Ocean.

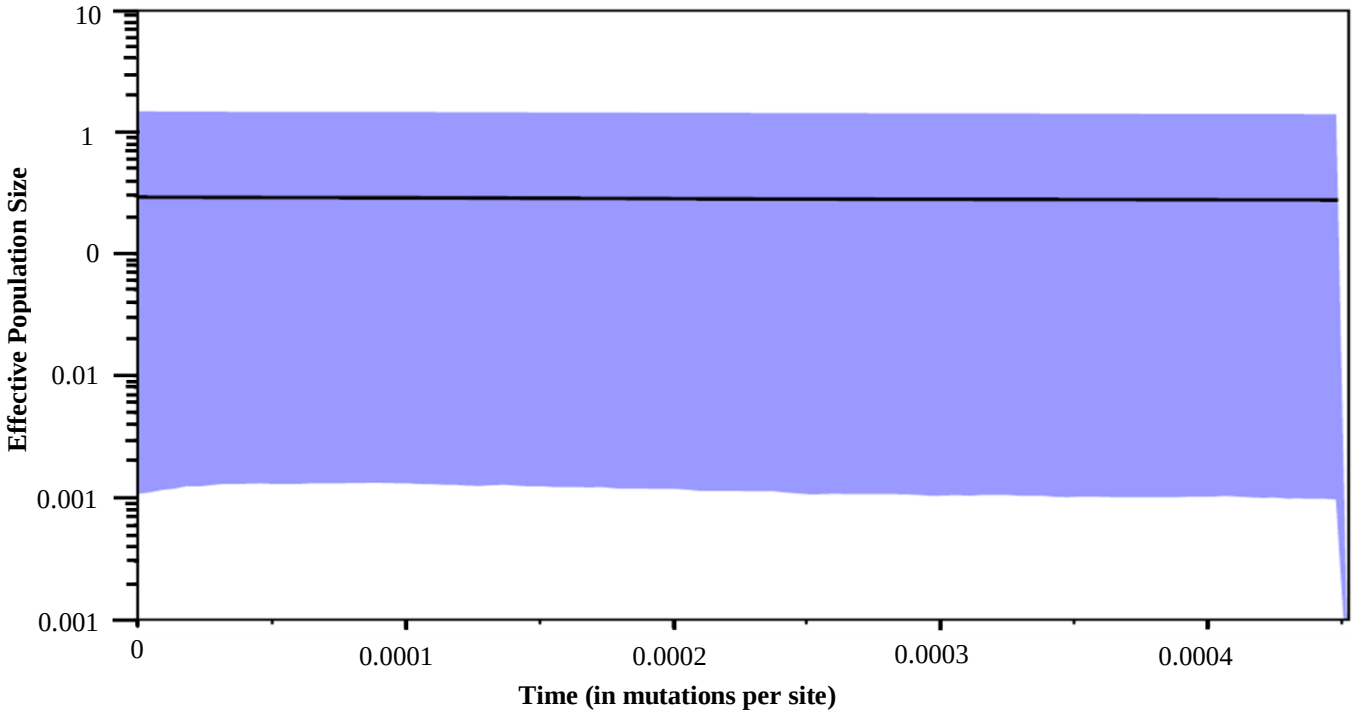

Figure 10: Bayesian Skyline Plot (BSP) based on the cytochrome oxidase subunit I (cox1) gene in the mtDNA showing the mean effective population size of female *N. lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal, Durban) over time. The blue shading on the graph represents the 95% confidence intervals of the calculated mean effective population size.

5.2 Discussion

5.2.1 Genetic Diversity

Genetic diversity of the cytochrome *c* oxidase subunit I (cox1) gene in the mitochondrial DNA of *D. capensis* populations in the SWIO was estimated to vary slightly. Madagascar and Durban demonstrated similar levels of genetic diversity (Table 1), however displayed differences in both haplotype composition and frequency (Figure 2). The Cape Vidal population was found to have the highest genetic diversity among the sampled populations (Table 1; Figure 2). High levels of haplotype diversity (0.700 to 0.822) were estimated for all three *D. capensis* populations. These levels of diversity were higher than those estimated for *D. capensis* by Henriques (2012) for the eastern coast of South Africa and are more comparable to those of the southern Angolan populations examined in Henrique's research. The nucleotide diversity of *D. capensis* in the SWIO (0.0021 to 0.0038) was also estimated to be higher than the diversity estimated by Henriques (2012) in *D. capensis* populations of southern Angola, Namibia and South Africa. This suggests that the populations in the KwaZulu-Natal region may have a higher genetic diversity than those of the southern coast of South Africa. However, the population sample size used in this study may have influenced the estimation of these molecular indices. Despite the geographic distance between them, Cape Vidal and Durban were estimated to vary greatly in genetic diversity (Figure 2). However, considering the number of populations sampled in this study, there may be missing populations between those sampled here that may give a more accurate representation of the pattern in genetic diversity in the SWIO.

The genetic diversity of *N. lithophilus* in the SWIO was found to vary among populations (Table 5). Haplotype diversity was estimated to range from 0.67 to 1.00, which is a fairly wide range. *Neoscorpis lithophilus* within the SWIO showed a wide range of nucleotide diversity, ranging from 0.0021 to 0.0088 (Table 5). This suggests that populations vary in genetic diversity between sampled populations. However, the sample sizes of the populations may call into question the validity of these genetic diversity indices, specifically for the Cape Vidal (n = 3) and Durban (n = 2) populations. *Neoscorpis lithophilus* individuals from the Madagascar population demonstrated the highest genetic diversity (Table 5; Figure 7). However, this high genetic diversity estimated for the Madagascar population is most likely a result of the sample 'Mad 146' which was estimated to be highly variant (see Figure 8). Highly variant individual samples have been demonstrated to reduce the ability to detect significant structuring between samples (Charlesworth 1998). This was most likely the case with the *N. lithophilus* data and the removal of this highly variant sample ('Mad 146'), allowed for a more accurate representation of the population structure.

The Cape Vidal population falls within the St Lucia Marine Reserve. High levels of genetic diversity have been suggested to illustrate the genetic importance of this region (von der Heyden 2009). The Cape Vidal population could be an important population for the diversity of the *D. capensis* individuals in the SWIO. For the *N. lithophilus* population, Madagascar demonstrates the highest genetic diversity (Table 5 and Figure 7). This could also suggest that populations on the southeast tip of Madagascar are genetically important. Unfortunately, not much research on genetic diversity of inshore fish species has been done in Madagascar, thus the genetic significance for fish species of this region in the greater SWIO is fairly unknown.

5.2.2 Population structure

The cytochrome oxidase subunit I (cox1) gene in the mitochondrial DNA of both *D. capensis* and *N. lithophilus* show no evidence for past genetic structuring between the populations in the SWIO. The haplotype networks (Figure 3 and 8), the phylogenetic tree (Figure 4) and AMOVAs (Table 2 and 6) all presented evidence that past populations of *D. capensis* and *N. lithophilus* were genetically similar with no structuring between South African and Madagascan populations. The haplotype network depicted ancestral haplotypes that are represented by individuals from all three populations in the centre of the network, from which all other unique haplotypes diverged from (Figure 3 and 8). This implies that there is incomplete lineage sorting for both *D. capensis* and *N. lithophilus*, where haplotype do not separate according to their geographic location (Zink et al. 2000). The formation of this pattern of lineage sorting indicates either that there is continuous gene flow between these regions or that this gene flow has ceased recently (Zink et al. 2000). The estimations of pairwise Φ_{pt} and F_{st} were not found to be significant (Table 3 and 7), thus the question of continuous gene flow could not be address with accuracy. This pattern of past panmictic populations in the SWIO has been illustrated in several marine species such as, *Xiphias gladius* (Muths et al. 2009); *Lutjanus kasmira* (Muths et al. 2012). This panmictic pattern of genetic structure is also evident in marine species in South Africa such as, *Chrysoblephus laticeps* (Teske et al. 2010), *Rhabdosargus holubi* (Oosthuizen 2006) and *Merluccius capensis* (von der Heyden et al. 2007). However, genetic structuring of marine populations in the SWIO has also been illustrated in other species: *Myripristis berndti* (Muths et al. 2011), *Palinurus delagoae* (Gopal et al. 2006), *Chelonia mydas* (Bourjea et al. 2007). Thus the historic genetic structure of marine populations in the SWIO is highly variable among species.

The lack of genetic structure illustrated by the populations of both *D. capensis* and *N. lithophilus* in the SWIO could be due to past gene flow between populations that was still evident in their mitochondrial DNA or that these regions have been colonized recently but gene flow has ceased and populations still share haplotypes. The estimated origins of the *Diplodus* genus eludes to a southward dispersal pattern from the East Atlantic to the southern tip of Africa through rapid colonization (Summerer et al. 2001). The mechanisms behind this divergence are largely unknown, however *Diplodus* individuals (either larvae or adults) would have had to cross the Lüderitz upwelling region in Namibia from southern Angola in order to colonise southern Africa, after which *D. capensis* speciated from the remnant *D. sargus* lineage. The upwelling cells within the Benguela system form a phylogenetic break for some species in this region, such as *Atractoscion aequidens* (Henriques et al. 2014). The ancestral colonizers would have then had to disperse south towards Cape Point, then up the eastern coast of South Africa, moving from the Atlantic Ocean into the Indian Ocean. This eastward flow of migration has been illustrated in marine fish of South Africa (von der Heyden 2009). In order to colonize Madagascar, individuals would have had to disperse across the Mozambique Channel and somehow to reach the Madagascan coastline. This illustrates the high dispersal ability of *D. capensis*. There is a paucity of information about the dispersal capability and larval swimming ability of *N. lithophilus*, however this research illustrates that past populations of this species were somehow able migrate across the Mozambique Channel to colonize either South Africa or Madagascar. Four tropical species from the Kyphosidae family (*Kyphosus bigibbus, K. sectatrix, K. cinerascens,* and *K. vaigiensis*) have wide global distributions that span from the Pacific Ocean to the Indian Ocean to the Atlantic Ocean, thus their dispersal is not hindered by known oceanographic barriers (Knudsen & Clements 2016). Knudsen and Clements (2016) illustrated that the remaining species of the *Kyphosus* genus are less widely distributed in comparison to these four species. The basal group of the *Kyphosus* genus, *K. azureus* and *K. cornelii*, have a concentrated distribution which is restricted to the eastern Pacific and the west

coast of Australia, respectively (Knudsen & Clements 2016). Though a restricted distribution range is also observed in that of *N. lithophilus*, this species could have either dispersed to either side of the Mozambique Channel.

5.2.3 Evolutionary history

The examination of the demographic histories of *D. capensis* and *N. lithophilus* presented evidence of demographic expansion and no demographic expansion (Figures 5 and 9; Tables 4 and 8). The demographic history of *D. capensis* illustrated past population expansion in the SWIO. This past demographic expansion was evident in the mismatch distributions of all populations, the estimations of the Raggedness Index (RI), as well as the Sum of Square Differences (SSD) (Figure 5 and Table 4). The estimation of Tajima's *D* was the only estimation of demographic history that suggested no demographic expansion, which is confirmed by the Bayesian Skyline Plot (Figure 6). However, Fu's Fs has been reported to be more sensitive to demographic expansion than Tajima's *D* (Fu 1997), thus this discrepancy had little influence on the outcome of this observation. The demographic history of *N. lithophilus* was slightly more difficult to discern. The mismatch distributions suggested that there was past demographic expansion (Figure 9) and the estimations of Tajima's *D*, RI, Fu's Fs and the Bayesian Skyline Plot (Figure 10) confirm this finding.

Considering the region in which this historic connectivity occurred is an aspect that must be examined when investigating the mechanisms behind this dispersal. The oceanography of the SWIO has been reviewed in Chapter Two. The position of both Africa and Madagascar, 84 – 66 Ma was 10° south of their current positions (Scotese et al. 1988, Yoder & Nowak 2006, Ali & Aitchison 2008). Ali and Huber (2010) suggested that the trajectory of ocean currents during this time was in an easterly direction instead of westerly as observed today, which would facilitate the dispersal of individuals from Africa to Madagascar. It has been suggested that these ocean currents also allowed the dispersal of terrestrial organisms through rafting from Africa to Madagascar (Agnarsson & Kuntner 2012). For fish species, this rafting dispersal could have been in association with floating object or aggregations in the water column. The trajectory of the Agulhas Current has been stable for the last 150 000 years (Winter & Martin 1990). Thus, it is quite likely that the trajectory of water masses in the SWIO have flowed in a north to south direction, as has been illustrated to flow today, for the last 150 000 years.

CHAPTER 6 – Contemporary population structure of *Diplodus capensis* and *Neoscorpis lithophilus* in the Southwest Indian Ocean based on microsatellites

6.1 Results

6.1.1 *Diplodus capensis*

A total of 2.4GB of read data was generated from the initial shotgun sequencing. A total of 6757 dinucleotide repeats, 290 trinucleotide repeats, and 253 tetranucleotide repeats were discovered using MSATCOMMANDER. A total of 30 primer sequences were designed and tested on the sampled populations of *D. capensis* in the SWIO. Of these, 24 primers amplified the targeted region of DNA and are listed in Table 9.

The results of the test for deviations from Hardy-Weinberg's expected equilibrium are also depicted in Table 9. Only six (DCMS 1, DCMS 2, DCMS 5, DCMS 6, DCMS 9 and DCMS 18) of the twenty-four microsatellites examined completely conformed to Hardy-Weinberg's expectations in all three populations examined. On the other hand, six loci (DCMS 3, DCMS 10, DCMS 12, DCMS 16, DCMS 29, and DCMS 30) were shown to deviate from Hardy-Weinberg's expectations, with *p*-values that were found to be lower than 0.05 in all three populations and were, therefore, excluded from further analyses. Three of these microsatellites (DCMS 10, DCMS 16 and DCMS 30) demonstrated a highly significant deviation from Hardy-Weinberg's Equilibrium. Thus, only 18 microsatellite loci were used for further analysis. A total of 33 *D. capensis* samples were used in this study. The linkage disequilibrium test estimated that seven loci (DCMS 5; 11; 13; 17; 21; 22; and 23) showed significant linkages to other loci across all populations. The Probability of Identity test conducted in GenAlEx estimates the probability of an identical multilocus genotype occurring between two unrelated individuals (Figure 11). This test showed that the Madagascar and Durban populations have a similar profile and thus, there is a similar probability that samples within each of these populations have the same multilocus genotype. The Cape Vidal population was shown to have a lower probability of identity on the left of this graph, which indicates that the likelihood of two unrelated individuals from this population having the same multilocus genotype is lower than that of individuals from the Madagascar and Durban populations having the same genotype. This is to be expected as the Cape Vidal population showed the highest genetic diversity among the sampled populations of this study.

Table 9: Characterization of the 24 amplified microsatellite loci used to examine the population connectivity of *Diplodus capensis* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban). These characters include: Product size range, mean number of alleles (N_a) and their standard deviation, observed and expected heterozygosity (H_o and H_e, respectively) and the *P*-value from the test of deviation from Hardy-Weinberg's equilibrium.

Locus name	Repeat motif	Product size range (in bp)	Mean N_a (± std. dev.)	Madagascar			Cape Vidal			Durban		
				H_o	H_e	*P*	H_o	H_e	*P*	H_o	H_e	*P*
DCMS1	$(AAAC)_8$	208 - 224	4.00 (± 1.0)	0.31	0.29	ns	0.58	0.49	ns	0.25	0.24	ns
DCMS2	$(AAC)_8$	221 - 230	3.33 (±0.6)	0.62	0.60	ns	0.67	0.68	ns	1.00	0.69	ns
DCMS3	$(AAG)_8$	192 - 204	4.00	0.23	0.71	***	0.25	0.48	*	0.38	0.79	*
DCMS4	$(AAT)_8$	114 - 120	3.67 (±1.2)	0.08	0.56	***	0.17	0.52	**	0.50	0.63	ns
DCMS5	$(AATC)_8$	104 - 124	6.67 (±0.6)	0.77	0.82	ns	0.75	0.76	ns	0.63	0.80	ns
DCMS6	$(AG)_8$	140 - 150	3.33 (±0.6)	0.77	0.64	ns	0.50	0.58	ns	0.75	0.57	ns
DCMS7	$(AGG)_8$	136 - 148	3.00 (±1.0)	< 0.01	0.15	*	0.17	0.31	ns	< 0.01	0.43	**
DCMS8	$(AT)_8$	136 - 140	2.67 (±1.2)	0.31	0.52	ns	0.08	0.37	**	0.13	0.33	ns
DCMS9	$(ATC)_8$	437 - 446	3.33 (±1.5)	0.38	0.46	ns	0.50	0.42	ns	0.38	0.33	ns
DCMS10	$(ATCC)_8$	144 - 168	4.00 (±1.7)	< 0.01	0.63	***	0.08	0.76	***	0.13	0.66	**
DCMS11	$(CCG)_8$	108 - 120	4.67 (±1.2)	0.62	0.75	ns	0.50	0.71	ns	0.38	0.74	**
DCMS12	$(AAGG)_9$	168 - 192	5.33 (±0.6)	0.15	0.41	**	0.33	0.75	**	0.25	0.78	**
DCMS13	$(ACAG)_9$	111 - 127	6.00 (±1.0)	0.23	0.79	***	0.50	0.84	ns	0.38	0.83	*
DCMS16	$(CG)_9$	162 - 176	6.33 (±0.6)	0.38	0.75	***	0.08	0.83	***	0.13	0.88	***
DCMS17	$(AAAC)_{10}$	137 - 162	7.00 (±1.7)	0.69	0.86	*	0.58	0.86	*	0.63	0.79	ns
DCMS18	$(ATC)_{10}$	157 - 169	5.33 (±0.6)	0.92	0.76	ns	1.00	0.76	ns	0.63	0.67	ns
DCMS19	$(ACC)_{11}$	126 - 138	4.00 (±1.0)	0.46	0.75	*	0.25	0.63	**	0.38	0.57	ns
DCMS21	$(AGC)_{11}$	117 - 147	6.67 (±1.5)	0.54	0.82	ns	0.33	0.76	**	0.38	0.89	**
DCMS22	$(AC)_{13}$	133 - 153	6.33 (±2.3)	0.62	0.73	*	0.42	0.86	***	0.63	0.75	ns
DCMS23	$(AG)_{13}$	90 - 112	6.33 (±2.1)	0.46	0.78	***	0.25	0.81	***	0.63	0.64	ns
DCMS24	$(AGC)_{13}$	163 - 175	5.00	0.31	0.75	***	0.58	0.78	ns	0.75	0.84	ns
DCMS25	$(AC)_{14}$	172 - 190	8.33 (±1.2)	0.62	0.83	*	1.00	0.91	ns	0.75	0.88	*
DCMS29	$(AC)_{18}$	130 - 154	9.00 (±2.6)	0.46	0.82	***	0.83	0.92	**	0.50	0.90	*
DCMS30	$(AAG)_{20}$	150 - 156	3.33 (0.6)	< 0.01	0.39	***	< 0.01	0.67	***	< 0.01	0.57	**

ns = not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

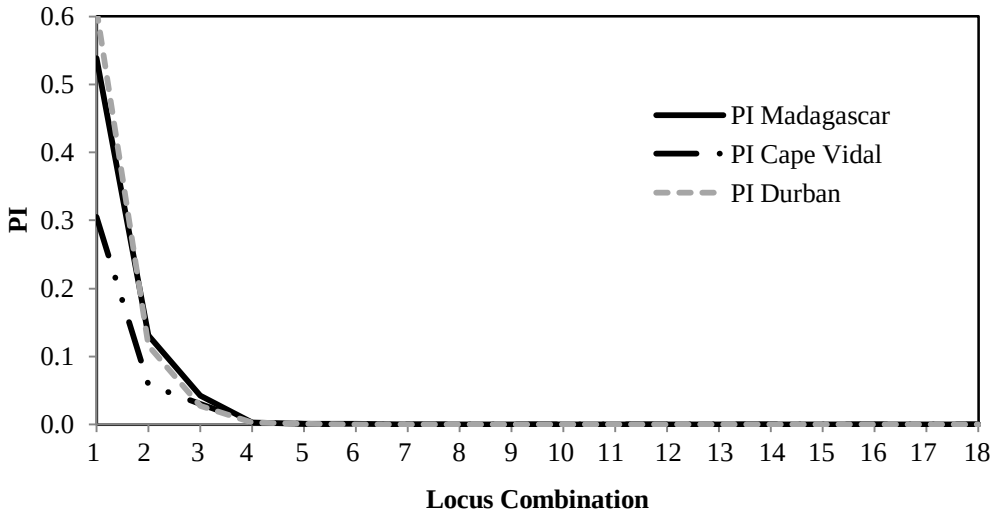

Figure 11: The Probability of Identity of multilocus genotypes of 18 microsatellites for each population of *D. capensis* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban).

Based on the allelic patterns observed within each population, which are illustrated by the bar graphs in Figure 12, the Madagascar and Cape Vidal populations had similar levels of genetic diversity. The Cape Vidal population, however, did show a slightly greater number of alleles, number of effective alleles, Shannon's Index and number of private alleles (Figure 12). The Durban population showed the lowest genetic diversity (Figure 12), although the differences between the molecular diversity indices for this population were still close to those calculated for the Cape Vidal and Madagascar populations.

The results of the population assignment test conducted in GENECLASS2 demonstrated that 91 % of all *D. capensis* individuals sampled were assigned to the populations from which they were sampled (Figure 12). The remaining 9 % of all sampled individuals were assigned to populations they were not sampled from, which suggests that several individuals are exchanged between populations. Though this is a small percentage of the sampled individuals, it still demonstrates evidence of recent gene flow between the sampled populations. Of the 13 individuals sampled from Madagascar, only one was assigned to a different population. This individual was assigned to the Cape Vidal population. Of the 12 individuals sampled from the Cape Vidal population, one individual was assigned to the Madagascar population. Of the eight individuals sampled from the Durban population, only one individual was assigned to the Madagascar population. No migrants from the Cape Vidal population were observed in the Madagascar or Durban populations.

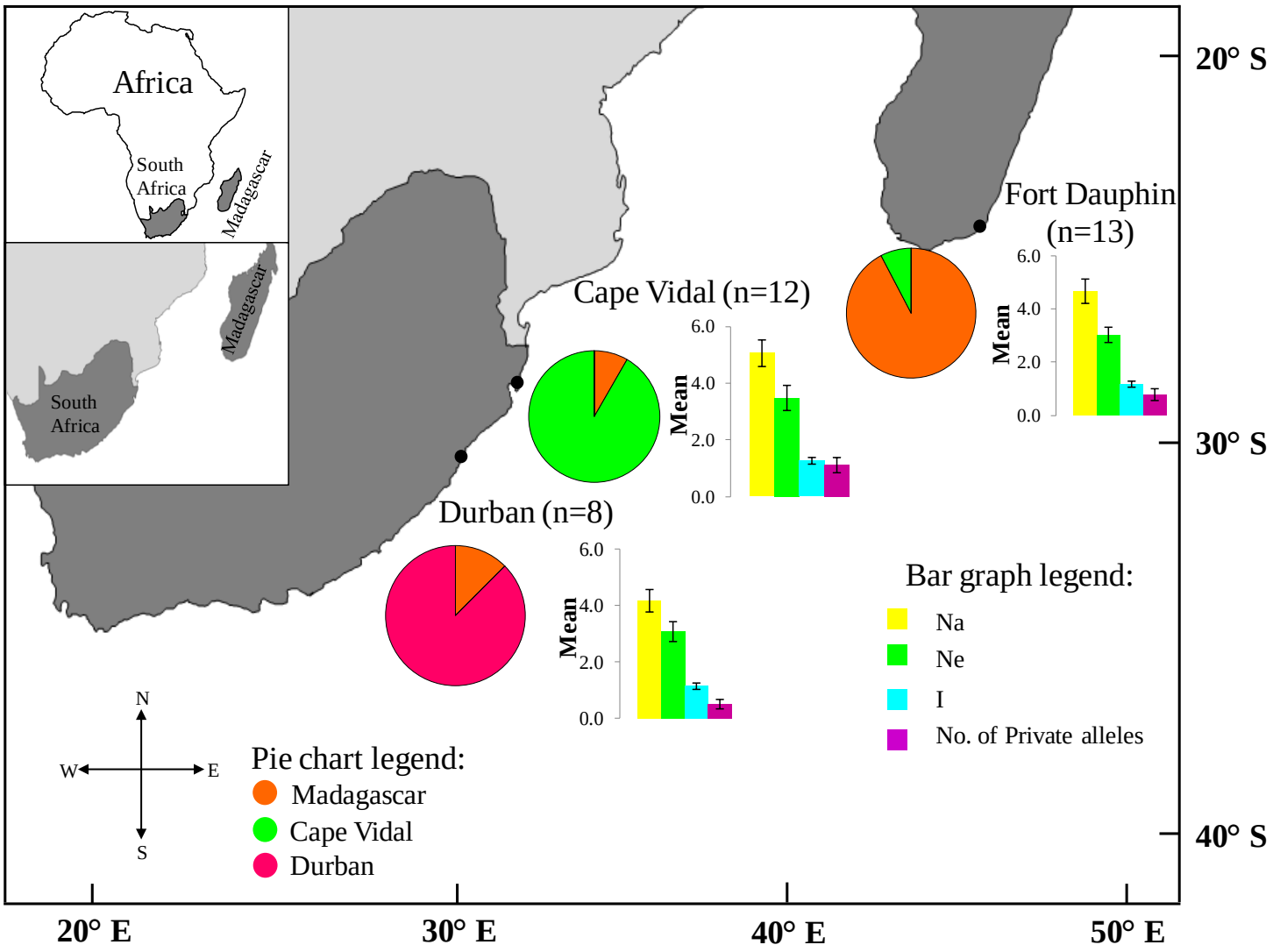

Figure 12: Population Assignment test constructed in GENECLASS2 based on 18 microsatellite loci of the *D. capensis* populations sampled in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban) are illustrated using pie charts. The colours represent the assigned populations within each population. The adjacent bar graphs illustrate the genetic diversity indices (Na - number of alleles; Ne - number of effective alleles; I - Shannon's Index; and Number of Private alleles) calculated for each population.

The Principle Coordinate Analysis (PCA) was conducted to illustrate the genetic difference between *D. capensis* samples in the SWIO. This analysis shows that the first component (PCO 1) accounts for 21.46% of the overall variance observed within the microsatellite data (Figure 13). The second component (PCO 2) and third components (PCO 3) accounts for 18.37% and 16.78%, respectively, of the overall variance. These three components collectively account for 56.61% of the observed variation among all the sampled individuals. Most of the samples are clustered in the middle of the plot which indicates that the eigenvalues used for this analysis have not been able to clearly separate the *D. capensis* samples. This suggests that *D. capensis* individuals are genetically similar to each other.

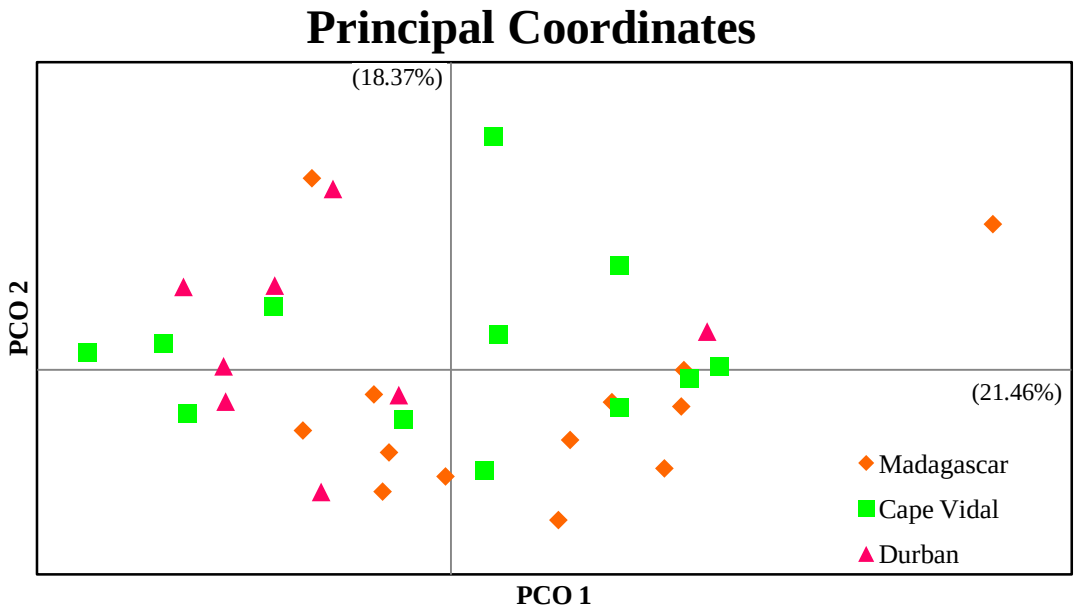

Figure 13: Results of the Principle Coordinate Analysis (PCA) demonstrating the first two components of genetic distance among *D. capensis* populations in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban) based on 18 microsatellites.

The Analysis of Molecular Variance (AMOVA) between populations of *D. capensis* in the SWIO was conducted using the R_{st} statistic and the F_{st} statistic (Table 10). The analysis conducted using the R_{st} statistic illustrated that there is no significant difference in genetic diversity (R_{st} = 0.04 with a *p*-value of 0.5) between populations of *D. capensis* in the SWIO. However, the analysis using the F_{st} statistic demonstrated that there is a significant difference (F_{st} = 0.034 with a *p*-value of 0.008) between *D. capensis* populations. The variation among populations accounts for 4.42% and 3.38% of the observed variation for both the R_{st} statistic and the F_{st} statistic, respectively (Tables 10). The variation within populations accounts for 95.58% and 96.62% of the observed variation for both the R_{st} statistic and the F_{st} statistic, respectively. The R_{st} statistic has been illustrated to be a weaker index of genetic variation (Neigel 2002). Thus, the F_{st} statistic was used to estimate the pairwise genetic differentiation among populations. All pairwise estimates of F_{st} and G_{st} for the populations of *D. capensis*, showed moderate levels of genetic differences between populations (0.024 – 0.049 and 0.139 – 0.172, respectively Table 11). The pairwise estimations of F_{st} between Madagascar and the remaining two populations were found to be significant which gives evidence for slight structuring between Madagascar and the South African populations (Cape Vidal and Durban). This pattern is also observed in the G_{st} estimations. However, pairwise estimation of F_{st} between Durban and Cape Vidal was not significant (*p*-value of 0.17).

Table 10: Results of the Analysis of Molecular Variance (AMOVA) using the R_{st} and F_{st} statistic calculated with ARLEQUIN showing the molecular variance in 18 microsatellite loci among and within all populations (Madagascar, Cape Vidal and Durban) of *D. capensis* within the Southwest Indian Ocean.

Statistic	Source of variation	df	Estimated variation	Variation %	Fixation Index
R_{st}	Among Populations	2	943.93	4.42%	0.04
	Within Populations	63	20397.16	95.58%	
F_{st}	Among Populations	2	0.208	3.38%	**0.034**
	Within Populations	63	5.939	96.62%	

Bold values illustrate statistically significant values ($p < 0.05$)

Table 11: The pairwise estimation of population differentiation of 18 microsatellites calculated using the F_{st} statistic (below the diagonal) and G_{st} statistic (above the diagonal) between populations of *D. capensis* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban).

	Madagascar	Cape Vidal	Durban
Madagascar		0.146	0.172
Cape Vidal	**0.030**		0.139
Durban	**0.049**	0.024	

Bold values illustrate statistically significant values ($p < 0.05$)

The Mantel test was used to test for isolation-by-distance by comparing genetic distance (based on pairwise F_{st}) and geographic distance (in km) between populations of *D. capensis* in the SWIO. This analysis demonstrated no significant relationship between geographic distance and genetic distance (r = 0.79; *p*-value = 0.169). This demonstrates that the distance between populations has no influence of the genetic variation observed between populations.

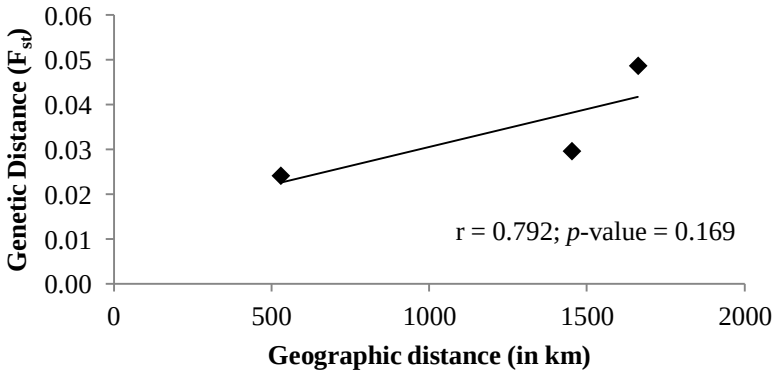

Figure 14: Correlation between geographic distances and genetic distances conducted using a Mantel test. The genetic distance was based on the pairwise estimations of F_{st} using 18 microsatellite loci, and geographic distance was based on the distance (in km) between three sample sites (Fort Dauphin, Madagascar; Cape Vidal and Durban, South Africa).

The Bayesian clustering analysis identified six clusters (K = 6) of *D. capensis* microsatellite genotypes when location is not used as a prior (Figure 15A), with a delta K value of 3.25. These genetic populations will be referred to as 'K' populations. The results of the mean Ln P(K) supports this estimation of K. The Bayesian clustering analysis that used location as a prior identified two genetic clusters (K = 2), with a delta K value of 4.57. The mean Ln P(K) values do not support this estimation of K, as a change in the trends between K = 1 and K = 2 were not very large.

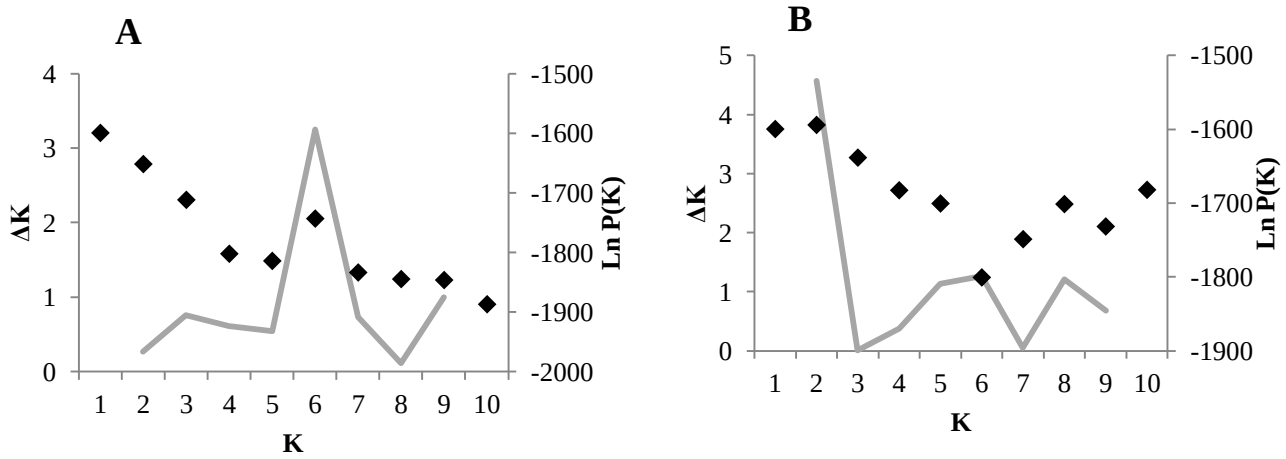

Figure 15: inferring the best K from the results of the Bayesian cluster analysis conducted in STRUCTURE using the Delta K method described by Evanno et al. (2005) based on 18 microsatellites of *D. capensis* populations in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban). The two runs where either the 'locprior' function was or was not used are depicted in the figures labelled 'A' and 'B', where 'A' illustrate the results where samples were not clustered according to location and 'B' illustrates the results of when they were. The line in these graphs depict the absolute value of $L''(K)$ for each estimation of K, while the black dots depict the average L(K) with each estimation of K.

Oddly, the admixture bar graphs do not support the findings mentioned above. The graph depicting six 'K' populations fails to congruently for six clusters based on the microsatellite DNA (Figure 15A). The admixture graph depicting the clustering of two populations, adequately illustrates the difference in genetic signatures (Figure 15B). Thus this analysis was inconclusive, as it failed to adequately cluster the microsatellite data for *D. capensis.*

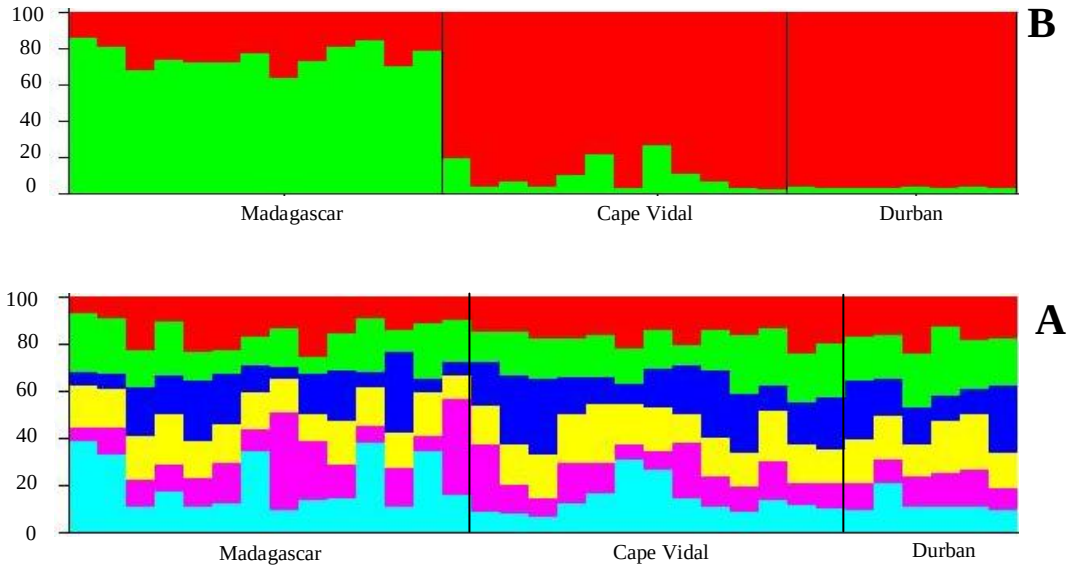

Figure 16: Bar graph constructed in STRUCTURE illustrating six 'K' populations (graph A) and two 'K' populations (graph B) observed within the *D. capensis* populations in the Southwest Indian Ocean. The probability of assignment for each sampled individual from Madagascar, Cape Vidal and Durban is represented by each bar, based on the genotype of 18 microsatellite loci.

6.1.2 *Neoscorpis lithophilus*

For *N. lithophilus*, unassembled data were used during the microsatellite discovery step due to the low number of primers outputs when the assembled data was used. A total of 5 723 trinucleotide and 14 863 tetranucleotide repeats were discovered during the microsatellite discovery step. A total of 9 246 primer sequences were designed by Primer3 for the discovered microsatellite region; however, these were filtered to remove repeat sequences. Of the 30 microsatellite primers tested, only 11 primers amplified the desired microsatellite regions (Table 12). The results of the test of deviation from the Hardy-Weinberg's expected equilibrium for each locus in each population found that only NLMS27 deviated in each population. This locus was therefore excluded from further analyses. No linkage disequilibrium among loci pairs was found.

A total of 44 *N. lithophilus* samples were used in this study. An estimation of the molecular diversity (number of alleles, number of effective alleles, Shannon's Index, number of private alleles) illustrated that the Madagascan population of *N. lithophilus* was the most genetically diverse amongst the sampled populations (Figure 17). The Durban population was found to be the least diverse as number of alleles, number of effective alleles, Shannon's Index

and number of private alleles was estimated to be the lowest in this population (Figure 17). The genetic diversity of the sampled populations of *N. lithophilus* seems to decrease from the northern - most (Madagascar) to the southernmost population (Durban).

The population assignment test conducted in GENECLASS2 demonstrated that 91% of the individuals were assigned to different populations; thus, the remaining 9 % of all sampled individuals were assigned to non-sampled populations (Figure 17). This analysis estimated that four samples present evidence of gene flow between populations. Two of these samples were from Durban, while the remaining two samples were from Cape Vidal and Madagascar. Of the 14 individuals sampled in Madagascar, only two individuals were assigned to different populations, the Durban and Cape Vidal populations. Of the 26 individuals sampled from Cape Vidal, one individual was assigned to the Madagascar population and one individual was assigned to the Durban population. None of the four individuals sampled from the Durban population was assigned to a different population.

Table 12: Characterization of the 11 amplified microsatellite loci used to examine the population connectivity of *Neoscorpis lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban). These characters include: Product size range, mean number of alleles (N_a) and their standard deviation, observed and expected heterozygosity (H_o and H_e, respectively) and the *P*-value from the test of derivation from Hardy-Weinberg's equilibrium.

Locus name	Repeat motif	Product size range (in bp)	Mean N_a (± std. dev.)	Madagascar			Cape Vidal			Durban		
				H_o	H_e	*P*	H_o	H_e	*P*	H_o	H_e	*P*
NLMS1	$(AACT)_8$	176 - 204	6.33 (± 1.2)	0.64	0.83	ns	0.62	0.76	ns	0.34	0.73	**
NLMS2	$(AATG)_8$	121 - 131	4.00	0.14	0.71	***	0.62	0.65	ns	0.13	0.69	***
NLMS8	$(AAG)_9$	189 - 205	6.00 (± 1.0)	0.29	0.85	***	0.12	0.77	***	0.13	0.61	***
NLMS9	$(AAT)_9$	230 - 239	3.33 (±0.6)	0.14	0.46	***	0.23	0.65	***	< 0.01	0.57	**
NLMS13	$(AGC)_9$	137 - 163	6.67 (±8.1)	0.57	0.93	***	-	-	-	0.13	0.24	ns
NLMS14	$(AGG)_9$	144 - 153	3.67 (±1.2)	0.57	0.47	ns	0.58	0.58	**	0.13	0.43	*
NLMS 15	$(AAG)_{10}$	120 -129	1.67 (±1.2)	< 0.01	0.26	**	-	-	-	-	-	-
NLMS 20	$(ACC)_{11}$	192 – 210	5.33 (±2.1)	0.93	0.80	ns	0.77	0.76	ns	< 0.01	0.43	**
NLMS 23	$(CCG)_{11}$	115 - 130	5.67 (±1.5)	0.50	0.80	*	0.54	0.69	**	0.25	0.44	*
NLMS 26	$(ATC)_{13}$	179 - 196	4.00 (±1.7)	0.36	0.52	ns	0.19	0.25	ns	< 0.01	0.40	*
NLMS 27	$(AAAG)_{14}$	92 - 108	4.00 (±1.7)	0.29	0.52	**	0.08	0.74	***	< 0.01	0.40	*

ns = not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

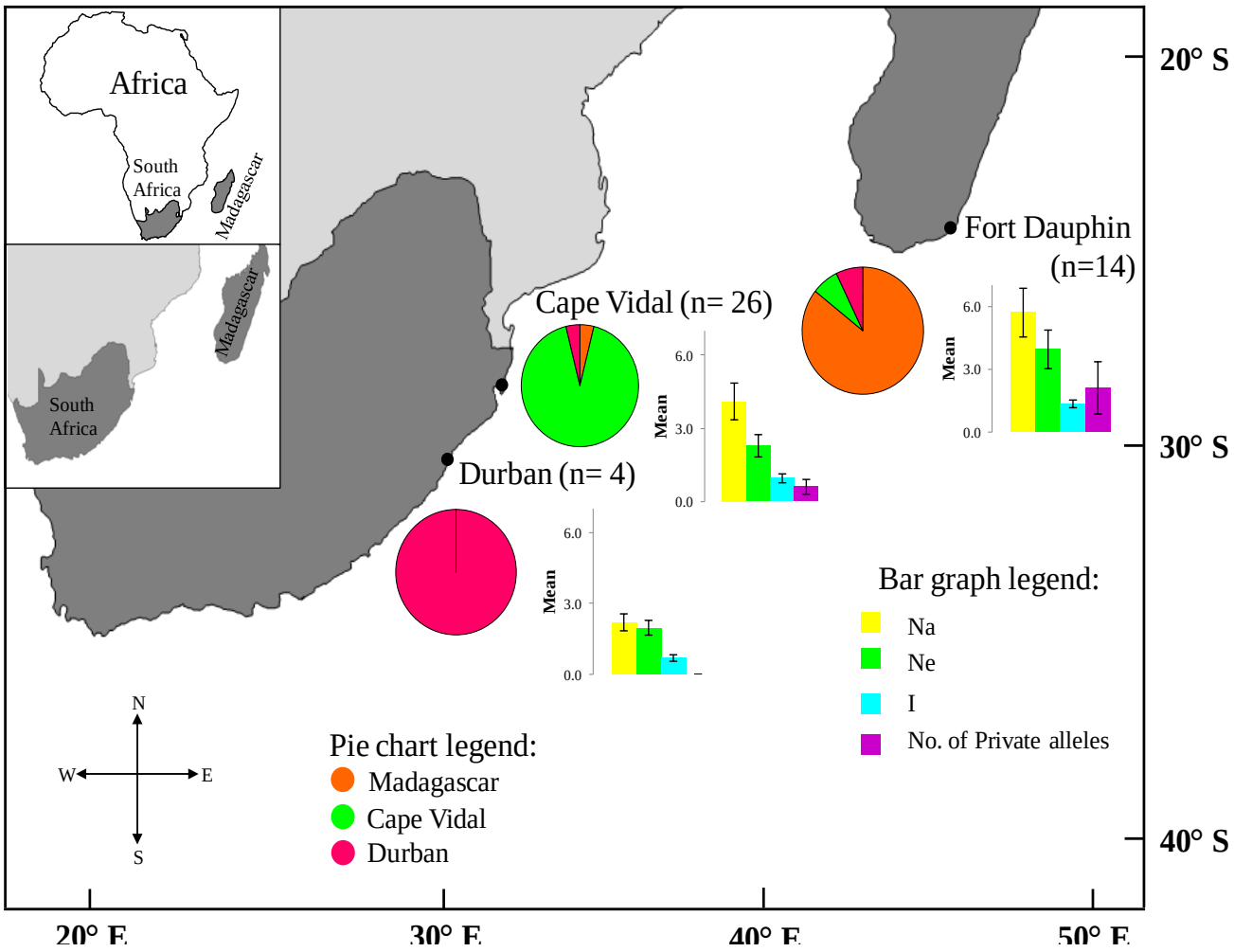

Figure 17: Population Assignment test conducted in GENECLASS2 based on 10 microsatellites of the *N. lithophilus* populations sampled in the Southwest Indian Ocean (Madagascar, Cape Vidal, and Durban) are illustrated using pie charts. The colours represent the assigned populations of samples within each population. The adjacent bar graphs illustrate the genetic diversity indices (number of alleles; number of effective alleles; Shannon's Index; and Number of Private alleles) calculated for each population.

The Principle Coordinate Analysis (PCA) was conducted to determine the genetic difference between *N. lithophilus* populations (Figure 18). These results illustrate a separation among *N. lithophilus* samples with the formation of two groups (labelled 'A' and 'B' in Figure 18) based on the first principle component. Both groups consist of individuals from all sampled populations. Cluster 'A' consists mainly of individuals from the Madagascar population with slightly fewer individuals from the Cape Vidal population. This analysis estimated that this separation was mainly based on the first principle component (PCO1), which accounted for 24.12% of the overall variation observed in *N. lithophilus* samples (Figure 18). The second principle component (PCO2) accounted for 19.77 % of the variation, while the third principle component (PCO3) accounted for 17.18 % of the overall variation. Together these three components account for 61.06 % of the overall observed variation within *N. lithophilus* individuals in this study.

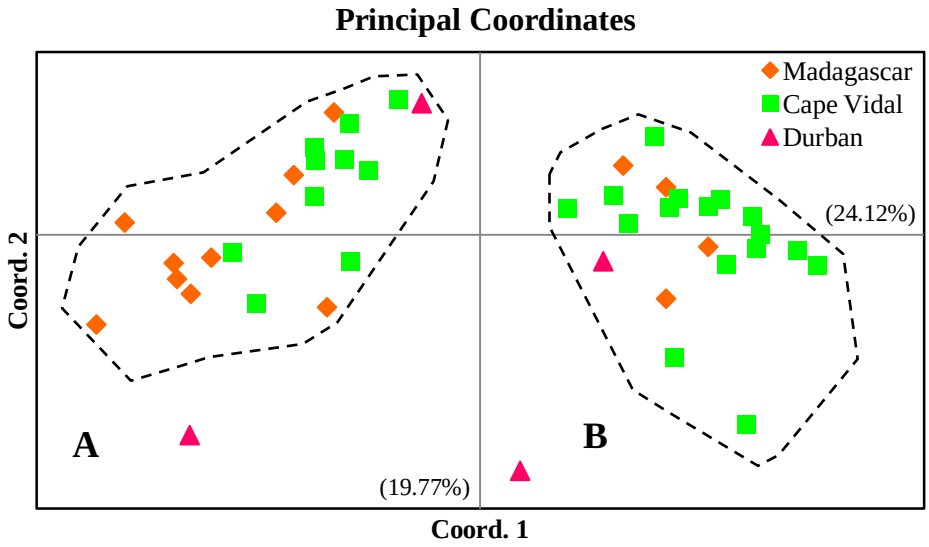

Figure 18: Results of the Principle Component Analysis (PCA) demonstrating the first two components of genetic distance among *N. lithophilus* populations in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban) based on 10 microsatellites. The two clusters formed are labelled 'A' and 'B'.

The Analysis of Molecular Variance (AMOVA) was used to test for genetic differentiation between populations of *N. lithophilus* within the SWIO using the R_{st} statistic and the F_{st} statistic. Both of these analyses demonstrate that there is significant differentiation between populations of *N. lithophilus* in the SWIO, with a significant R_{st} value of 0.37 (with a *p*-value less than 0.00005) and F_{st} value of 0.092 (with a *p*-value less than 0.00005) (Table 13). These analyses estimated that variation among populations accounted 36.8% and 9.2% of the observed variation for both the R_{st} statistic and F_{st} statistic, respectively. The variation within populations was estimated to account for 63.2% and 90.80% of the observed variation for both the R_{st} statistic and F_{st} statistic, respectively. The pairwise estimation of F_{st} and G_{st} demonstrate a significant level of structuring (Table 14). The lowest pairwise estimation of F_{st} was between Madagascar and Durban, with a statistically insignificant value of 0.07, while the lowest G_{st} was observed between Durban and Cape Vidal which was estimated to be 0.400. The highest values of both F_{st} and G_{st} were estimated to be between Madagascar and Cape Vidal (0.098 and 0.739, respectively). This suggests that there is slight structuring between Madagascar and Cape Vidal and between Cape Vidal and Durban. Thus, the Cape Vidal population is slightly isolated from both the Madagascar and Durban populations of *N. lithophilus.*

Table 13: Results of the Analysis of Molecular Variance (AMOVA) using the R_{st} statistic and F_{st} statistic using 10 microsatellite loci among and within populations of *N. lithophilus* within the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban).

Statistic	Source of variation	df	Estimated variation	Variation %	Fixation Index
R_{st}	Among Populations	2	10975.26	36.82%	**0.37**
	Within Populations	85	18829.02	63.18%	
F_{st}	Among Populations	2	0.288	9.20 %	**0.092**
	Within Populations	85	2.844	90.80 %	

Bold values illustrate statistically significant values

Table 14: The pairwise estimation of population differentiation using the F_{st} (below the diagonal) and G_{st} statistic (above the diagonal) based on 10 microsatellite loci between populations of *N. lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban).

	Madagascar	Cape Vidal	Durban
Madagascar		0.739	0.693
Cape Vidal	**0.098**		0.400
Durban	0.070	**0.087**	

Bold values illustrate statistically significant values (*P*-value < 0.05)

The Mantel test was conducted to test for a relationship between the genetic variation (based on pairwise F_{st}) and geographic distance (in km) between populations of *N. lithophilus* populations in the SWIO (Figure 19). This analysis failed to find a significant relationship between geographic distance and genetic distance. This suggests that geographic distance between populations fails to account for the genetic variation between them. Thus, populations of *N. lithophilus* in the SWIO are not isolated by geographic distance.

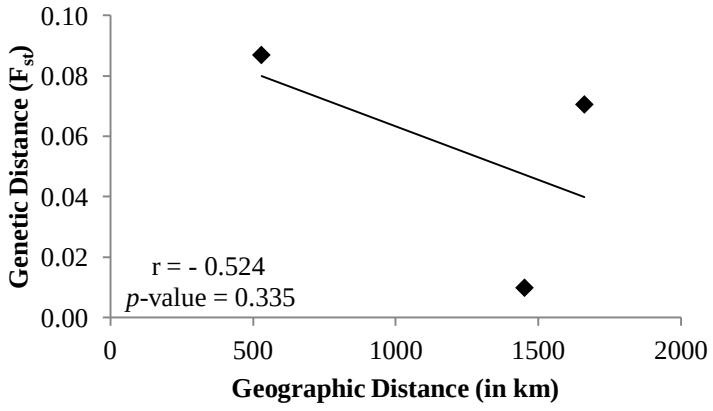

Figure 19: Mantel test conducted to determine the difference between geographic location (in km) and genetic distance (based on F_{st}) using 10 microsatellite loci in the *N. lithophilus* populations in the SWIO (Madagascar, Cape Vidal and Durban).

The Bayesian clustering analysis revealed that the most likely number of populations or the best estimate of 'K' is either 3 or 4 (Figure 20A and B), with Delta K values of 1.445 and. 2.071. The mean Ln P(K) estimations in Figure 20A, did not concur with the Delta K estimation of K. The Ln P(K) values of the clustering analysis conducted using location as a prior concur with the estimation of Delta K in this analysis and estimated that the best number of populations within the *N. lithophilus* samples to be three (Figure 20 B).The bar graph in Figure 21 did not support either of the estimations of K, as neither of the graphs illustrate the formation of groups that differ in their genetic signature. However, one could suggest that Figure 21B demonstrates the formation of three genetic clusters, however differ in their assignment to populations. T assignment to the population represented by the red colour seems to have occurred largely in samples from Madagascar. Samples from Cape Vidal and Durban showed the highest probability of being assigned to the population illustrated in green. The assignment to the populations represented in blue seems to have been the lowest in all populations. Each population seems to have a unique combination of these genetic signatures.

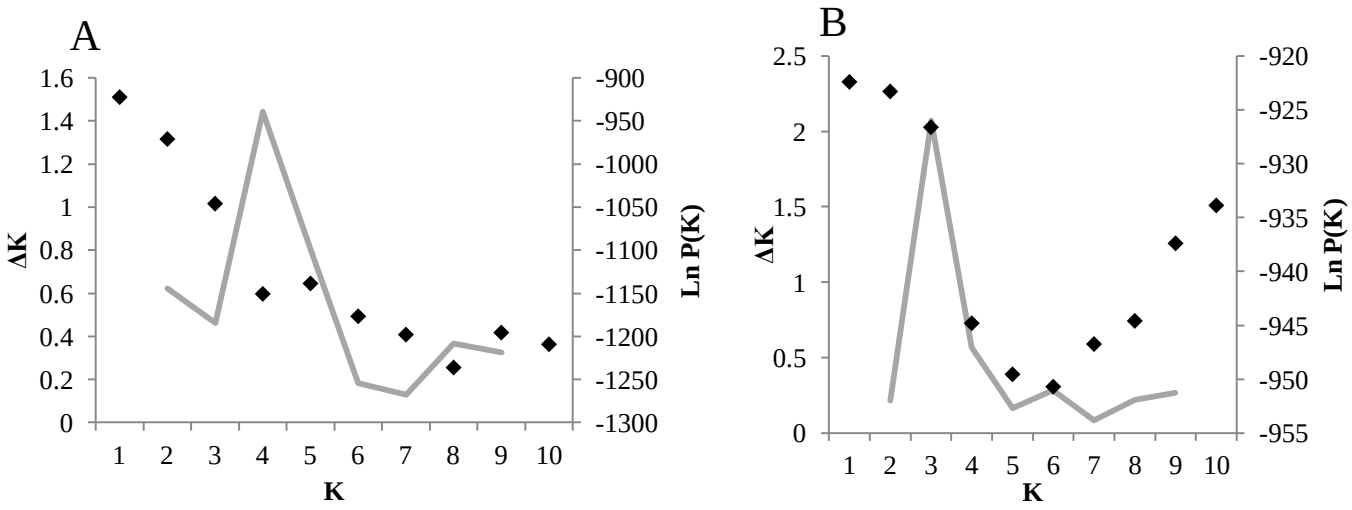

Figure 20: inferring the best estimate of the number of populations or 'K' using Delta K and Ln P(K) methods as described by Evanno et al. (2005) from the Bayesian clustering analysis of *N. lithophilus* in the Southwest Indian Ocean (Madagascar, Cape Vidal and Durban), based on 10 microsatellite loci. Graph A represents the results of the analysis run without location as a prior. Graph B represents the result of the analysis run with the location as a prior. The grey line on these graphs demonstrates the results of the Delta K analysis, while the black lines represent the mean Ln P(K) for every estimation of K.

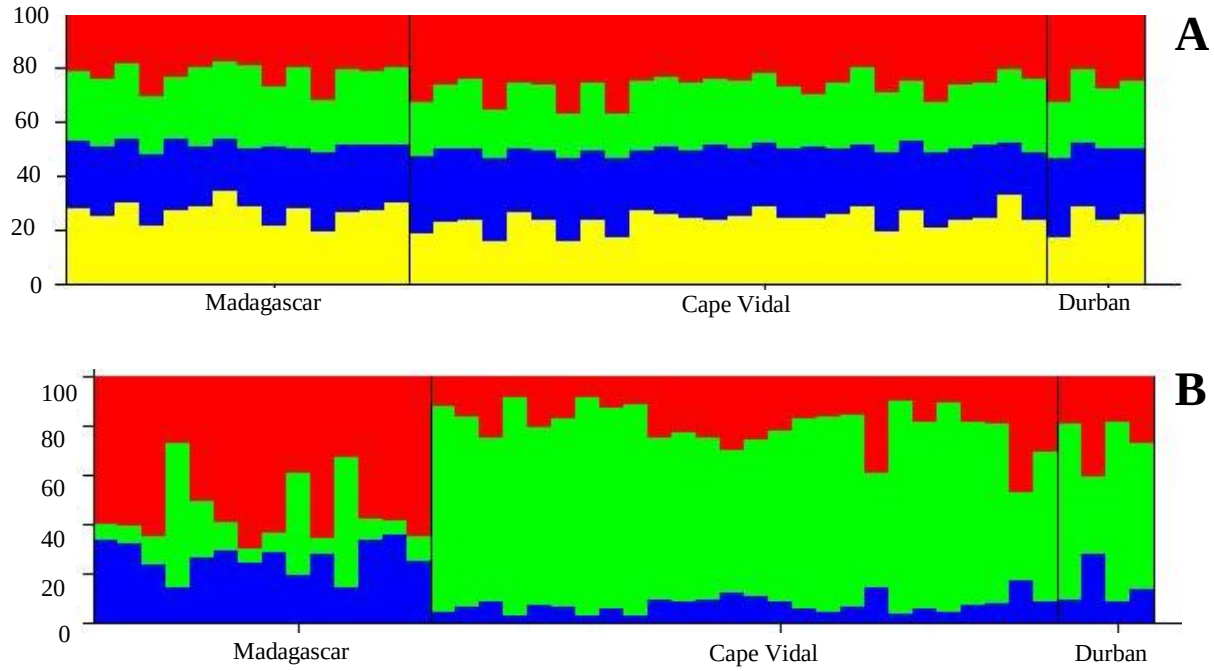

Figure 21: Bar graph constructed in STRUCTURE illustrating the genetic clusters (K = 3 and 4) observed within the *N. lithophilus* populations in the Southwest Indian Ocean. Graph A depict the clustering of samples where the analysis did not use location as a prior of clustering. Graph B illustrates the result of the analysis where location was used as a prior to cluster samples. Each bar represents the probability of each sampled individual from Madagascar, Cape Vidal and Durban, being assigned to a certain cluster, based on the genotype of 10 microsatellite loci. The different colours represent the number of assumed populations.

6.2 Discussion

6.2.1 Contemporary genetic diversity

The genetic diversity of the microsatellites used to examine the connectivity of *D. capensis* and *N. lithophilus* in the SWIO was estimated to be highly variable across all loci (Tables 9 and 12). For the eighteen microsatellites amplified in *D. capensis,* the average observed heterozygosity (H_o) was estimated to range from 0.06 to 0.85 and estimates of the average expected heterozygosity (H_e) ranged from0.30 to 0.87. For the ten microsatellites amplified in *N. lithophilus,* the average observed heterozygosity was estimated to range from 0.18 to 0.85 and estimates of the average expected heterozygosity ranged from 0.26 to 0.77. Though these ranges for *N. lithophilus* are quite wide, it should be noted that almost all estimations of both observed and expected heterozygosity were moderately lower (less than 0.5 for H_o; less than 0.7). These estimations of genetic diversity are lower than most of those estimated for other marine fish by previous authors such as: *D. capensis* H_o = 0.77 – 0.84 (Henriques 2012); *D. capensis* H_o = 0.18 – 1.00 (Reid et al. 2012); *D. sargus* H_o = 0.51 – 0.79 (Perez et al. 2008); *Lithognathus lithognathus* H_o = 0.25 – 0.90 (Reid et al. 2012); *Rhabdosargus holubi* H_o = 0.69 - 0.85 (Oosthuizen 2006); *Chysoblephus laticeps* H_o = 0.78 – 0.89 (Teske et al. 2010).

Diplodus capensis and *N. lithophilus* demonstrated different patterns of genetic diversity across populations in the SWIO (Figures 12 and 17). Though, the genetic diversity of all three populations of *D. capensis* (Madagascar, Cape Vidal and Durban) were quite similar, the Cape Vidal population demonstrated the highest diversity. On the other hand, the Madagascar population demonstrated the highest genetic diversity for *N. lithophilus* populations. This implies that these populations are important for the genetic diversity of these species. However, the *N. lithophilus* sample size of the Durban population was quite small and thus these results fail to give accurate representation of the genetic diversity of this population.

6.2.2 Contemporary population structure and connectivity

6.2.2.1 Population structure in the SWIO

Populations of both *D. capensis* and *N. lithophilus* in the SWIO both illustrated patterns of fine-scale structuring. The microsatellite DNA used to determine the contemporary genetic structure of *D. capensis* illustrated that populations were genetically different from each other. This fine-scale genetic structure was supported by the AMOVA (Table 10), and the estimations of pairwise F_{st} (Table 11). This suggests that there is restricted genetic exchange between populations of *D. capensis* in the SWIO. However, the results of the PCA failed to illustrate this genetic difference between populations and demonstrated no difference between the samples across multiple loci (Figure 13).The admixture graphs also failed to demonstrate any conclusive clustering between samples of *D. capensis* The microsatellite DNA used to determine the contemporary genetic structure between *N. lithophilus* populations demonstrated fine-scale structuring which was evident in the PCA (Figure 18), the AMOVA (Table 13), the pairwise F_{st} (Table 14). This means that populations of *N. lithophilus* in the SWIO are genetically different from each other. Genetic structure between populations could be a result of limited gene flow

between populations which results in the accumulation of local mutations. A barrier to recent dispersal could result in such fine-scale structure (Palumbi 2003).

The Bayesian clustering analyses illustrated that more than one genetically unique population of both *D. capensis* and *N. lithophilus* is present in this study. This analysis suggested that *D. capensis* samples cluster into two genetically distinct groups (Figure 15B) while the *N. lithophilus* samples cluster into three distinct populations (Figure 20B). However the admixture graph failed to successfully depict these genetic clusters. The geographic distance between South Africa and Madagascar could have influenced the genetic variation observed between these two regions and was tested using the Mantel test. The Mantel test demonstrated that distance between populations of both *D. capensis* and *N. lithophilus,* had no significant influence on the genetic differentiation (Figure 14). Thus, the genetic structure observed was not due to the geographic distance between populations.

Available literature on the connectivity of marine organisms within the SWIO is quite low (Ridgway & Sampayo 2005). However, the literature that is available suggests that historic and contemporary connectivity within this region varies greatly between species. Some authors present evidence for connectivity, as seen in *Lutjanus kasmira* (Muths et al. 2012), *Chrysoblephus puniceus* (Duncan et al. 2015) and *Xiphias gladius* (Muths et al. 2009), while others present evidence suggesting that there is limited connectivity at some level in this region, as seen for *Myripristis berndti* (Muths et al. 2011), *Penaeus monodon* (Forbes et al. 1999); *Scarus ghobban* (Visram et al. 2010) and *Chelonia mydas* (Bourjea et al. 2007). Genetic structure is influenced by both the characteristics of the organism, as well as the environment in which it lives (Lowe et al. 2009). As these organisms are distributed across the same region, one would assume that the functions of genetic structure between them would be similar. Thus, the differences in genetic structure between them may be a result of their difference in biological characteristics.

The pairwise estimation of F_{st} was used to determine where this genetic structuring was observed. The estimation for *D. capensis* illustrated that significant genetic structure is observed between Madagascar and Cape Vidal and between Madagascar and Durban (F_{st} = 0.03 and 0.49, respectively; Table 11). The estimation for *N. lithophilus,* also illustrates that the genetic structure is between Madagascar and Cape Vidal and between Cape Vidal and Durban (F_{st} = 0.10 and 0.07, respectively; Table 14). This contradicts the results illustrated by the admixture bar graph in Figure 21, which suggests that Cape Vidal and Durban originated from the same ancestral population. The estimation F_{st} for these comparisons was greater than 0.025, which suggests that less than 10 migrant individuals have been exchange between these populations (Palumbi 2003). Estimations of F_{st} between a metapopulation may be influence by both direct and indirect connections between populations, such as indirect connections between a source population and another, more isolated population through other populations within the network (Neigel 2002). Similarly, estimations of G_{st} demonstrated that there is genetic structuring between all three populations with high values estimated for both *D. capensis* and *N. lithophilus* (0.139 - 0.172 and 0.4 – 0.739). An estimation of G_{st} of 0.01 suggests that 20 – 50 migrants are exchanged between populations (Palumbi 2003) Based on the G_{st} value estimated in this study, both *D. capensis* and *N. lithophilus* populations exchange less than twenty individuals.

6.2.2.2 Connectivity in the SWIO

The population assignment test for *D. capensis* and *N. lithophilus* assigned three and four individuals, respectively, to populations that they were not sampled in (Figure 12 and 17). This indicates that individuals of these two fish species are able to cross the Mozambique Channel in both an eastward and westward direction. This is supported by the observation that eddies which are generated off the southeast tip of Madagascar travel westward to meet the KwaZulu-Natal coast at around Sodwana Bay (Morris et al. 2013). Hancke et al. (2014) has illustrated that particles can be transported from Madagascar to Mozambique, as well as from Mozambique to Madagascar by using surface drifters, as mentioned in Chapter 2. The result of the Population Assignment test conducted in GENECLASS2 for *D. capensis* also illustrated that individuals are able to travel from Madagascar to regions south of Sodwana Bay, suggesting that the dispersal range of this species along the KwaZulu-Natal coast extends at least 529 km to include both Cape Vidal and Durban. The sample sizes of the Durban population for both *D. capensis* and *N. lithophilus* are small and may have influenced the outcome of the Population Assignment test by limiting the number of individuals used to estimate the multilocus genotype for this population; this would in turn affect the assignment of individuals to this population.

Approximately 9.1% of *D. capensis* individuals and 8.3% of the *N. lithophilus* individuals within populations in the SWIO show evidence of recent genetic exchange. The low percentage genetic exchange into populations may not be enough to allow sufficient homogenizing of allele frequencies across the metapopulation, thus resulting in the fine-scale structuring observed in this study. This low percentage also suggests that the dispersal events across the Mozambique Channel are not frequent (Mokhtar-Jamaï et al. 2011). Mokhtar-Jamaï et al. (2011) studied the connectivity of the red gorgonian (*Paramuricea clavata*) in the Mediterranean Sea and found limited dispersal of individuals between populations and structure between populations. This could also be the case for *D. capensis* and *N. lithophilus* population in the SWIO.

Both *D. capensis* and *N. lithophilus* demonstrate similar patterns of genetic structure between Madagascan and South African populations, though the data for *N. lithophilus* was more compelling. This could either present evidence that similar factors influence the connectivity of inshore fish species in the SWIO. However, this could also be discounted as the genetic structure of only two species was presented in this study and their population structures could be unrelated. In the case of the reef fish *Myripristis berndti*, Muths et al. (2011) demonstrated that there is a level of restricted connectivity between populations of this species in the SWIO despite its dispersal capability. Similarly, *D. capensis* is a wide-spread species, however the results presented in this study illustrate that there is slight restriction to connectivity in the SWIO. *Neoscorpis lithophilus* has a restricted distribution range which suggests that this species may have a lower dispersal capability than *D. capensis*, though some level of migration does occur within this species. This study did, however, only concentrated on populations in KwaZulu-Natal, South Africa and the southern tip of Madagascar, and not the entire distribution of these species.

Ultimately, the level of gene flow between populations is dictated by the mobility of the species, its gamete or larval dispersal ability and levels of isolation between populations, which can be temporal, physical and/or ecological (Lowe et al. 2009). The pathway through which successful migrant individuals have travelled should be evident from their genetics and should provide an indirect measure of population connectivity (Hellberg et al. 2002). Measuring the gene flow between populations allows for the estimation of local adaptations and local changes over time within populations (Whitlock & McCauley 1999). However, it must be considered

that the estimation of the number of first generation migrants is not a direct measure of gene flow (Bohonak 1999, Whitlock & McCauley 1999). For example, these migrants may be unable to reproduce and thus fail to contribute to the gene pool of the next generation in their current population (Bohonak 1999, Whitlock & McCauley 1999). Though these migrants may not contribute to the population genetically, they may have indirect effects on the resident population, such as increased competition for resources or the introduction of new diseases (Whitlock & McCauley 1999). Thus, the number of migrants per generation within any population may be estimated, however the genetic contribution to the next generation of these individuals is unknown. Similarly, estimations of high gene flow between populations do not translate directly to a high number of migrants, as this high genetic exchange may be represented by a small number of migrants (Palumbi 2003). The successful dispersal of individuals occurs when they successfully reproduce to produce offspring that grow to adulthood which represents their contribution to the gene pool of the next generation (Whitlock & McCauley 1999, Hellberg 2009). This is probably not the case for *D. capensis* and *N. lithophilus* populations in the SWIO. The limited genetic exchange between populations of *D. capensis* and *N. lithophilus* in the SWIO could be an indication that successful dispersal between populations may not occur in this region.

Though the information gathered through this study has demonstrated that gene exchange occurs between populations in contemporary times, it however does not elucidate the mechanisms used during this dispersal. Ocean currents have been presented a potential explanations for genetic patterns of connectivity in marine populations (White et al. 2010). The oceanography of the SWIO consists of eddies that move water through the Mozambique Channel as well as off the southern tip of Madagascar, as discussed in Chapter 2. Hancke et al. (2014) showed evidence of the movement of water masses between eddies that travel from the southeast tip of Madagascar to Mozambique. This could be one way through which these individuals are transported from one population to another across the Mozambique Channel. About four of these eddies are formed a year (de Ruijter et al. 2004). The low frequency of these eddies per year potentially influences the contemporary migration of individuals across the Mozambique Channel by limiting the availability of the migration pathways. This would explain the low number of dispersal events suggested by the population assignment tests (Figures 12 and 17). Spawning of *D. capensis* fish occurs over 7 months (May to December) and 6 months (July to January) for *N. lithophilus* in the South African populations (Joubert 1981, Coetzee 1986, Mann & Buxton 1998, Whitfield 1998, Beckley 2000, Pattrick & Strydom 2008, Strydom 2008, van der Elst 2012). The use of the dispersal pathway presented by these eddies by these fish larvae would thus depend on the timing of eddy development. The larval stage of the fishes, which last for 45 days for *D. capensis*, is another factor that could influence the frequency of dispersal events across the Mozambique Channel. Considering the research conducted by Hancke et al. (2014) which illustrated the travel time of particles across the Mozambique Channel (fastest transport time = 15 – 85 days from Mozambique to Madagascar), the larval stage of *D. capensis* may be too short to allow the larvae to travel from one population to another before they need to recruit. This may result in larvae dying before they are able to recruit to a suitable habitat.

CHAPTER 7 – CONCLUSIONS

7.1 Main conclusions

The Southwest Indian Ocean (SWIO) is a unique system in terms of both its oceanography and the marine species found there, that show a high level of endemism (de Ruijter et al. 2005, Griffiths et al. 2010). The highly variable ocean currents occur within the Southwest Indian Ocean due to the position of Madagascar (Penven et al. 2006). Despite this highly variable marine environment, identical coastal species have been described in Mozambique, South Africa and Madagascar, including *D. capensis* and *N. lithophilus* (van der Elst 2012). The present study was aimed at examining the connectivity of *D. capensis* and *N. lithophilus* in the SWIO. The results of this study would give insight to the connectivity of both widely distributed coastal marine fish such as *D. capensis* and more endemic coastal species with a limited distribution such as *N. lithophilus* in the SWIO.

The evolutionary history of *D. capensis* and *N. lithophilus* populations in the SWIO explored using mtDNA showed no evidence of genetic structuring. This illustrated that past populations of *D. capensis* and *N. lithophilus* in SWIO were genetically similar with incomplete lineage sorting between populations. The contemporary investigation of population structure of *D. capensis* and *N. lithophilus* using microsatellites in this region demonstrates that these populations are difference from each other. This could mean that historic populations exchange genetic information at a rate that allowed the homogenization of allele frequencies and a recent barrier, possible due to oceanographic currents, to dispersal has resulted in recent differentiation between populations. Thus, this discrepancy could be explained by the difference in the time frame revealed by these two markers. Mitochondrial DNA is maternally inherited, thus only reveals the evolutionary history of females within and among species while microsatellites undergo recombination, thus illustrating reveal a pattern of genetic structure for both sexes (Hurst & Jiggins 2005). Thus, the difference in genetic structure patterns could be a result of sex-bias dispersal. However, the potential of adult dispersal is low for *D. capensis*, as it has been shown to be a resistant species (Joubert 1981, Attwood & Bennett 1995, Mann 2013), thus any possible dispersal would occur during their larval phase. The adult dispersal of *N. lithophilus* is unknown, thus available information about this species may limit the interpretation of the lack of genetic structure estimated in this study.

The fine-scale structure observed for *D. capensis* and *N. lithophilus* is surprising, given the available information on the trajectory of ocean currents in the SWIO showed evidence that water masses in this region are able to cross the Mozambique Channel and provide possible pathways of transport between Madagascar and South Africa, as reviewed in Chapter 2. This study did also illustrate that there is evidence of weak genetic exchange between South African and Madagascan populations, which however, did not allow the sufficient homogenizing of allelic patterns across the metapopulation which was illustrated by the low percentage of potential dispersal. The frequency of the formation of eddies off the southern tip of Madagascar could also not be sufficient to allow the formation of continuous pathways of connectivity. The properties of the fish biology, such as the duration of the larval stage and spawning season, could also influence the changes of the larvae using the pathways of connectivity across the Mozambique Channel.

Despite the differences in dispersal capability of *D. capensis* and *N. lithophilus*, they show the same patterns of connectivity in the SWIO. This could be the same for other coastal fish that are similar in biology and distribution in this region. It is clear that there may be some

recent genetic exchange between South Africa and Madagascar but allelic patterns of these populations have begun to diverge from each other. It is also clear that the oceanography in the SWIO has the potential to help explain patterns of connectivity in coastal fish species within this region. However, the biology of these marine species and it's interaction with the oceanography may be more complex than previously anticipated which could have affected recent patterns of connectivity within the SWIO.

7.2 Future research

One of the shortcomings of this study is that only three populations of *D. capensis* and *N. lithophilus* were included and used to draw conclusions on connectivity in the whole SWIO. Thus more populations from Madagascar, Mozambique, South Africa and locations throughout the distribution range of these marine fish are required to gain a more holistic view of the connectivity in this region. Once this is done, the genetic source populations can be determined. The number of microsatellite loci used, especially for *N. lithophilus*, can be increased to produce more robust data and estimation of genetic indices. More research into the reproduction, population dynamics such as sex ratio, as well as larval dispersal, recruitment, behaviour and swimming ability are required to better understand the underlying mechanisms behind the patterns of genetic structure observed in this study. This is especially required for *N. lithophilus* populations, where such information is unknown. The movement of eddies in the SWIO potentially plays a major role in the dispersal of marine species in this region. The directionality of the available pathways across the Mozambique has been investigated, however where or not marine species use these pathway has not yet been established. This type of investigation may require systematic sampling of eddies on either side of the Mozambique Channel.

REFERENCES

Adcodk G, Bernal Ramázrez J, Hauser L, Smith P, Carvalho G (2000) Screening of DNA polymorphisms in samples of archived scales from New Zealand snapper. Journal of Fish Biology 56:1283-1287

Agnarsson I, Kuntner M (2012) The generation of a biodiversity hotspot: biogeography and phylogeography of the western Indian Ocean islands. Current topics in phylogenetics and phylogeography of terrestrial and aquatic systems Rijeka: In Tech Publishers:33-82

Alerstam T, Hedenström A, Åkesson S (2003) Long-distance migration: evolution and determinants. Oikos 103:247-260

Ali JR, Aitchison JC (2008) Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). Earth-Science Reviews 88:145-166

Ali JR, Huber M (2010) Mammalian biodiversity on Madagascar controlled by ocean currents. Nature 463:653-656

Allsop DJ, West SA (2004) Sex-ratio evolution in sex changing animals. Evolution 58:1019-1027

Amoroso E (1960) Viviparity in fishes Symp Zool Soc London, p 153-181

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Molecular Biology and Evolution 13:494-504

Armsworth P (2000) Modelling the swimming response of late stage larval reef fish to different stimuli. Marine Ecology Progress Series 195:231-247

Attwood C, Bennett B (1995) Modelling the effect of marine reserves on the recreational shore-fishery of the south-western Cape, South Africa. South African Journal of Marine Science 16:227-240

Atz JW (1965) Hermaphroditic fish. Science 150:789-792

Avise JC (2004) Molecular Markers, Natural History, and Evolution, Vol. Sinnauer Kluwer Academic Publishers

Avise JC (2009) Phylogeography: retrospect and prospect. Journal of biogeography 36:3-15

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual review of ecology and systematics:489-522

Baguette M, Van Dyck H (2007) Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. Landscape ecology 22:1117-1129

Banks SC, Piggott MP, Williamson JE, Bové U, Holbrook NJ, Beheregaray LB (2007) Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecology 88:3055-3064

Barson N, Cable J, Van Oosterhout C (2009) Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source–sink metapopulation structure, founder events and population bottlenecks. Journal of evolutionary biology 22:485-497

Barton NH, Charlesworth B (1984) Genetic revolutions, founder effects, and speciation. Annual Review of Ecology and Systematics 15:133-164

Beckley LE (1983) The ichthyofauna associated with *Zostera capensis* Setchell in the Swartkops estuary, South Africa. South African Journal of Zoology 18:15-24

Beckley LE (2000) Species composition and recruitment of tidal pool fishes in KwaZulu-Natal, South Africa. African Zoology 35:29-34

Bennett AF (1999) Linkages in the landscape: the role of corridors and connectivity in wildlife conservation, Vol. Iucn

Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in ecology & evolution 16:314-321

Berthold P, Kaatz M, Querner U (2004) Long-term satellite tracking of white stork (*Ciconia ciconia*) migration: constancy versus variability. Journal of Ornithology 145:356-359

Bohonak AJ (1999) Dispersal, gene flow, and population structure. Quarterly review of biology:21-45

Botsford LW, Brumbaugh DR, Grimes C, Kellner JB, Largier J, O'Farrell MR, Ralston S, Soulanille E, Wespestad V (2009) Connectivity, sustainability, and yield: bridging the gap between conventional fisheries management and marine protected areas. Reviews in Fish Biology and Fisheries 19:69-95

Bourjea J, Lapegue S, Gagnevin L, Broderick D, Mortimer J, Ciccione S, Roos D, Taquet C, Grizel H (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. Molecular Ecology 16:175-186

Bowne DR, Peles JD, Barrett GW (1999) Effects of landscape spatial structure on movement patterns of the hispid cotton rat (*Sigmodon hispidus*). Landscape Ecology 14:53-65

Budker P (1971) The life of sharks, Vol. George Weidenfeld & Nicholson Ltd, Great Britain

Buxton CD, Garratt PA (1990) Alternative reproductive styles in seabreams (Pisces: Sparidae). Environmental Biology of Fishes 28:113-124

Carr MH, Neigel JE, Estes JA, Andelman S, Warner RR, Largier JL (2003) Comparing marine and terrestrial ecosystems: implications for the design of coastal marine reserves. Ecological Applications 13:90-107

Castoe TA, Poole AW, Gu W, Jason de Koning A, Daza JM, Smith EN, Pollock DD (2010) Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. Molecular Ecology Resources 10:341-347

Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. Molecular biology and evolution 15:538-543

Chistiakov DA, Hellemans B, Volckaert FA (2006) Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. Aquaculture 255:1-29

Chopelet J, Helyar S, Mann B, Mariani S (2009_a) Novel polymorphic microsatellite loci for the protogynous hermaphrodite slinger sea bream (*Chrysoblephus puniceus*, Sparidae). Molecular ecology resources 9:1223-1226

Chopelet J, Waples RS, Mariani S (2009_b) Sex change and the genetic structure of marine fish populations. Fish and Fisheries 10:329-343

Coetzee P (1986) Diet composition and breeding cycle of blacktail, *Diplodus sargus capensis*(Pisces: Sparidae), caught off St. Croix Island, Algoa Bay, South Africa. S AFR J ZOOL/S-AFR TYDSKR DIERKD 21:237-243

Compagno LJV (1988) Sharks of the order Carcharhiniforms, Vol. Princeston University Press, New Jersey, USA

Coward K, Bromage N, Hibbitt O, Parrington J (2002) Gamete physiology, fertilization and egg activation in teleost fish. Reviews in Fish Biology and Fisheries 12:33-58

Cowen R, Paris C, Srinivasan A (2006) Scaling of connectivity in marine populations. Science 311:522-527

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Annual Review of Marine Science 1:443-466

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499-510

de Queiroz A (2005) The resurrection of oceanic dispersal in historical biogeography. Trends in ecology & evolution 20:68-73

de Ruijter W, Biastoch A, Drijfhout S, Lutjeharms J, Matano R, Pichevin T, Leeuwen Pv, Weijer W (1999_a) Indian-Atlantic interocean exchange: Dynamics, estimation and impact. Journal of Geophysical Research: Oceans (1978–2012) 104:20885-20910

de Ruijter WP, Aken HMv, Beier EJ, Lutjeharms JR, Matano RP, Schouten MW (2004) Eddies and dipoles around south Madagascar: formation, pathways and large-scale impact. Deep Sea Research Part I: Oceanographic Research Papers 51:383-400

de Ruijter WP, Ridderinkhof H, Lutjeharms JR, Schouten MW, Veth C (2002) Observations of the flow in the Mozambique Channel. Geophysical Research Letters 29:140 - 141, 140 - 143

de Ruijter WP, Ridderinkhof H, Schouten MW (2005) Variability of the Southwest Indian Ocean. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 363:63-76

de Ruijter WP, van Leeuwen PJ, Lutjeharms JR (1999_b) Generation and evolution of Natal Pulses: solitary meanders in the Agulhas Current. Journal of physical oceanography 29:3043-3055

Divanach P, Kentouri M, Paris J (1982) Etapes du developpement embryonnaire et larvaire du sar, *Diplodus sargus* L., en elevage. Aquaculture 27:339-353

Doherty PJ, Planes S, Mather P (1995) Gene flow and larval duration in seven species of fish from the Great Barrier Reef. Ecology:2373-2391

Domingues VS, Santos RS, Brito A, Alexandrou M, Almada VC (2007) Mitochondrial and nuclear markers reveal isolation by distance and effects of Pleistocene glaciations in the northeastern Atlantic and Mediterranean populations of the white seabream (*Diplodus sargus*, L.). Journal of Experimental Marine Biology and Ecology 346:102-113

Donovan TM, Thompson FR, Faaborg J, Probst JR (1995) Reproductive success of migratory birds in habitat sources and sinks. Conservation Biology 9:1380-1395

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214

Duncan M, James N, Fennessy ST, Mutombene RJ, Mwale M (2015) Genetic structure and consequences of stock exploitation of Chrysoblephus puniceus, a commercially important sparid in the South West Indian Ocean. Fisheries Research 164:64-72

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources 4:359-361

Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. American journal of human genetics 49:746

Ekblom R, Galindo J (2010) Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107:1-15

Ellegren H (2000) Microsatellite mutations in the germline: Implications for evolutionary inference. Trends in genetics 16:551-558

Ellegren H (2004) Microsatellites: Simple sequences with complex evolution. Nature reviews genetics 5:435-445

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. Trends in ecology & evolution 29:51-63

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology 14:2611-2620

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources 10:564-567

Faircloth BC (2008) Msatcommander: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources 8:92-94

Fisher R, Bellwood DR, Job S (2000) Development of swimming abilities in reef fish larvae. Marine Ecology-Progress Series 202:163-173

Fisher R, Leis JM, Clark DL, Wilson SK (2005) Critical swimming speeds of late-stage coral reef fish larvae: Variation within species, among species and between locations. Marine Biology 147:1201-1212

Forbes A, Demetriades N, Benzie J, Ballment E (1999) Allozyme frequencies indicate little geographic variation among stocks of giant tiger prawn *Penaeus monodon* in the South-West Indian Ocean. South African Journal of Marine Science 21:271-277

Fountain T, Nieminen M, Sirén J, Wong SC, Hanski I (2016) Predictable allele frequency changes due to habitat fragmentation in the Glanville fritillary butterfly. Proceedings of the National Academy of Sciences 113:2678-2683

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925

Fumagalli M (2013) Assessing the effect of sequencing depth and sample size in population genetics inferences. PLoS One 8:e79667

Gaina C, Torsvik TH, van Hinsbergen DJ, Medvedev S, Werner SC, Labails C (2013) The African Plate: A history of oceanic crust accretion and subduction since the Jurassic. Tectonophysics 604:4-25

Galtier N, Nabholz B, Glémin S, Hurst G (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular ecology 18:4541-4550

Ghiselin MT (1969) The evolution of hermaphroditism among animals. Quarterly Review of Biology:189-208

Gibbons AD, Whittaker JM, Müller RD (2013) The breakup of East Gondwana: assimilating constraints from Cretaceous ocean basins around India into a best-fit tectonic model. Journal of geophysical research: solid earth 118:808-822

Gilchrist JDF, Thompson WW (1908) Description of fishes from the coast of Natal, Vol. South African Museum

Gilg MR, Hilbish TJ (2003) The geography of marine larval dispersal: coupling genetics with fine-scale physical oceanography. Ecology 84:2989-2998

Gopal K, Tolley K, Groeneveld J, Matthee C (2006) Mitochondrial DNA variation in spiny lobster *Palinurus delagoae* suggests genetically structured populations in the Southwestern Indian Ocean. Marine Ecology Progress Series 319:191-198

Griffiths CL, Robinson TB, Lange L, Mead A (2010) Marine biodiversity in South Africa: an evaluation of current states of knowledge. PLoS One 5:e12008

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series 41:95-98

Halo I, Penven P, Backeberg B, Ansorge I, Shillington F, Roman R (2014) Mesoscale eddy variability in the southern extension of the East Madagascar Current: Seasonal cycle, energy conversion terms, and eddy mean properties. Journal of Geophysical Research: Oceans 119:7324-7356

Hancke L, Roberts M, Ternon J-F (2014) Surface drifter trajectories highlight flow pathways in the Mozambique Channel. Deep Sea Research Part II: Topical Studies in Oceanography 100:27-37

Hanski I, Pakkala T, Kuussaari M, Lei G (1995) Metapopulation persistence of an endangered butterfly in a fragmented landscape. Oikos:21-28

Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human biology:591-600

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Current Anthropology:483-496

Hasegawa M, Kishino H, Yano T-a (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution 22:160-174

Heath D (1977) Simultaneous hermaphroditism; cost and benefit. Journal of Theoretical Biology 64:363-373

Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B: Biological Sciences 270:313-321

Hedenström A, Petterson J (1987) Migration routes and wintering areas of willow warblers *Phylloscopus trochilus*(L.) ringed in Fennoscandia. Ornis Fennica 64:137-143
Heemstra PC, Heemstra E (2004) Coastal fishes of southern Africa, Vol. NISC (PTY) LTD
Hellberg ME (2009) Gene flow and isolation among populations of marine animals.
Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002) Genetic assessment of connectivity among marine populations. Bulletin of marine science 70:273-290
Henriques R, Potts WM, Santos CV, Sauer WH, Shaw PW (2014) Population connectivity and phylogeography of a coastal fish, *Atractoscion aequidens* (Sciaenidae), across the Benguela Current Region: Evidence of an ancient vicariant event. PloS one 9:e87907
Henriques RN (2012) Influence of the Benguela Current in genetic sub-structuring of commercially exploited fish species. Royal Holloway, University of London
Hewitt GM (2004) The structure of biodiversity–insights from molecular phylogeography. Frontiers in Zoology 1:1-16
Hinch SG, Bratty J (2000) Effects of swim speed and activity pattern on success of adult sockeye salmon migration through an area of difficult passage. Transactions of the American Fisheries Society 129:598-606
Hoar WS (1957) The gonads and reproduction. The physiology of fishes 1:287-321
Hodgson S, Quinn TP (2002) The timing of adult sockeye salmon migration into fresh water: adaptations by populations to prevailing thermal regimes. Canadian Journal of Zoology 80:542-555
Hoffman SG, Schildhauer MP, Warner RR (1985) The costs of changing sex and the ontogeny of males under contest competition for mates. Evolution:915-927
Hogarth PJ, Hogarth PJ, Hogarth PJ, Hogarth PJ (1976) Viviparity, Vol. Edward Arnold London
Hubbs C, Lagler K (1947) Fishes of the Great Lakes region. Cranbrook Inst Sci Bull 26:1-186
Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular ecology resources 8:3-17
Hurst GD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proceedings of the Royal Society of London B: Biological Sciences 272:1525-1534
Jones FRH (1968) Fish migration, Vol. Edward Arnold Ltd, London, UK
Joubert C (1981) Aspects of the biology of five species of inshore reef fishes on the Natal coast, South Africa, Vol. Oceanographic Research Institute
Kamel SJ, Grosberg RK, Addison JA (2014) Multiscale patterns of genetic structure in a marine snail (Solenosteira macrospira) without pelagic dispersal. Marine biology 161:1603-1614
Kircher M, Sawyer S, Meyer M (2011) Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic acids research:gkr771
Knudsen SW, Clements KD (2013) Revision of the fish family Kyphosidae (Teleostei: Perciformes). Zootaxa 3751:1-101
Knudsen SW, Clements KD (2016) World-wide species distributions in the family Kyphosidae (Teleostei: Perciformes). Molecular phylogenetics and evolution 101:252-266
Kuwamura T (1997) Evolution of female egg care in haremic triggerfish, *Rhinecanthus aculeatus*. Ethology 103:1015-1023
Leis JM, Carson-Ewart BM (1997) In situ swimming speeds of the late pelagic larvae of some Indo-Pacific coral-reef fishes. Marine Ecology Progress Series 159:165-174
Lindenmayer D, Possingham H (1996) Modelling the inter-relationships between habitat patchiness, dispersal capability and metapopulation persistence of the endangered species, Leadbeater's possum, in south-eastern Australia. Landscape Ecology 11:79-105
Liu J-X, Gao T-X, Zhuang Z-M, Jin X-S, Yokogawa K, Zhang Y-P (2006) Late Pleistocene divergence and subsequent population expansion of two closely related fish species, Japanese anchovy (*Engraulis japonicus*) and Australian anchovy (*Engraulis australis*). Molecular phylogenetics and evolution 40:712-723

Liu Y-G, Liu L-X, Wu Z-X, Lin H, Li B-F, Sun X-Q (2007) Isolation and characterization of polymorphic microsatellite loci in black sea bream (*Acanthopagrus schlegeli*) by cross-species amplification with six species of the Sparidae family. Aquatic Living Resources 20:257-262

Lobel PS, Johannes RE (1980) Nesting, eggs and larvae of triggerfishes (Balistidae). Environmental Biology of Fishes 5:251-252

Lombard A (2004) Marine component of the National Spatial Biodiversity Assessment for the development of South Africa's National Biodiversity Strategic and Action Plan. National Botanical Institute 101

Lowe A, Harris S, Ashton P (2009) Ecological genetics: design, analysis, and application, Vol. John Wiley & Sons

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1:18

Lutjeharms J, de Ruijter W (1996) The influence of the Agulhas Current on the adjacent coastal ocean: possible impacts of climate change. Journal of Marine Systems 7:321-336

Lutjeharms J, Machu E (2000) An upwelling cell inshore of the East Madagascar Current. Deep Sea Research Part I: Oceanographic Research Papers 47:2405-2411

Lutjeharms J, Monteiro P, Tyson P, Obura D (2001) The oceans around southern Africa and regional effects of global change: START Regional Syntheses. South African Journal of Science 97:p. 119-130

Lutjeharms J, Roberts H (1988) The Natal pulse: An extreme transient on the Agulhas Current. Journal of Geophysical Research: Oceans (1978–2012) 93:631-645

Lutjeharms J, Valentine H, Van Ballegooyen R (2000a) The hydrography and water masses of the Natal Bight, South Africa. Continental Shelf Research 20:1907-1939

Lutjeharms J, Wedepohl P, Meeuwis J (2000b) On the surface drift of the East Madagascar and Mozambique Currents. South African Journal of Science 96

Machu E, Lutjeharms J, Webb A, Van Aken H (2002) First hydrographic evidence of the Southeast Madagascar upwelling cell. Geophysical research letters 29:5-1-5-4

Mäkinen TS, Niemelä E, Moen K, Lindström R (2000) Behaviour of gill-net and rod-captured Atlantic salmon (*Salmo salar* L.) during upstream migration and following radio tagging. Fisheries Research 45:117-127

Mann B, Buxton C (1992) Diets of *Diplodus sargus capensis* and *D. cervinus hottentotus* (Pisces: Sparidae) on the Tsitsikamma coast, South Africa. Koedoe 35:27-36

Mann B, Buxton C (1997) Age and growth of *Diplodus sargus capensis* and *D. cervinus hottentotus* (Sparidae) on the Tsitsikamma coast, South Africa. Cybium 21:135-147

Mann B, Buxton C (1998) The reproductive biology of *Diplodus sargus capensis* and *D. cervinus hottentotus* (Sparidae) off the south-east Cape coast, South Africa. Cybium 22:31-47

Mann BQ (2013) Southern African marine linefish species profiles.

Mann BQ, Fennessy ST, Govender A, van der Walt BA (2002) Age and growth and a preliminary stock assessment of stonebream *Neoscorpis lithophilus* (Pisces: Scorpididae) along the KwaZulu-Natal coast, South Africa. Marine and freshwater research 53:131-138

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer research 27:209-220

Mardis ER (2008) The impact of next-generation sequencing technology on genetics. Trends in genetics 24:133-141

Mariani S, Sala-Bozano M, Chopelet J, Benvenuto C (2013) Spatial and temporal patterns of size-at-sex-change in two exploited coastal fish. Environmental biology of fishes 96:535-541

Mate BR, Krutzikowsky GK, Winsor MH (2000) Satellite-monitored movements of radio-tagged bowhead whales in the Beaufort and Chukchi seas during the late-summer feeding season and fall migration. Canadian Journal of Zoology 78:1168-1181

Metzgar D, Wills C (2000) Evidence for the adaptive evolution of mutation rates. Cell 101:581-584

Meyer A, Lutjeharms J, De Villiers S (2002) The nutrient characteristics of the Natal Bight, South Africa. Journal of marine systems 35:11-37

Miller NG, Wassenaar LI, Hobson KA, Norris DR (2012) Migratory connectivity of the monarch butterfly (*Danaus plexippus*): patterns of spring re-colonization in eastern North America. PLoS One 7:e31891

Moe MA (1969) Biology of the red grouper Epinephelus morio (Valenciennes) from the eastern Gulf of Mexico, Vol. Florida Department of Natural Resources, Marine Research Laboratory

Mokhtar-Jamaï K, Pascual M, Ledoux JB, Coma R, FÉRAL JP, Garrabou J, Aurelle D (2011) From global to local genetic structuring in the red gorgonian Paramuricea clavata: the interplay between oceanographic conditions and limited larval dispersal. Molecular Ecology 20:3291-3305

Morris T, Lamont T, Roberts M (2013) Effects of deep-sea eddies on the northern KwaZulu-Natal shelf, South Africa. African Journal of Marine Science 35:343-350

Munday P, Molony B (2002) The energetic cost of protogynous versus protandrous sex change in the bi-directional sex-changing fish *Gobiodon histrio*. Marine Biology 141:1011-1017

Muths D, Gouws G, Mwale M, Tessier E, Bourjea J (2012) Genetic connectivity of the reef fish *Lutjanus kasmira* at the scale of the western Indian Ocean. Canadian Journal of Fisheries and Aquatic Sciences 69:842-853

Muths D, Grewe P, Jean C, Bourjea J (2009) Genetic population structure of the Swordfish (*Xiphias gladius*) in the Southwest Indian Ocean: Sex-biased differentiation, congruency between markers and its incidence in a way of stock assessment. Fisheries Research 97:263-269

Muths D, Tessier E, Gouws G, Craig M, Mwale M, Mwaluma J, Mwandya A, Bourjea J (2011) Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. Marine Ecology-progress Series 443

Nei M (1968) The frequency distribution of lethal chromosomes in finite populations. Proceedings of the National Academy of Sciences 60:517-524

Nei M (1987) Molecular evolutionary genetics, Vol. Columbia university press

Neigel JE (2002) Is F ST obsolete? Conservation Genetics 3:167-173

Nelson J (2006) Fishes of the World. 4th eds. New York: John

Nikolsky GV (1963) The ecology of fishes, Vol. London

Nylander J (2004) MrModeltest ver. 2.3. Program distributed by the author Evolutionary Centre, Uppsala Univ

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC (2006) Origin, evolution and genome distribution of microsatellites. Genetics and Molecular Biology 29:294-307

Oosthuizen C (2006) Intraspecific genetic variation of the endemic estuarine-dependent sparid, Rhabdosargus holubi (Steindachner 1881). MSc thesis, University of Pretoria, South Africa

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular ecology 13:55-65

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. Ecological applications:S146-S158

Pattrick P, Strydom NA (2008) Composition, abundance, distribution and seasonality of larval fishes in the shallow nearshore of the proposed Greater Addo Marine Reserve, Algoa Bay, South Africa. Estuarine, Coastal and Shelf Science 79:251-262

Pattrick P, Strydom NA (2009) Swimming abilities of wild-caught, late-stage larvae of *Diplodus capensis* and *Sarpa salpa* (Pisces: Sparidae) from temperate South Africa. Estuarine, Coastal and Shelf Science 85:547-554

Peakall R, Smouse P (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular ecology notes 6:288-295

Peakall R, Smouse PE, Huff D (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. Molecular Ecology 4:135-148

Penven P, Lutjeharms J, Florenchie P (2006) Madagascar: a pacemaker for the Agulhas Current system? Geophysical Research Letters 33

Perez L, Infante C, Ponce M, Crespo A, Zuasti E, Funes V, Catanese G, Manchado M (2008) Characterization of eight microsatellite markers in the white sea bream, Diplodus sargus (Teleostei, Sparidae). Molecular ecology resources 8:1291-1293

Piñera J, Bernardo D, Blanco G, Vázquez E, Sánchez J (2006) Isolation and characterization of polymorphic microsatellite markers in *Pagellus bogaraveo*, and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*. Molecular Ecology Notes 6:33-35

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of heredity 95:536-539

Plough LV, Marko PB (2013) Characterization of microsatellite loci and repeat density in the Gooseneck Barnacle, *Pollicipes elegans*, using Next Generation Sequencing. Journal of Heredity 105:136-142

Potts WM, Booth AJ, Richardson TJ, Sauer WH (2014) Ocean warming affects the distribution and abundance of resident fishes by changing their reproductive scope. Reviews in Fish Biology and Fisheries 24:493-504

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Pruett CL, Winker K (2008) The effects of sample size on population genetic diversity estimates in song sparrows Melospiza melodia. Journal of Avian Biology 39:252-256

Puurtinen M (2004) Evolution of hermaphroditic mating systems in animals, Vol. University of Jyväskylä

Quartly G, Srokosz M (2004) Eddies in the southern Mozambique Channel. Deep Sea Research Part II: Topical Studies in Oceanography 51:69-83

Rambaut A, Drummond A (2010) Tracer v 1.5. Program distributed by the authors. Oxford: University of Oxford

Reid K, Hoareau TB, Bloomer P (2012) High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. Molecular ecology resources 12:740-752

Richardson T, Potts W, Santos C, Sauer WH (2011_a) Comparison of the population structure and life-history parameters of *Diplodus capensis* (Sparidae) in exploited and unexploited areas of southern Angola. African Journal of Marine Science 33:191-201

Richardson T, Potts W, Sauer WH (2011_b) The reproductive style of *Diplodus capensis* (Sparidae) in southern Angola: rudimentary hermaphroditism or partial protandry? African Journal of Marine Science 33:321-326

Richardson TJ (2010) The taxonomy, life-history and population dynamics of blacktail, *Diplodus capensis* (Perciformes: Sparidae), in southern Angola. Rhodes University

Ridderinkhof H, De Ruijter W (2003) Moored current observations in the Mozambique Channel. Deep Sea Research Part II: Topical Studies in Oceanography 50:1933-1955

Ridgway T, Sampayo EM (2005) Population genetic status of the Western Indian Ocean: what do we Know? Western Indian Ocean Journal of Marine Science 4:1-10

Roberts MJ, Ternon J-F, Morris T (2014) Interaction of dipole eddies with the western continental slope of the Mozambique Channel. Deep Sea Research Part II: Topical Studies in Oceanography 100:54-67

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Molecular biology and evolution 9:552-569

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574

Roques S, Galarza JA, Macpherson E, Turner GF, Carreras-Carbonell J, Rico C (2007) Isolation of eight microsatellites loci from the saddled bream, *Oblada melanura* and cross-species amplification in two sea bream species of the genus *Diplodus*. Conservation Genetics 8:1255-1257

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174-175

Rozen S, Skaletsky H (1999) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols. Springer, p 365-386

Rubenstein DR, Hobson KA (2004) From birds to butterflies: animal movement patterns and stable isotopes. Trends in Ecology & Evolution 19:256-263

Sadovy de Mitcheson Y, Liu M (2008) Functional hermaphroditism in teleosts. Fish and Fisheries 9:1-43

Sadovy Y, Shapiro DY (1987) Criteria for the diagnosis of hermaphroditism in fishes. Copeia:136-156

Sala-Bozano M, Ketmaier V, Mariani S (2009) Contrasting signals from multiple markers illuminate population connectivity in a marine fish. Molecular ecology 18:4811-4826

Sale PF, Cowen RK, Danilowicz BS, Jones GP, Kritzer JP, Lindeman KC, Planes S, Polunin NV, Russ GR, Sadovy YJ (2005) Critical science gaps impede use of no-take fishery reserves. Trends in ecology & evolution 20:74-80

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152:1079-1089

Schouten MW, de Ruijter WP, van Leeuwen PJ (2002) Upstream control of Agulhas ring shedding. Journal of Geophysical Research: Oceans (1978–2012) 107:23-21-23-11

Scotese CR, Gahagan LM, Larson RL (1988) Plate tectonic reconstructions of the Cretaceous and Cenozoic ocean basins. Tectonophysics 155:27-48

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology letters 9:615-629

Senan S, Kizhakayil D, Sasikumar B, SHEEJA TE (2014) Methods for development of microsatellite markers: an overview. Notulae Scientia Biologicae 6:1

Shamoun-Baranes J, Baharad A, Alpert P, Berthold P, Yom-Tov Y, Dvir Y, Leshem Y (2003) The effect of wind, season and latitude on the migration speed of white storks *Ciconia ciconia*, along the eastern migration route. Journal of Avian Biology 34:97-104

Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. Ecological applications:S159-S169

Shulman MJ, Bermingham E (1995) Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. Evolution:897-910

Siddiqui PJ, Amir SA, Masroor R (2014) The sparid fishes of Pakistan, with new distribution records. Zootaxa 3857:071-100

Sinclair EA, Hobbs RJ (2009) Sample size effects on estimates of population genetic structure: implications for ecological restoration. Restoration Ecology 17:837-844

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457-462

Smith J (1931) New and little known fish from the south and east coasts of Africa. Records of the Albany Museum of Grahamstown 4:145-160

Smouse PE, Long JC (1992) Matrix correlation analysis in anthropology and genetics. American Journal of Physical Anthropology 35:187-213

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic zoology 35:627-632

Söndgerath D, Schröder B (2002) Population dynamics and habitat connectivity affecting the spatial spread of populations–a simulation study. Landscape Ecology 17:57-70

Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han J-H (2001) Haplotype variation and linkage disequilibrium in 313 human genes. Science 293:489-493

Stobutzki I, Bellwood D (1998) Nocturnal orientation to reefs by late pelagic stage coral reef fishes. Coral Reefs 17:103-110

Storey M, Mahoney JJ, Saunders AD, Duncan RA (1995) Timing of hot spot-related volcanism and the breakup of Madagascar and India. Science 267:852

Strydom NA (2008) Utilization of shallow subtidal bays associated with warm temperate rocky shores by the late-stage larvae of some inshore fish species, South Africa. African Zoology 43:256-269

Summerer M, Hanel R, Sturmbauer C (2001) Mitochondrial phylogeny and biogeographic affinities of sea breams of the genus *Diplodus* (Sparidae). Journal of Fish Biology 59:1638-1652

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology & Evolution 15:199-203

Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Molecular Biology and Evolution 13:510-524

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

Teske P, Forget F, Cowley P, von der Heyden S, Beheregaray L (2010) Connectivity between marine reserves and exploited areas in the philopatric reef fish Chrysoblephus laticeps (Teleostei: Sparidae). Marine Biology 157:2029-2042

Teske PR, von der Heyden S, McQuaid CD, Barker NP (2011) A review of marine phylogeography in southern Africa. South African Journal of Science 107:43-53

Torsvik TH, Cocks LRM (2013) Gondwana from top to base in space and time. Gondwana Research 24:999-1030

Tóth G, Gáspári Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome research 10:967-981

van der Elst R (2012) A guide to the common sea fishes of southern Africa, Vol. Struik Nature, Cape Town, South Africa

Visram S, Yang M-C, Pillay RM, Said S, Henriksson O, Grahn M, Chen CA (2010) Genetic connectivity and historical demography of the blue barred parrotfish (*Scarus ghobban*) in the western Indian Ocean. Marine Biology 157:1475-1487

von der Heyden S (2009) Why do we need to integrate population genetics into South African marine protected area planning? African Journal of Marine Science 31:263-269

von der Heyden S, Lipinski MR, Matthee CA (2007) Mitochondrial DNA analyses of the Cape hakes reveal an expanding, panmictic population for Merluccius capensis and population structuring for mature fish in Merluccius paradoxus. Molecular Phylogenetics and Evolution 42:517-527

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular ecology 10:249-256

Wallace JH, van der Elst R (1983) Marine linefish programme priority species list. Report No. 0798826894, National Scientific Programmes Unit: CSIR

Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution:385-400

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences 360:1847-1857

Warner RR (1975) The adaptive significance of sequential hermaphroditism in animals. American Naturalist:61-82

Webster MS, Marra PP, Haig SM, Bensch S, Holmes RT (2002) Links between worlds: unraveling migratory connectivity. Trends in Ecology & Evolution 17:76-83

Weersing KA, Toonen RJ (2007) Population genetics, larval dispersal, and demographic connectivity in marine systems. Marine Ecology Progress Series 393

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. Proceedings of the Royal Society B: Biological Sciences:1685–1694

Whitfield AK (1998) Biology and ecology of fishes in southern African estuaries, Vol. J.L.B. Smith Institute of Ichthyology

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: FST≠ 1/(4Nm+ 1). Heredity 82:117-125

Wilcove DS, McLellan CH, Dobson AP (1986) Habitat fragmentation in the temperate zone. Conservation biology 6:237-256

Winter A, Martin K (1990) Late Quaternary history of the Agulhas current. Paleoceanography 5:479-486

Wootton RJ (1998) Ecology of teleost fishes, Vol. Kluwer Academic Publishers, Dordrecht, The Netherlands

Wourms JP (1981) Viviparity: the maternal-fetal relationship in fishes. American Zoologist 21:473-515

Wright S (1946) Isolation by distance under diverse systems of mating. Genetics 31:39

Wright S (1949) The genetical structure of populations. Annals of eugenics 15:323-354

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution:395-420

Yoder AD, Nowak MD (2006) Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annual Review of Ecology, Evolution, and Systematics:405-431

Zink RM, Barrowclough GF, Atwood JL, Blackwell-Rago RC (2000) Genetics, taxonomy, and conservation of the threatened California gnatcatcher. Conservation Biology 14:1394-1405

APPENDIX

Appendix 1: the forward and reverse primer sequences of the genetic markers (cytochrome *c* oxidase subunit I and microsatellite loci) used to amplify DNA from *D. capensis* and *N. lithophilus* samples from the Southwest Indian Ocean. The primers for cytochrome *c* oxidase subunit I (coxI) were used to amplify mitochondrial DNA from both *D. capensis* and *N. lithophilus* samples.

Genetic marker	Marker ID	Forward Primer (5' – 3')	Reverse Primer (5' – 3')	Product Size (in bp)	Reference
coxI	Fish F2t1 & Fish R2t1	TCGACTAATCATAAAGATATCGGCAC	ACTTCAGGGTGACCGAAGAATCAGAA	± 655	Ward et al. 2005
D. capensis microsatellites	DCMS1	TGATGACGGCTGCAGAGAC	GAACTGTAACTTTCTGCACGC	216	-
	DCMS2	TGTAACAGCATCCAGTTCCTG	GTGAACAGACATCCAGCAGC	230	-
	DCMS3	CCTCATCTAAAGACCGCACG	GAACTCTGTCTGTGCATGGC	189	-
	DCMS4	CTGCTCTGTCTTACTGTGTTTG	CAGCAGCAACATGACAGTTC	117	-
	DCMS5	GGTCATCCATCCATCCAGAAC	GTTCAGACCTGCTCGATTTGG	112	-
	DCMS6	GATTGGACAGGCAGGCGG	CGGAGTCATGATGTGTGCAG	150	-
	DCMS7	CCGGGCCTCTCGATCATATG	CCTCTACAGCCTCCTCGAAG	142	-
	DCMS8	CAAGGCTGCAGTCTGTTCTC	AGTCATAGTCTTGCCATGTTGC	136	-
	DCMS9	CATGCAATGGGTAGCGGTG	TTGTAACCAGGCTACCGTCC	443	-
	DCMS10	TGTCCACTTACCTCCACCAAG	AACCAACCCACCATTCATCC	152	-
	DCMS11	TTTCTACATGAGTCTGTCGGC	CCCATAGTGATTGGGCAGGG	117	-
	DCMS12	GGGAACGTGTGAGGGAATTG	TTTACTGAGCGACACGATGC	180	-
	DCMS13	TTCAAACGCCAGATCACACC	TAACAGGCCACAAACCACTG	123	-
	DCMS16	AGCCTAACCTCACTTCTGCC	TTATTAAACAGAGGCAAGCGTG	164	-
	DCMS17	AGCGCAGGGACCTTATCATG	GTTGTTCCTTCATGTCCCGC	150	-
	DCMS18	GTGCCAACCCAAAGACCATC	AGGACAGCTGAATCCTAACCC	160	-
	DCMS19	AGGCGTTGCAGTCTTTGATG	AGGCCGTCTTAGTGTCCTTC	141	-
	DCMS21	CTGTGATGTGAATGCCGATTG	GCAACAACGACTGCAACCG	132	-
	DCMS22	AGGCACAATAAGGACAAGGAG	GAGGAGCAGCTTGACTCAAC	139	-
	DCMS23	ATCACCGTGCACACAAACTG	GACCGCAGAGCTGTTTGTG	100	-
	DCMS24	CAACTTGGAGGACTGAGCG	AGGGATCACTGCACTGAGAG	178	-
	DCMS25	TCTGTCCACCTTGTCCAGTG	ATTCCATTCAGCGCCAAAGG	176	-
	DCMS29	TCTCCACCCACTTTGACAGG	GGGTCGCTGTGGTTACTGG	144	-
	DCMS30	GCTCCTCGTTCACCTTCTTG	TGTTTGTGACGATCAGCAGC	165	-
N. lithophilus microsatellites	NLMS1	TATCTACGGTGCCTCAGTGC	GCAAGTCGTACATGTCGGAC	192	-
	NLMS2	TCTCAGATGTCAAAGCAGCAC	GCCATGTCACTGTCACCATC	125	-
	NLMS8	GCACATGAGGCGATGAGAAAC	GAGACAGGTGGAGTGTAGGG	202	-
	NLMS9	GATTGCAGACTTGACAGCGG	GAGTGTTGTTCGGGAAGCAG	236	-
	NLMS13	GCAGTGAATGTGGCTCTCC	TGTTACGGCACACAGAGGAG	152	-
	NLMS14	AGTAACGGAGCACTGCAAAG	AGAGACAGCCCATACACCTC	150	-
	NLMS15	GCAGACAGAAGCGTTAGTGC	CGAGCTGTGCGTTTGTAGAG	130	-
	NLMS20	TCAGAGGTGAGTTGTGTCGC	TCACTGCTGTCAACCAACAC	207	-
	NLMS 23	TATTGTTACCCAGCTCCGGG	AATTCAGCCGAGAGTGGAGG	130	-
	NLMS26	CGCGCAGAGGAATAATGAGG	TGCTACCACATACTCAGTTCAG	188	-
	NLMS27	ACAGAGCAGAAGAAAGCAAGC	GAGACAGCCACTCTCCTACAG	108	-

Appendix 2: Known microsatellite primer sequences that were included in this study. The label the primer was given for this study, the repeat motif, the primer sequences, the annealing temperature (T_a), PCR product size, the species these primers were used in by previous authors (*Pachymetopon blochii*, *Pagellus bogaraveo*, *Acanthopagrus schlegeli* and *Pagrus auratus*), the primer sequence references, and the available GenBank accession numbers for each primer pair are included in this table.

Primer label	Repeat motif	F. Primer sequence (5'-3')	R. Primer sequence (5'-3')	T_a (°C)	Product Size (bp)	Species	Reference	GenBank accession
DCMS14	$(AGAT)_9$	AGGTTCCCCACAGAAGGTC	AGTACCTGGGAAACAGCCC	61	224	*P. blochii*	(Reid et al. 2012)	JQ688073
DCMS20	$(AGAT)_{11}$	CATTTTGGCAGTCCGTCCG	CATTTTGGCAGTCCGTCCG	61	243	*P. blochii*	(Reid et al. 2012)	JQ688075
DCMS26	$(AC)_{15}$	ACGGCTGTGAGGTCAGAA	ATGGAGCGTGTGGTCAGT	60	109–221	*P. bogaraveo*	(Piñera et al. 2006)	-
DCMS27	$(AG)_{16}$	ACGGACAGAGAGGGAGTGG	CATCATCATCAGTCAGAGCTG	55	80–110	*A. schlegeli* and *P. auratus*	(Adcodk et al. 2000, Liu et al. 2007)	-
DCMS28	$(AC)_{17}$	CAGATACAGGCAGAGGAGC	CAATTAGAGGAGGGAGAACG	55	160–190	*A. schlegeli* and *P. auratus*	(Adcodk et al. 2000, Liu et al. 2007)	-

Appendix 3: Tracer v1.5 output illustrating the mean and ESS values for statistics based on the cytochrome *c* oxidase subunit I (coxI) of mitochondrial DNA for *D. capensis* populations in the SWIO. These statistics are based on the combination of two independent runs.

Statistic	Mean	ESS
posterior	-3046.409	17980.023
prior	-2230.997	17985.389
likelihood	-815.412	16931.487
treeModel.rootHeight	0.04163	14657.960
skyline.popSize1	0.408	1527.400
skyline.popSize2	0.242	1357.557
skyline.popSize3	0.177	1450.633
skyline.popSize4	0.124	1413.145
skyline.popSize5	0.081	1126.452
skyline.popSize6	0.060	1767.327
skyline.popSize7	0.036	1478.010
skyline.popSize8	0.022	1276.858
skyline.popSize9	0.014	1667.586
skyline.popSize10	0.009	1843.256
skyline.groupSize1	2.358	4852.486
skyline.groupSize2	2.294	6163.638
skyline.groupSize3	2.314	7737.413
skyline.groupSize4	2.309	8209.701
skyline.groupSize5	2.266	9090.490
skyline.groupSize6	2.291	8462.397
skyline.groupSize7	2.278	8573.848
skyline.groupSize8	2.312	7793.327
skyline.groupSize9	2.306	6094.042
skyline.groupSize10	2.273	5513.538
CP1+2.kappa	3.572	1738.038
CP3.kappa	0.774	1595.594
CP1+2.frequencies1	0.239	2892.360
CP1+2.frequencies2	0.276	2691.458
CP1+2.frequencies3	0.200	2537.952
CP1+2.frequencies4	0.285	1917.854
CP3.frequencies1	0.257	2620.231
CP3.frequencies2	0.298	2292.126
CP3.frequencies3	0.121	1835.581
CP3.frequencies4	0.324	2486.910
CP1+2.alpha	0.352	1475.736
CP3.alpha	0.366	1365.040
CP1+2.mu	0.995	11519.743
CP3.mu	1.010	11519.743
ucld.mean	0.758	1042.607
ucld.stdev	0.436	9946.207
meanRate	0.743	1027.369
coefficientOfVariation	0.458	9516.147
covariance	-0.019	16672.375
CP1+2.treeLikelihood	-548.124	17011.371
CP3.treeLikelihood	-267.288	12267.208
skyline	38.244	17999.575

Appendix 3: Tracer v1.5 output illustrating the mean and ESS values for statistics based on the cytochrome *c* oxidase subunit I (coxI) of mitochondrial DNA for *N. lithophilus* populations in the SWIO. These statistics are based on the combination of two independent runs.

Statistic	Mean	ESS
posterior	-3022.684	17922.674
prior	-2240.877	17932.796
likelihood	-781.807	15994.794
treeModel.rootHeight	0.060	8353.509
skyline.popSize1	0.293	3822.855
skyline.popSize2	0.153	2458.404
skyline.popSize3	0.101	3276.576
skyline.popSize4	0.060	1668.060
skyline.popSize5	0.042	1303.077
skyline.popSize6	0.031	1122.323
skyline.popSize7	0.025	1074.124
skyline.popSize8	0.023	1360.418
skyline.popSize9	0.025	1387.868
skyline.popSize10	0.030	1487.885
skyline.groupSize1	1.191	14643.116
skyline.groupSize2	1.192	15694.406
skyline.groupSize3	1.187	17552.224
skyline.groupSize4	1.200	16833.748
skyline.groupSize5	1.202	1800100
skyline.groupSize6	1.203	16992.939
skyline.groupSize7	1.208	17827.945
skyline.groupSize8	1.207	15912.608
skyline.groupSize9	1.207	16761.590
skyline.groupSize10	1.202	16650.052
CP1+2.kappa	2.818	2684.173
CP3.kappa	2.521	1042.625
CP1+2.frequencies1	0.212	2750.456
CP1+2.frequencies2	0.300	2365.884
CP1+2.frequencies3	0.168	2857.935
CP1+2.frequencies4	0.320	2455.630
CP3.frequencies1	0.239	2651.903
CP3.frequencies2	0.288	2754.382
CP3.frequencies3	0.174	2559.453
CP3.frequencies4	0.299	2368.966
CP1+2.alpha	0.512	1516.736
CP3.alpha	0.528	1400.316
CP1+2.mu	1.233	17142.324
CP3.mu	0.531	17142.324
ucld.mean	0.813	1336.737
ucld.stdev	0.671	513.024
meanRate	0.781	1274.951
coefficientOfVariation	0.730	493.841
covariance	-0.016	9481.175
CP1+2.treeLikelihood	-538.25	15527.700
CP3.treeLikelihood	-243.56	15233.777
skyline	34.401	17986.580

Appendix 4: product size (in bp) of each of the eighteen microsatellite loci amplified in *D. capensis* samples (n = 33) from the SWIO (Madagascar, Cape Vidal and Durban).

Sample	Population	DCMS1		DCMS2		DCMS4		DCMS5		DCMS6		DCMS7		DCMS8		DCMS9		DCMS11		DCMS13		DCMS17		DCMS18		DCMS19		DCMS21		DCMS22		DCMS23		DCMS24		DCMS25	
5.2	Madagascar	216	216	224	224	117	117	112	112	148	141	142	142	136	136	443	443	114	117	113	113	158	145	157	160	129	129	135	135	139	141	100	103	-9	-9	176	176
6	Madagascar	216	216	224	224	-9	-9	108	112	148	141	-9	-9	136	138	-9	-9	-9	-9	-9	-9	145	145	158	160	-9	-9	129	135	-9	-9	-9	-9	163	163	-9	-9
11	Madagascar	216	220	224	227	117	117	120	124	148	141	142	142	136	136	443	443	117	120	123	123	-9	-9	157	163	132	138	126	126	139	139	-9	-9	163	163	176	178
12	Madagascar	216	216	227	230	117	117	104	120	148	150	142	142	136	136	443	443	108	117	113	113	137	145	158	160	132	135	135	135	139	139	98	100	163	166	180	182
13	Madagascar	216	216	224	227	114	117	108	124	148	150	142	142	138	138	437	443	108	117	123	123	137	149	157	157	132	138	132	132	139	141	98	100	163	169	174	176
15	Madagascar	208	216	224	224	120	120	107	108	150	150	142	142	138	138	443	443	111	117	123	123	-9	-9	157	166	129	129	132	138	139	141	98	100	-9	-9	182	184
16	Madagascar	216	216	224	224	117	117	117	117	148	148	142	142	138	138	443	446	111	111	115	115	137	150	157	160	129	129	132	138	139	141	-9	-9	169	169	176	178
23	Madagascar	216	216	230	230	118	118	108	112	150	150	142	142	136	138	443	443	108	108	113	117	137	154	157	158	129	132	132	135	137	137	90	90	163	163	188	188
26	Madagascar	216	216	224	227	120	120	108	124	148	150	142	142	136	136	443	446	117	117	113	117	145	154	157	158	138	138	138	138	135	135	-9	-9	166	175	184	186
28	Madagascar	212	216	224	227	117	117	108	120	148	150	142	142	136	138	443	443	117	117	123	123	150	158	158	160	138	138	123	135	137	139	98	100	169	166	174	176
29	Madagascar	216	216	224	230	117	117	108	112	148	150	142	142	136	138	443	443	114	117	115	115	145	145	158	166	135	138	135	147	137	139	100	100	169	169	176	176
30	Madagascar	216	220	224	227	117	117	120	120	148	150	142	142	136	136	443	446	114	117	123	125	142	158	157	160	129	138	135	138	137	139	110	112	169	169	178	180
31	Madagascar	216	216	224	230	117	117	108	112	141	148	142	142	138	138	440	443	111	117	115	115	137	154	158	160	129	129	144	144	137	139	-9	-9	163	163	176	176
39	Cape Vidal	216	216	227	230	-9	-9	112	112	148	148	142	142	-9	-9	443	443	108	108	-9	-9	145	145	157	166	129	129	132	135	137	137	106	106	169	169	186	188
41	Cape Vidal	216	216	224	224	117	117	107	116	148	148	142	142	136	136	443	443	108	108	119	119	149	154	157	163	132	132	-9	-9	137	139	104	106	163	163	186	188
42	Cape Vidal	216	220	224	227	117	117	108	112	141	148	142	145	136	136	440	443	111	114	121	123	150	150	157	158	132	138	132	132	139	139	-9	-9	166	172	180	182
44	Cape Vidal	216	216	224	227	117	117	108	116	140	148	142	142	136	136	443	443	108	111	123	127	149	149	158	161	132	132	132	135	143	145	98	100	169	172	178	180
45	Cape Vidal	216	216	224	230	120	120	112	112	140	140	142	142	136	136	443	446	111	111	123	123	145	154	157	160	129	129	123	135	139	139	-9	-9	163	172	188	190
46	Cape Vidal	216	216	224	224	117	120	112	124	148	148	142	142	136	136	440	443	108	114	117	123	145	150	157	163	129	129	135	135	139	141	-9	-9	163	163	178	180
53	Cape Vidal	216	218	230	230	117	117	108	112	148	150	142	142	136	136	443	446	108	108	115	115	137	137	157	160	126	126	132	132	153	153	96	96	172	172	176	178
56	Cape Vidal	216	224	224	224	117	117	112	120	140	150	142	142	136	136	443	446	108	114	115	119	145	145	157	163	129	132	135	135	141	143	98	100	169	169	174	176
57	Cape Vidal	216	220	224	227	117	120	108	112	148	148	136	136	136	138	443	446	117	117	117	117	153	157	157	160	129	129	132	132	135	137	112	112	163	166	184	186
59	Cape Vidal	212	216	227	230	117	117	107	116	141	148	142	142	140	140	443	443	111	117	115	115	145	150	157	160	129	132	117	135	-9	-9	-9	-9	163	175	182	184
60	Cape Vidal	216	220	227	230	-9	-9	112	120	148	148	142	142	136	136	443	443	111	114	119	123	149	162	157	160	132	132	-9	-9	133	133	-9	-9	163	169	178	184
61	Cape Vidal	212	216	227	230	117	117	108	108	148	150	142	148	136	136	443	443	108	108	115	117	137	154	160	163	132	132	123	123	139	139	102	102	172	175	188	190
213	Durban	216	216	224	230	117	120	108	108	148	148	142	142	136	136	443	443	108	108	123	123	150	150	157	160	132	135	129	129	141	149	98	100	169	172	176	178
214	Durban	216	224	224	227	117	120	108	120	148	148	142	142	138	138	443	443	108	111	117	119	149	149	157	160	129	132	120	129	137	137	100	100	169	172	172	172
215	Durban	216	216	221	230	114	114	116	120	148	150	148	148	136	136	440	443	117	117	111	111	137	149	157	160	129	129	126	126	139	141	-9	-9	163	175	182	184
217	Durban	216	216	227	230	117	117	104	111	148	150	142	142	136	136	440	443	111	114	115	115	154	154	169	169	129	135	141	141	139	139	98	100	166	175	176	178
218	Durban	216	216	227	230	114	120	108	112	148	150	-9	-9	136	136	443	443	108	108	119	119	149	150	157	157	129	129	123	144	135	137	98	100	166	166	174	176
220	Durban	216	220	227	230	117	117	108	108	141	148	142	142	136	136	440	443	108	108	117	117	137	149	157	157	129	129	123	123	137	139	100	100	163	172	182	182
221	Durban	216	216	227	230	117	120	115	115	148	150	142	142	136	136	443	443	114	114	117	123	137	150	157	158	129	129	132	132	139	139	98	100	163	163	178	180

Appendix 5: product size (in bp) of each of the ten microsatellite loci amplified in *N. lithophilus* samples (n = 44) from the SWIO (Madagascar, Cape Vidal and Durban).

Sample	Population	NLMS1		NLMS2		NLMS3		NLMS4		NLMS5		NLMS6		NLMS7		NLMS8		NLMS9		NLMS10	
133	Madagascar	184	184	121	121	-9	-9	230	230	137	138	150	150	-9	-9	198	201	-9	-9	188	188
134	Madagascar	184	184	133	133	202	202	230	230	153	160	150	150	-9	-9	198	201	121	124	181	181
135	Madagascar	176	188	121	121	189	190	236	239	161	163	144	150	-9	-9	201	207	121	121	188	188
136	Madagascar	184	188	125	125	193	193	230	230	153	153	150	153	-9	-9	198	201	118	121	188	188
137	Madagascar	184	184	121	121	193	193	236	236	138	139	144	150	-9	-9	198	201	115	121	188	188
140	Madagascar	188	192	129	129	-9	-9	230	230	152	156	150	153	-9	-9	198	204	124	124	182	182
141	Madagascar	196	204	129	129	-9	-9	230	230	150	151	150	150	-9	-9	198	210	121	124	185	188
146	Madagascar	184	188	129	129	196	199	230	230	154	155	150	150	-9	-9	192	207	124	124	188	188
147	Madagascar	176	180	125	125	193	202	236	236	152	153	144	150	129	129	201	201	-9	-9	182	188
157	Madagascar	180	192	121	121	190	190	230	230	142	142	150	153	-9	-9	195	201	124	130	188	188
158	Madagascar	188	188	125	129	196	202	236	239	147	147	150	150	-9	-9	204	207	-9	-9	188	188
159	Madagascar	176	176	125	129	196	196	230	230	-9	-9	150	150	-9	-9	198	201	115	121	182	188
160	Madagascar	176	184	129	129	202	202	230	230	-9	-9	144	150	120	120	198	207	115	115	182	188
161	Madagascar	192	196	129	129	196	196	230	230	-9	-9	150	153	-9	-9	204	207	115	124	188	194
162	Cape Vidal	188	189	129	129	-9	-9	236	236	-9	-9	-9	-9	-9	-9	207	207	-9	-9	188	188
163	Cape Vidal	184	188	129	129	-9	-9	230	230	-9	-9	150	153	-9	-9	204	207	121	121	188	188
164	Cape Vidal	176	188	129	133	-9	-9	233	236	-9	-9	150	150	-9	-9	198	207	-9	-9	188	188
165	Cape Vidal	188	188	125	129	-9	-9	236	236	-9	-9	150	150	-9	-9	201	207	121	121	188	188
166	Cape Vidal	184	188	125	125	-9	-9	236	236	-9	-9	150	153	-9	-9	198	210	124	130	188	188
167	Cape Vidal	184	192	125	129	193	193	230	230	-9	-9	150	150	-9	-9	204	207	121	124	188	188
168	Cape Vidal	180	188	125	125	196	196	236	236	-9	-9	150	153	-9	-9	204	207	-9	-9	188	188
169	Cape Vidal	188	192	125	133	193	193	230	230	-9	-9	150	153	-9	-9	210	210	121	121	188	188
171	Cape Vidal	180	180	125	129	196	196	236	236	-9	-9	150	150	-9	-9	204	207	118	121	188	188
173	Cape Vidal	184	188	129	133	196	205	230	230	-9	-9	153	153	-9	-9	198	204	121	121	188	188
174	Cape Vidal	192	192	121	125	193	193	233	236	-9	-9	150	150	-9	-9	195	207	121	124	188	188
175	Cape Vidal	180	188	125	133	190	202	230	230	-9	-9	150	150	-9	-9	204	204	121	121	182	188
176	Cape Vidal	192	196	121	129	202	202	236	236	-9	-9	150	153	-9	-9	201	207	121	121	188	188
177	Cape Vidal	188	188	121	133	202	202	236	239	-9	-9	153	153	-9	-9	207	210	115	121	188	188
178	Cape Vidal	184	188	125	125	193	193	230	230	-9	-9	150	153	-9	-9	207	207	115	124	188	191
179	Cape Vidal	180	180	125	125	193	193	230	230	-9	-9	150	153	-9	-9	207	210	121	124	182	188
180	Cape Vidal	180	189	125	129	193	205	236	236	-9	-9	150	150	-9	-9	204	204	115	127	188	188
181	Cape Vidal	188	188	125	129	190	190	233	236	-9	-9	150	153	-9	-9	210	210	118	121	188	188
182	Cape Vidal	188	188	125	129	196	196	230	230	-9	-9	150	153	-9	-9	198	207	124	127	196	196
183	Cape Vidal	188	188	129	129	193	193	233	236	-9	-9	150	153	-9	-9	204	207	115	121	188	188
184	Cape Vidal	184	188	125	125	-9	-9	233	233	-9	-9	150	153	-9	-9	201	204	121	121	179	188
185	Cape Vidal	180	184	125	125	-9	-9	236	236	-9	-9	150	153	-9	-9	201	204	121	124	188	188
186	Cape Vidal	188	189	129	129	193	193	230	230	-9	-9	144	150	-9	-9	201	207	121	121	188	188
188	Cape Vidal	184	192	125	133	-9	-9	236	236	-9	-9	144	150	-9	-9	198	207	121	121	188	188
190	Cape Vidal	184	184	125	129	-9	-9	230	230	-9	-9	144	147	-9	-9	204	207	115	124	188	191
193	Cape Vidal	192	192	125	129	196	196	233	239	-9	-9	150	150	-9	-9	198	207	121	127	188	188
223	Durban	184	188	129	129	196	202	236	236	152	153	150	150	-9	-9	207	207	121	124	188	188
224	Durban	184	188	129	129	199	199	230	230	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9
225	Durban	188	189	125	125	193	193	230	230	-9	-9	150	153	-9	-9	201	201	121	127	188	188
226	Durban	-9	-9	133	125	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9