MOLECULAR PHYLOGENY OF SOUTH AFRICAN GOBIID FISHES OF THE GENUS *CAFFROGOBIUS* SMITT, 1900 AND INTRASPECIFIC GENETIC VARIATION OF *CAFFROGOBIUS GILCHRISTI* BOULENGER, 1900

by

CASWELL MAVIMBELA

RESEARCH DISSERTATION

Submitted in fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

ZOOLOGY

in the

FACULTY OF SCIENCE & AGRICULTURE

(School of Molecular & Life Sciences)

at the

UNIVERSITY OF LIMPOPO

SUPERVISOR: Prof. S.M. Dippenaar

CO-SUPERVISOR: Dr. M. Mwale

2011

DECLARATION

I declare that the dissertation hereby submitted to the University of Limpopo, for the degree of Master of Science in Zoology has not previously been submitted by me for a degree at this or any other university; that it is my work in design and in execution, and that all material contained herein has been duly acknowledged.

C. MAVIMBELA (Mr) Date

Student Number [redacted]

ABSTRACT

The molecular phylogeny of the southern African endemic genus *Caffrogobius* and the variability in intraspecific DNA sequences of the prison goby, *Caffrogobius gilchristi,* from the South African coastline, were investigated. The genus *Caffrogobius* (Gobiidae) consists of seven nominal western Indian Ocean species, six of which occur in the southern African region (from Mozambique to Namibia) with the seventh species, *C. dubius,* from Seychelles and do not form part of this study. These are small, benthic fishes that generally inhabit shallow-water habitats in estuaries while their larvae are marine. There is a high degree of morphological similarity among the species of this genus, hampering species identification using traditional dichotomous morphological keys. In a study of the phylogeny of the genus, the cytochrome *b* (cyt-*b*) and cytochrome oxidase subunit I (COI) loci were employed to resolve taxonomic problems.

Six *Caffrogobius* species (*C. gilchristi*, *C. nudiceps*, *C. saldanha*, *C. agulhensis*, *C. caffer* and *C. natalensis*) were used for molecular phylogenetic analyses. Two partial mitochondrial DNA genes, cyt-*b* (281 bp) and COI (535bp) were amplified for 28 specimens using Polymerase Chain Reaction (PCR). Three phylogenetic tree reconstruction methods were employed namely, maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) using *Bathygobius* sp. 1 and *Bathygobius* sp. 2 as outgroups. According to the estimated topologies from the family Gobiidae, the genus *Caffrogobius* is monophyletic and consists of three clades, the *C. agulhensis* and *C. saldanha* clade with *C. natalensis* as sister group, the sister groups *C. gilchristi* and *C. caffer* and the basal clade *C. nudiceps*.

Phylogeographic patterns of the prison goby, *C. gilchristi* using sequences from cyt-*b* (425 bp from 63 specimens) and COI (524 bp from 109 specimens) may reveal patterns related to intraspecific population genetic structure. Demographic history was investigated with neutrality tests (Tajima's D, Fu's Fs and Ramos-Onsins & Rozas R_2), mismatch distributions and haplotype networks which all indicated population expansion. Spatial groupings of populations were examined using AMOVA. Coalescent analyses were used to estimate gene flow between sampling locations and indicated a high level of gene flow among localities. AMOVA showed no consistent

patterns of differences in the genetic structure within populations and thus AMOVA and parsimony network analyses revealed panmictic populations (i.e. no significant population structure detected), suggesting a recent divergence. Furthermore, parsimony networks support recent population expansion of only one population amongst all the localities analysed. Several possible life-history adaptations could be responsible for maintaining gene flow across phylogeographical distribution range. These may include a long pelagic larval stage and larval behavior, as well as oceanographic features such as the flow of sea currents. The presence of a marine pelagic larval stage in gobies may account for this lack of population structuring. It is also reasonable to assume that estuaries farther away from a particular spawning site would receive fewer recruits from that site, than would be the case for estuaries situated at a closer proximity to the breeding event. This means that conservation should not be focused on individual estuaries, but on a wider geographical scale. However this must still be tested further.

ACKNOWLEDGEMENTS

'Everything that happens was once a dream' – David Baird

'Only dead fish swim with the stream' – David Baird

I owe a debt of gratitude to those who were so generous with their time and expertise. Words can't explain the gratitude, support, guidance and encouragement I received from my supervisors. I would like to thank my supervisors SM Dippenaar, M Mwale and my project manager O Gon for their guidance and help. This dissertation was financially and logistically supported by various organizations and institutes which I would like to acknowledge: South African Institute for Aquatic Biodiversity (SAIAB), National Research Foundation, Marine and Coastal Management, Rhodes University, and University of Limpopo (UL) (Turfloop campus).

SAIAB molecular systematics laboratory staff and students, led by Ms P Reddy, PA Maake, T Gonose, G Gouws, ER Swartz, and the SAIAB staff are gratefully thanked for comments and discussions on my work. The Department of Biodiversity from UL is sincerely thanked for their knowledge, wisdom and inspiration and for letting me pursue research in science.

Lastly, I would like to express my sincere gratitude to my family and friends for their love, encouragement and support.

The expertise in this dissertation belongs to those listed above. Any errors are mine.

I would love to give thanks and praise to almighty God, for making me who I am. Life is a school and we are all students.

'The future you get was paid for somewhere in your past' – David Baird

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE
GENERAL INTRODUCTION

1.1 Background

The emergence of molecular genetic methods during the last decades has enabled biologists from many fields to incorporate genetics into the study of the various processes that occur in biological systems (Neto 1998). Since biomolecules play a major role in evolutionary biology and intraspecific genetic variation, studies of biodiversity using molecular systematics have become very popular (Hillis & Moritz 1990; Hillis *et al.* 1996; Neto 1998). Hillis *et al.* (1996) identified three major areas of application for molecular systematics: (i) gene evolution, (ii) intraspecific, or population studies and (iii) interspecific studies. Many different studies can be conducted using these disciplines, including those on the origin of new genes or loci, tracing of allellic genealogies, gene flow, conservation genetics, estimation of species phylogenies, investigation of population genetics, as well as applications in the fields of taxonomy and systematics (Hillis *et al.* 1996; Neto 1998; Avise 2004).

Understanding the molecular basis of genetics has provided an avenue for examining evolutionary and ecological forces in nature (Dougherty *et al*. 1995; Avise *et al*. 1998; Grosberg & Cunningham 2001; Duran 2003). Evolutionary forces — such as selection, genetic drift, non-random mating and migration (gene flow) — play various roles in the genetic structuring of populations, while genetic variation among populations is created by mutation and/or recombination (Verspoor *et al.* 2007). The populations of most, if not all, species show some level of genetic structuring, shaped — at least to some extent — by environmental barriers, historical processes and life histories (Meyer 1994; Moore 1995; Kvist 2000; Duran 2003). Most importantly, when measuring a pattern of genetic variation, the main concern is to interpret the causes that underlie that pattern, and whether the pattern is due to recurrent evolutionary forces or historical factors (Pruett *et al*. 2005).

Since both evolutionary forces and historical factors can play a role in the evolutionary biology, the latter is defined as the study of change in organisms across generations (Ridley 1993; Rodriguez 2007). Fishes (gobies in particular) provide a unique perspective into evolutionary biology because they have undergone extensive adaptive radiation (Thacker 2003). Ecologically, when tracing the evolution of marine and/or estuarine gobies, a pattern of freshwater origins, followed by returns to salt-tolerance, has been revealed. It has, therefore, been generally assumed that gobies originated from a freshwater ancestor and returned a number of times to marine and/or estuarine environments (Thacker & Hardman 2005). According to a family-level phylogenetic analysis, Gobioidei is the sister taxon of the marine family Apogonidae (Thacker & Hardman 2005). Furthermore, phylogenetic analyses estimated that the basal genera *Rhyacichthys* Boulenger, 1901, *Odontobutis* Bleeker, 1874 and *Percottus* Dybowski, 1877 are exclusively freshwater taxa (Thacker & Hardman 2005). Genetic studies of various organisms provide knowledge that can be utilised for the conservation and management of such organisms. Very little research on the molecular phylogeny and population genetics of gobies has, however, been carried out. Such research could help to resolve taxonomic relationships and assess genetic structures within populations of Gobiidae.

Despite a growing interest in population genetics and phylogeography in South Africa, relatively little research on this topic has been carried out on marine and estuarine organisms, particularly in the eastern and southern Cape. Biologists have generally assumed that circulation in the coastal environment allows adequate gene flow, due to the high mobility of pelagic larvae throughout the region (Grant 1993). This assumption may have serious implications for conservation and management of widespread taxa (Grant 1993). The study of intraspecific genetic variation of estuarine-dependent species may help to test this assumption. Consider the following hypothetical situation: when one of two genetically-distinct populations that have been erroneously considered as one population, is threatened, then conservation authorities may intervene by introducing new stock into the threatened system. Genetic characteristics of a particular population represent the adaptations that have evolved in response to a changing environment and

such variability may be crucial to the survival of the particular gene pool. The translocation of genetically-different individuals into such a population may lead to reproductive incompatibilities, out-breeding depression and the loss of long-term evolutionary fitness (Grant 1993). Furthermore, several recent genetic studies on southern African estuarine fish species with a marine larval phase (Klopper 2005; Norton 2005) have revealed a high degree of gene diversity among populations, which contradicts pervious assumptions of extensive gene flow throughout the geographical range of species in marine and estuarine environments. The various factors that affect the structure of marine fish populations include larval behaviour, ocean circulation patterns, distribution of suitable habitats, as well as the geological and climatic history of these habitats (Grosberg & Cunningham 2001).

1.2 Rationale

This genus *Caffrogobius* Smitt, 1900 consists of seven nominal species, of which six occur in the southern African region (Mozambique to Namibia): *C. agulhensis* Banard, 1927, *C. caffer* Günther, 1874, *C. gilchristi* Boulenger, 1900, *C. natalensis* Günther, 1874, *C. nudiceps* Valenciennes, 1837 and *C. saldanha* Banard, 1927 (Goren 1996). The seventh species, *C. dubius* Smith, 1959, is from the Seychelles (Table 1.1). There is, however, some uncertainty about its validity: the only available specimen of this species is the holotype, which is in a poor condition. Since its original description, there have been no further reports of its occurrence from the Seychelles. Consequently, the genus *Caffrogobius*, together with other gobioid, was categorized as endemic to southern Africa by Hoese & Winterbottom (1979a). Three of the species in the genus (*C. saldanha*, *C. agulhensis* and *C. caffer*) are marine — occurring in bays and tidal pools — while the remaining three are estuarine-dependent species i.e (*C. gilchristi*, *C. natalensis* and *C. nudiceps*), occurring in the lower reaches of rivers and estuaries (Goren 1996) (Table 1.1).

The similarity of meristic counts, coloration and the supposedly-diagnostic cephalic sensory system (sensory papillae) of most *Caffrogobius* species has previously led to

much confusion among ichthyologists (Goren 1996). Species in this genus are characterised by a combination of characters, such as the transverse cheek papillae pattern, the number of dorsal and anal segmented rays, and several characters involving scale counts. There is, however, some overlap in the distinguishing features relating to these characters, making it difficult to identify different species and this has also compromised understanding of relationships at the species level. According to Maugé (1986), many fish collections contain large numbers of misidentifications, and long lists of synonyms. As an example of how to overcome such problems, it is worth considering the work of Goren (1996), who undertook a morphological review of *Caffrogobius* species. He recognized *Caffrogobius gilchristi* as a valid species and a senior synonym of the commonly-used *Caffrogobius multifasciatus* Smith, 1959. A molecular review of the relationships among species of the genus would complement the morphological review conducted by Goren (1996) and may provide additional support for the recognition of the different species.

Gobies are a dominant group of coastal fishes in South Africa and aquatic biologists have often misidentified species from this group as they are small, similar-looking fishes and difficult to work with (Hoese & Winterbottom 1979a,b; Hoese 1984; Hoese 1986; Mauge 1986; Hoese & Gill 1993). Hence, neither the phylogenetic placement nor the population structure of marine and estuarine-dependent species of *Caffrogobius* that occur along the South African coast (Figure 1.1) is well known, and there are no known sister taxon of this genus. An assessment of the phylogenetic relationships among members of the family Gobiidae may provide a greater understanding of their systematics, taxonomy, spatial and temporal distribution, and the relationships within and between different populations of the same species. The geographical distribution of organisms can be explained in terms of ecological and historical biogeography (Avise 2004).

The effect of the geological landscape and the Agulhas Current on the distribution of marine and estuarine species has been studied as part of population genetics in South Africa (Klopper 2005; Norton 2005; Neethling *et al.* 2008; Maake 2009). Neethling *et al.* (2008) found a homogeneous population structure for the marine *C. caffer* throughout its

distribution range and suggested that dispersing larval stages play a major role in the high rate of gene flow, thus implying that such larval dispersal mechanisms could have a significant effect on the intraspecific genetic variation in estuarine and marine organisms. Similarly, the estuarine-dependent *C. gilchristi* has a marine larval phase: the larvae migrate to sea for weeks and presumably return to their original estuaries (Whitfield 1998). If these larval stages indeed return to their original spawning habitats it can be expected that there will be population structure across the species distributional range. It may however also be hypothesised that estuarine species with a marine larval phase have a single population, due to dispersal by ocean currents. The validity, or otherwise, of these hypotheses has important consequences for the conservation and management of these widespread species.

1.3 The coastal environment of South Africa

The geomorphology of the South African coast and oceanography of its marine environment affects the biogeography of aquatic fauna in the region (Whitfield 1998). The South African coastal zone is divided into three biogeographical provinces: the cool temperate west coast, the warm-temperate south coast and the subtropical east coast (Figure 1.2). The exact locations of the boundaries between these provinces are still disputed (Lutjeharms 1988; Teske *et al.* 2007a). There is also evidence of a possible fourth tropical province in the north-east of southern Africa (Teske *et al.* 2007b). South Africa's 3000 km coastline stretches from the Orange River Mouth on the west coast (Atlantic Ocean) to the Mozambique border on the east coast (Indian Ocean) (Figure 1.1). The South African coastline is influenced by oceanographic features such as currents, tides, waves and upwelling and is one of the most variable and complex coastlines in Africa (Lutjeharms *et al.* 1993). The two currents that have a major influence on marine ichthyofauna (Figure 1.2) are the cold Benguela Current flowing northwards along the west coast and the warm Agulhas Current flowing southwards along the east coast (Lutjeharms *et al.* 1993). The hydrographic peculiarities of the Agulhas Current affect a number of factors — such as water temperature, nutrient availability, and vertical stability of the water column, as well as advection rate — that influence the distribution

of organisms. Biogeographical distributions of small organisms such as invertebrates (Teske *et al.* 2006, 2007c) all clearly reflect the strong gradients in the water characteristics of the Agulhas Current (Lutjeharms 2005). The resulting conditions may also influence the movement of fish larvae into, or out of, affected areas within this region (Wooldridge 1991). Thus, the Agulhas system has a marked influence on the distribution of a number of pelagic species in the south-west Indian, and south Atlantic, Oceans (Lutjeharms 2005). Because the borders between the southern African biogeographical regions are not clearly defined (Lutjeharms 1988; Teske *et al.* 2007a), comparing different organisms and criteria by applying modern and available historical data is critical in producing a comprehensive picture and understanding of the biogeography of the region (Lutjeharms & Gordon 1987; Lutjeharms & Cooper 1996; Lutjeharms & Ansorge 2001; Norton 2005).

Variation in climate, topography and catchment geology and the connections of estuaries with the marine environment gave rise to a wide variety of estuarine types (Van Niekerk 2007). Thus, the estuaries along the coast of South Africa and their associated fish assemblages are not uniformly distributed and have been categorised according to biological, physical and geographical criteria based largely on biogeographical assessments conducted by Whitfield (1998) and Harrison (2002, 2004). Based on their physical characteristics, Whitfield (1998) classified southern African estuarine systems as follows: estuarine bays, permanently open estuaries, river mouths, estuarine lakes and temporarily closed estuaries. These estuarine systems are driven by the fundamental variables of water quality and movement, which in turn influences the type of substratum that prevails in different parts of each system. Fish respond to these variables on temporal and spatial scales that vary from one estuary to another. The degree of the effect of the physical factors on the behaviour and life histories of fishes, may to some extent also affect the decline or growth of populations in these systems. These assessments have enabled the study of fish species, within designated groups or areas, to facilitate a greater understanding of the the biology of species, and their ecological behaviour (Whitfield & Bate 2007).

Factors that affect the structure of marine fish populations include larval behaviour, ocean circulation patterns, distribution of suitable habitats, and the geological and climatic history of these habitats (Grosberg & Cunningham 2001).

Many southern African studies on coastal invertebrates, fishes and aquatic weeds have made use of mitochondrial molecular markers to elucidate their phylogenies and phylogeography. Results have suggested that ecological (biogeographic regions) and ocean circulation patterns have a major impact on the distribution and genetic structure of species (Klopper 2005; Norton 2005; Neethling *et al.* 2008). Some studies suggest that barriers to gene flow in southern African marine organisms exist at Algoa Bay and Cape Agulhas in the eastern, and southern, Cape regions, respectively (Evans *et al.* 2004; Teske *et al*. 2004, 2006; Edkins *et al.* 2007; Teske *et al.* 2007a, c). The nature of the Agulhas Current begins to change dramatically as it flows southwards past Algoa Bay, following the edge of the Agulhas Bank south of South Africa (Lutjeharms 1988, 2005). After this change in direction, the Agulhas Current returns back to the south Indian Ocean as the Agulhas Return Current (Lutjeharms & van Ballegooyen 1984; 1988). The oceanography of the region between Port Elizabeth and Mossel Bay is therefore extremely complex, due to interactions between wind-driven offshore currents and the Agulhas Current.

1.4 Systematics

1.4.1 Molecular versus morphological phylogenetics

A phylogeny is a lineage of gene transmission that flows from generation to generation and continues through conspecific populations and species to higher taxonomic levels (Kocher & Stepien 1997). A phylogenetic approach can be used to help find answers to a variety of questions in evolutionary biology, such as resolving taxonomic relationships of species, studying the evolution of gene families, evaluating evolutionary rates in different lineages, and dating past historical events that affected species (Avise *et al.* 1987; Moritz *et al.* 1987; Kvist 2000). As is the case for morphological studies — in which the focus is on studying particular characters — molecular phylogeny also utilises various characters,

but in this case these are based on DNA polymorphisms. According to Hillis (1987), the greatest advantage of molecular data in systematics appears to be the large number of observable characters that can be used for analyses.

Whether a molecular or morphological approach is used in a phylogenetic study, the goal is to estimate evolutionary relationships. Traditional taxonomy and systematics have used morphological characters as the raw data for identification and classification, and to infer evolutionary relationships (Avise 2004; Page 2006). Studies of higher level (suborder and family level) relationships of gobiid fishes have made use of both approaches to infer phylogenies (Akihito *et al.* 2000; Thacker 2003; Thacker & Hardman 2005).

The development of molecular techniques has helped to invigorate studies of fish phylogenetics and biogeography (Kocher & Stepien 1997). Animal mitochondrial DNA (mtDNA) provides an ideal genetic system for studies of population and evolutionary biology. This is a small (15-20 kb) circular molecule, composed of 37 genes coding for 22 tRNAs, two rRNAs and 13 mRNAs (Avise 1986, Finne 2001). Mitochondrial DNA can be an effective marker for gathering information for population genetic analysis since it has unique characteristics such as maternal transmission, rapid rate of evolutionary change, transmission without recombination, and haploid inheritance (Brown *et al.* 1982; Avise 1994, 2000, Martins *et al.* 2003). Distinct mtDNA sequences are referred to as haplotypes rather than alleles, because these are inherited maternally as a single non-recombining unit (Avise 1986; Avise *et al.* 1987; Moritz *et al.* 1987; Meyer 1993; Kvist 2000). Different regions of the mitochondrial genome evolve at different rates, which enable researchers to select suitable regions for different kinds of study. Numerous mtDNA studies have revealed significant variation among and within taxa in the mtDNA sequence (Avise *et al.* 1987).

Morphological studies have also been extensively used to classify fish species and to estimate their phylogenetic relationships (Stepien & Rosenblatt 1991; Avise 1994; Goren 1996; Haponski & Stepien 2008; Harold *et al*. 2008). These phylogenies have usually been confirmed when tested with molecular approaches for various fish groups, including

some goby genera — for example, *Gobiodon* Bleeker, 1856 (Harold *et al*. 2008) — as well as studies at higher taxonomic levels; for example, family and suborder (Akihito *et al.* 2000; Thacker 2003; Thacker & Hardman 2005; Haponski & Stepien 2008). Molecular characters have, however, also revealed some cryptic species, and also identified some taxonomic groups that had been incorrectly 'split' on the basis of morphological classifications (Stepien & Rosenblatt 1991; Avise 1994; Kocher & Stepien 1997). The presence of cryptic taxa has confounded many traditional studies, since they are morphologically similar but genetically distinct (Haponski & Stepien 2008). Molecular markers can be used to distinguish cryptic taxa (see Haponski & Stepien 2008). Harold *et al.* (2008) studied phylogenetic relationships of Indo-Pacific coral gobies of the genus *Gobiodon*, based on morphological and molecular data inferred from mtDNA genes (12S and 16S rRNA). The combined data set helped to resolve the phylogeny of the genus *Gobiodon*. Therefore, approaches that combine data sets (such as those based on mtDNA), together with a re-examination of morphological characters, are powerful tools for resolving biogeographic and phylogeographic questions relating to cryptic species groups.

1.4.2 Phylogeography of estuarine-dependent species

Phylogeography is an integrative field of study that combines information from several disciplines — including molecular and population genetics, ethology, demography, phylogenetics and historical geography — to explain the genetic structure of populations (Avise & Walker 1998). It focuses on how contemporary and historical geographic factors have influenced the geographical distribution of gene lineages. The most common molecular tool used in phylogeography has been mtDNA (Avise 2004). This is because the mutation rate of mtDNA is well suited to studies examining events that took place during the last few million years, while for the study of more recent events, markers with higher mutation rates are needed (Avise *et al.* 1988; 1998; Hewitt 2004). Phylogeography is important in the context of conservation and can be used to identify management units and evolutionary significant units (ESUs) (e.g. Avise 2000). These may be employed to identify geographical regions (locations and regions) within which several species display

phylogenetically distinct (reciprocally monophyletic groups) populations and/or unique/different adaptations (Avise 2000; Crandall *et al*. 2000).

In southern Africa, studies based on molecular data have been conducted to elucidate the phylogeography of several estuarine and marine organisms (Klopper 2005; Norton 2005; Neethling *et al.* 2008). Most fish studies (Klopper 2005; Norton 2005; Neethling *et al.* 2008) have suggested single populations in southern Africa that appear to have experienced recent population expansions. However, some studies (Oosthuizen *et al*. 2004; Zardi *et al.* 2007; Teske *et al.* 2007b, c) based on marine invertebrates, have suggested barriers to gene flow between biogeographical regions in this area. Since there are different scientific views on the genetic structure of the southern African marine fauna, more research must be conducted on various marine and estuarine taxa to understand the evolutionary forces determining population stucture in the South African coastline fish community.

1.4.3 Taxonomy of southern African gobies

1.4.3.1 Generic level relationships

A number of genera in the subfamily Gobiinae are found in southern Africa's estuarine habitats, including: *Caffrogobius* Smitt, 1900; *Awaous* Valenciennes, 1837; *Glossogobius* Gill, 1860; *Psammogobius* Smith, 1936 and *Redigobius* Herre, 1927 (see Hoese & Winterbottom 1979a, b; Hoese 1986; Goren 1996; Eschmeyer & Fong 2008), with *Caffrogobius* being the most speciose. There is a lot of similarity between the genus *Caffrogobius* and some European goby genera, such as *Gobius* Linnaeus, 1758 and *Zosterisessor* Whitley, 1935. These genera all have paired anterior interorbital pores and a characteristic head shape with similar papillae rows (Hoese & Winterbottom 1979b). The relationship among several genera of Gobiidae has been estimated by Miller (1978), Hoese (1986), Van Tassell *et al.* (1988), Birdsong *et al.* (1988), and Gill & Bradley (1992). Miller (1978) considered *Caffrogobius* to be closely related to *Gobius* and *Zebrus* de Buen, 1930, within the subfamily Gobiinae, due to the similarity of their cephalic sensory system. Gill & Bradley (1992) further suggested that the cephalic papillae pattern

is also a valid indicator of goby relationships, and that this character implied that *Caffrogobius* is closely related to *Monishia* Smith, 1959, *Coryogalops* Smith, 1958, *Hetereleotris* Bleeker, 1874, *Gorogobius* Miller, 1978 and *Nematogobius* Boulenger, 1910, all of which share a similar transverse cephalic papillae pattern and axial skeleton structure. The genera *Bathygobius* Bleeker, 1878, *Glossogobius* and *Lesuerigobius* Whitley, 1950 of the subfamily Gobiinae, are less closely related to *Caffrogobius* as they have a longitudinal head papillae pattern (Goren 1996). Hoese (1986) regarded *Callogobius*, *Monishia* and *Gobiosoma* as possible relatives of *Caffrogobius*, due to similarities in head and body forms. Furthermore, Gill & Bradley (1992) also suggested that the genus *Caffrogobius* is very similar to *Monishia*, but differs from the latter genus in that it reaches a larger size, and has higher longitudinal and transverse scale counts. These contradictory estimations of relationships among the Gobiinae genera make it difficult to identify the sister group of the genus *Caffrogobius*. Additionally, the lack of a morphological phylogeny for the species of the genus *Caffrogobius*, as well as the number of misidentifications noted by Goren (1996), has complicated the study of this genus and its species.

1.4.3.2 Species-level relationships within the genus *Caffrogobius*

Hoese (1986) provided the first major in-depth study on morphological variation among species of *Caffrogobius* and revised the genus in his checklist of South African gobioid fishes. Six species are found around the southern African coastline (Mozambique to Namibia) (see Section 1.2) while the seventh species, *C. dubius*, was recorded from the Seychelles. Four species of *Caffrogobius* — namely, *C. caffer*, *C. gilchristi*, *C. natalensis* and *C. nudiceps* — were previously classified as belonging to the genus *Gobius* Linnaeus, 1758. Hoese & Winterbottom (1979a) reported that the genus *Gobius* was generally used as a 'dumping ground' for lodging species of uncertain relationships, which exacerbated problems associated with using morphological characters for identifying goby species.

The genus *Caffrogobius* has been recognised as a southern African endemic genus (Hoese & Winterbottom 1979a; Goren 1996). Goren (1996) utilised the cephalic sensory

system (sensory papillae) on both sides of the head — scale counts (longitudinal and transverse series; predorsal), and fin rays and gill raker counts — to define the genus *Caffrogobius*. This author proposed a set of morphological characters that diagnose the seven named species included in his revision, but did not use formal cladistic or phenetic analyses to generate a hypothesis about the relationships among these species. Goren (1996) specifically noted the extremely high variability in longitudinal and transverse scale counts in *C. nudiceps* populations and suggested that this could indicate a potential mixture of two closely related species. To date no study of the evolutionary relationships of the species of *Caffrogobius* has focussed on molecular or morphological characters.

1.5 Molecular markers – an overview

A larger variety of neutral markers — such as mini- and microsatellites, different kinds of restriction fragments, and DNA sequences — are utilized today than was the case during the mid 1960's (Kvist 2000). For a genetic marker to be useful, it must vary intraspecifically or interspecifically and have a known means of inheritance. A well developed theoretical background facilitates the interpretation of both co-dominant nuclear markers and maternally transmitted mtDNA. Other factors worth considering when choosing a genetic marker include: the expense and time required to survey a sufficient number of populations and individuals to resolve the question of interest, the quantity and quality of sample tissue needed to obtain individual genotypes, and the likelihood that a chosen marker might be subjected to strong evolutionary forces other than genetic drift and gene flow (i.e., selection) (Moritz *et al.* 1987; Avise 2000).

Genetic variation is introduced into a population by mutation, recombination or immigration and can also be affected by natural selection or random fluctuations of allele frequencies (Dobzhansky *et al.* 1977; Kvist 2000). Molecular markers that have recently been widely and successfully used for population structure studies include mtDNA regions, cytochrome *b* (cyt-*b*) (Kocher & Stepien 1997; Akihito *et al*. 2000; Klopper 2005; Maake 2009) and cytochrome oxidase subunit I (COI) (Trontelj *et al.* 2005; Wang *et al.* 2007; Gerard *et al.* 2008; Gruenthal & Burton 2008).

The cyt-*b* gene, which codes for a transmembrane protein that plays an important role in the respiratory chain of cellular metabolism, has recently been the most prevalent source of sequence data in fish studies (Kocher & Stepien 1997; Moore & DeFillipss 1997; Akihito *et al.* 2000; Kvist 2000; Doadrio *et al.* 2002; Klopper 2005). The cyt-*b* gene is believed to be appropriate for the investigation of recent Cenozoic divergences (Harrison 2002). This gene is under strong evolutionary constraints, with some parts of the gene being more conserved than others, due to functional restrictions (Meyer 1994). Nevertheless, the rate of evolution in some positions is relatively fast (Irwin *et al.* 1991), although it evolves slowly in terms of non-synonymous substitutions. Therefore, the cyt-*b* gene is thought to be variable enough for population-level questions, and conserved enough for clarifying complex phylogenetic relationships (Avise *et al.* 1987).

The cyt-*b* gene is probably the most-studied mtDNA gene in ichthyological assessments and has been successfully used to determine molecular phylogeny and phylogeography of gobies (Kocher & Stepien 1997; Akihito *et al.* 2000; Harada *et al.* 2002; Stefanni & Thorley 2003; Thacker & Hardman 2005; Harold *et al.* 2008; Haponski & Stepien 2008; Neethling *et al*. 2008; Maake 2009). Previous studies on the phylogeny and phylogeography of gobies, using the cyt-*b* gene and combined mtDNA data sets, include a study of the genetic diversity and historical population structure of *Schizopygopsis pylzovi* Kessler, 1876 in the Qinghai-Tibetan Plateau (Qi *et al*. 2007) and the systematics and phylogeography of the common goby, *Pomatoschistus microps* Krøyer, 1838 (Gysels *et al.* 2004).

The COI gene is present in all animals and sequence comparisons are not complicated because insertions and deletions are rare (Huang *et al.* 2007). COI is also the slowest-evolving gene of the mitochondrial protein-coding genes (Simon *et al.* 1994; Vandewoestijne *et al.* 2004). The gene is composed of interspersed, highly conserved membrane-bound regions and variable extra-membrane loops (Saraste 1990; Lunt *et al.* 1996; Roe & Sperling 2007). Studies of COI molecular evolution have shown that its structure leads to functional constraints on the mutation rate, particularly in the

membrane-bound regions (Saraste 1990; Lunt *et al.* 1996; Caterino & Sperling 1999; Roe & Sperling 2007). Because the COI gene is composed of both highly conserved, and more variable, regions, it is suitable for simultaneous analysis of both closely- and more distantly- related taxonomic groups (Lunt *et al.* 1996; Stahls & Nyblom 2000). COI gene sequence analysis is therefore an effective tool for assessing intraspecific population genetics in fishes (Trontelj *et al.* 2005; Wang *et al.* 2007; Gerard *et al.* 2008; Gruenthal & Burton 2008) and for inferring relationships among closely-related, and morphologically-similar, species (Huang *et al.* 2007).

1.6 Outlines of the present study

The work presented here is part of the South African Institute for Aquatic Biodiversity (SAIAB) goby project that was initiated in 2005. The overall contribution of this dissertation to the project was to study the systematics of the genus *Caffrogobius*.

1.6.1 Aims and objectives

The purpose of this study is to examine phylogenetic relationships within the genus *Caffrogobius* and also to focus on intraspecific genetic variation within one of the species, *C. gilchristi*. More specifically, the study will investigate the systematics of *Caffrogobius* in South Africa using molecular markers (cyt-*b* and COI) in an attempt to clarify the biogeographic distribution and species status.

The objectives of the study are:

- To investigate phylogenetic relationships of the species of the gobiid fish genus *Caffrogobius* in southern Africa using mitochondrial DNA genes (cyt-*b* and COI).
- To examine intraspecific genetic variation of the prison goby (*Caffrogobius gilchristi*) using mitochondrial DNA genes (cyt-*b* and COI). This will test the hypothesis that *C. gilchristi* individuals along the southern African coast constitute a single population.

The following questions will be addressed:

a. What is the nature of the genetic structure in populations of the estuarine-dependent goby, *C. gilchristi* along the southern African coastline?
b. If population sub-structures exist, are these structured geographically?
c. What are the phylogeographic relationships between the sub-structures?
d. What is the extent of gene flow between the sampled locations?
e. What is the extent of genetic variation within the populations and between various locations/regions?
f. Does the level of genetic differentiation between sub-structures justify recognition of distinct units of biodiversity for conservation purposes?

The current study will attempt to contribute towards a greater understanding of phylogenetics and genetic diversity of marine and estuarine species of *Caffrogobius* in southern Africa. Knowledge of genetic variation in this species will provide some insight into the conservation of estuarine gobies, which could be useful for identifying appropriate units for conservation management. This knowledge could therefore contribute towards efforts to conserve marine and estuarine fish diversity in South Africa.

Table 1.1: An ecological classification of the genus *Caffrogobius*, after Goren (1996) and Hoese (1986) (habitat is indicated by ED = estuarine dependent and M = marine species, distribution is indicated by CT = cool temperate; WT = warm temperate; ST = subtropical; TR = tropical. * = *Caffrogobius dubius* reported from Seychelles).

SPECIES	**HABITAT**	***DISTRIBUTION***
C. gilchristi	*ED*	CT, WT, ST
C. natalensis	*ED*	WT, ST
C. caffer	*M*	CT, WT
C. agulhensis	*M*	CT, WT
C. saldanha	*M*	CT, WT
C. nudiceps	*ED*	CT, WT
C. dubius*	*M*	TR

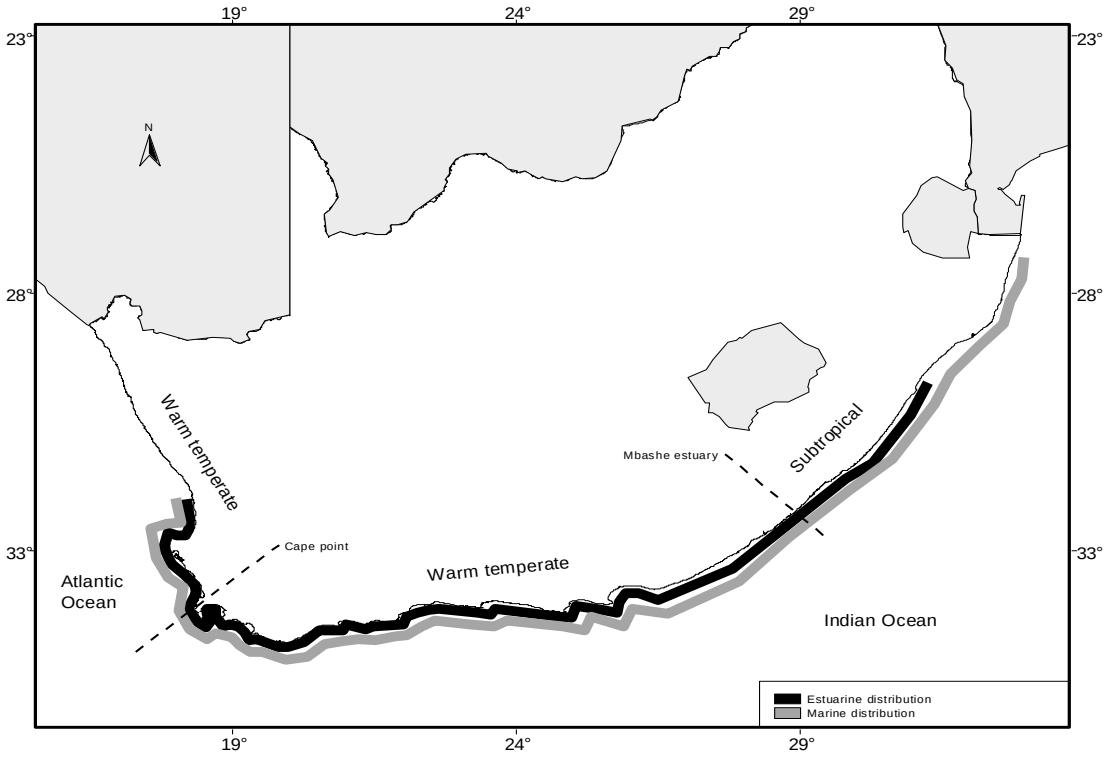

Figure 1.1: The South African coastal zone, indicating the estuarine and marine distribution of naturally-occurring *Caffrogobius* species: the black solid line represents the distribution of *Caffrogobius gilchristi*; the solid grey line represents the distribution of other *Caffrogobius* species.

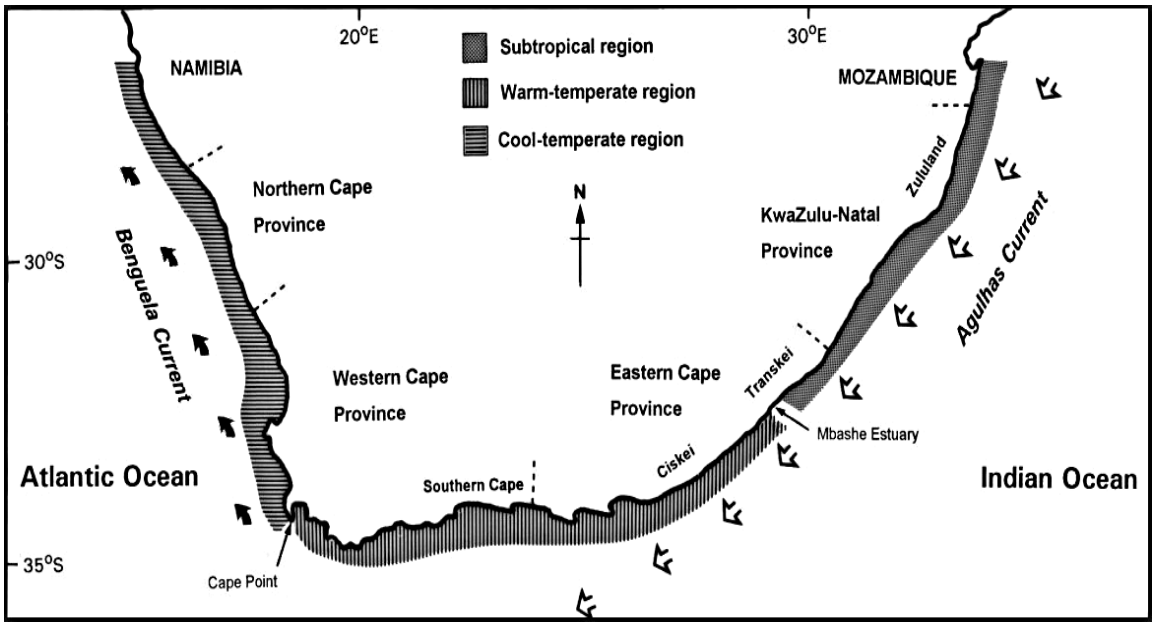

Figure 1.2: The South African coastline illustrating the three biogeographic regions. Also shown are the two main current systems (warm Agulhas and cold Benguela) that influence the South African coastline (from Whitfield 1998).

CHAPTER TWO

MATERIAL AND METHODS

2.1 Study area, sampling and preservation procedures

This chapter provides an overview of the material and methods used in the two chapters that follow:

- Chapter Three: The study of the molecular phylogeny of the South African endemic gobiid fishes of the genus *Caffrogobius*; and
- Chapter Four: Intraspecific genetic variation of *Caffrogobius gilchristi*.

The methods included direct sequencing of mitochondrial DNA regions (cytochrome oxidase subunit I (COI) and cytochrome *b* (cyt-*b*)). Analyses included the relevant genetic statistical techniques and use of computer software for gene and haplotype diversity as well as phylogenetic procedures. Morphological observations using light dissecting microscopes and scanning electron microscopy (SEM) were made for species identifications and comparisons.

For the study of the molecular phylogeny of *Caffrogobius* species, 1–10 specimens of each species of *Caffrogobius* were collected from localities listed in Table 2.1 (Fig 2.1). Two species of *Bathygobius* were sampled for outgroup taxa in the phylogenetic analysis.

For the study of the intraspecific variation of *C. gilchristi*, 2–21 specimens of *Caffrogobius gilchristi* were collected at the sampled sites listed in Table 2.2 (Fig 2.1). The specimens collected were frozen or preserved whole in 96% ethanol. A sample of muscle tissue was removed from the right side of each specimen and stored in 1.5 µl Eppendorf tube at -70˚C in the SAIAB biomaterial bank collection.

The localities for phylogeographic analysis were subdivided into four regions: KwaZulu-Natal (KZN), Eastern Cape (EC), southern Cape (SC) and Western Cape (WC) (Table 2.2). These localities ranged from the subtropical biogeographic region (Wild Coast in

the Eastern Cape and KZN) to the cool temperate zone (western and southern Cape). Specimens were collected using a 5 m beach seine net, with mesh size of 7 mm, but with the middle portion having a mesh size of 4 mm. The net was dragged by two individuals for about 50 m and then pulled out for observation. Specimens from the KZN coastal region were collected in a fast-flowing small freshwater tributary and in the turbid water of the upper reaches of the Umhlatuzi Estuary. Eastern and southern Cape specimens were collected on rocky, muddy and sandy bottoms, as well as on algae in estuaries and tidal rock pools. Some specimens from Gordon's Bay (WC) were collected in tidal rock pools and on a sandy beach. Geographical coordinates were obtained in the field using a Global Positioning System (GPS) handset for all sampling sites under the WGS84 system (Table 2.1). Additional samples for phylogenetic analyses within the complex species of *C. agulhensis* and *C. saldanha* were randomly selected from the National Fish Collection at the South African Institute for Aquatic Biodiveristy (SAIAB) for morphological verifications of second dorsal (D2) rays count (see Chapter three Table 3.1).

2.2 Laboratory procedures

2.2.1 DNA extraction

Total genomic DNA was extracted using a Promega Wizard® Genomic DNA Purification kit, following the manufacturer's animal tissue extraction protocol. The initial stage of this process involved an overnight, or three-hour incubation at 55˚C in an EDTA/Nuclei Lysis and Proteinase K solution. An RNAse treatment was included by adding 0.1 mg RNAse A (Roche Diagnostics) to each sample, followed by incubation at 37˚C for one hour. The total genomic DNA from each sample was eluted in 50 microlitres (µl) ultra pure water and stored at 4°C. The DNA concentration and extraction was verified by running it on a 1% low-melting-point agarose gel, stained with Ethidium bromide and visualized under UV light.

2.2.2 PCR amplification and sequencing

The extracted DNA was used as a template for the amplification of the partial sequences of the COI and cyt-*b* mtDNA regions by means of polymerase chain reactions (PCR). PCR amplification and sequencing was done to determine the nucleotide sequence of the DNA fragments. Specific primers for the cyt-*b* and COI regions were selected, based on published sequences of closely related species and universal primers for fish species, after testing them for desirable results in these gobies (see Table 2.3).

All PCR reactions were performed in a total volume of 50 µl in thin-walled 200 µl micro-centrifuge tubes. In addition to the DNA template, the reaction mixture contained 1.5 mM $MgCl_2$, 5 µL of the 10x reaction buffer, 0.2 mM of each of the four nucleotides (dNTP's; Promega), 1.0 unit Super-therm® Taq DNA polymerase (Southern Cross Biotechnology, South Africa) and 50 pmol of each of the respective forward and reverse primers, in accordance with each of the particular regions amplified.

2.2.2.1 Cytochrome-*b* (cyt-*b*) gene amplification

The amplification conditions for the cyt-*b* region included aliquots of 2-5 µl DNA templates for the PCR procedure in a Thermo Hybaid Px2 Thermal Cycler and/or Multiblock System. For phylogenetic analysis, a 281 base pairs (bp) segment from 28 specimens of the cyt-*b* region of mitochomdrial DNA was amplified using PCR with Oscytb-F1 and Acytb-R1 primers (Thacker & Hardman 2005) (Table 2.3). To estimate intraspecific gnetic variation of *Caffrogobius gilchristi* speciemens, 425 fragments of cyt-*b* were amplified for 63 specimens using primers Oscytb-F1 and Acytb-R1 (Table 2.3). PCR cycling conditions involved an initial denaturing step of 4 minutes (min) at 94ºC, followed by 35 cycles of 30 seconds at 94ºC, annealing for 1 min at 56.2ºC, extension for 1 min at 72ºC, and a final elongation step for 7 min at 72ºC before holding at 4°C. Amplification was always carried out with a no-template control to test for contamination. All PCR reactions were checked on a one percent low-melting-point agarose gel, stained with ethidium bromide, and visualized under UV light. PCR purification using a Qiagen PCR Purification Kit was done following the manufacturer's instructions and was completed with elution in 10 µl of DNA-free water. Cycle-

sequencing was performed in 20 µl volumes with the reaction mixture containing 2 µl purified PCR template, 2 µl of one primer diluted 10 times, 2 µl 5X PCR buffer, 10 µl of water and 4 µl BigDye Terminator Sequencing Ready Reaction V 3.1 kit (Applied Biosystems) under the following conditions: 94ºC for 1 min followed by 25 cycles at 94 ºC for 10 seconds, 48 ºC for 5 seconds, and 72 ºC for 7 min. Samples were precipitated using 2 µl EDTA, 2 µl 3M Sodium Acetate and 50 µl 100% EtOH. Automated thermal sequencing was performed on an ABI 3100 sequencer at the Rhodes University sequencing facility and Macrogen (South Korea).

2.2.2.2 Cytochrome oxidase subunit I (COI) gene amplification

The amplification conditions for the COI region also included aliquots of two-five µl template DNA. A 535 bp fragment of the COI was amplified for *Caffrogobius* species for phylogenetic analyses (28 specimens) while a 524 bp fragment was amplified for intraspecific genetic variation of *C. gilchristi* (109 specimens) using primers FishR1 and FishF1 (Ward *et al*. 2005) (Table 2.3) in a Thermo Hybaid PX2 Thermal Cycler and/or Multiblock System. PCR cycling conditions involved an initial denaturing step of four minute at 94 ºC, followed by 35 cycles of 30 seconds at 94 ºC, annealing for 1 min at 54.5 ºC, extension for one minute at 72 ºC, and a final elongation step for 7 min at 72 ºC. PCR purification and cycle-sequencing were performed according to the procedure described for the cyt-*b* region.

2.3 Statistical analyses

2.3.1 Sequence alignment

Obtained sequence chromatogram files of specimens were imported into SeqMan Pro (DNASTAR® Lasergene® 7.2) to correct for possible base ambiguities and compute consensus sequences after alignment. All of the *Caffrogobius* sequences were analysed along with the sequenced COI/cyt-*b* regions of *Bathygobius* species, the only available close relative sampled species along with *Caffrogobius* species for the project. This was done in order to verify the origin of the obtained sequences. Alignment of the consensus

was performed with the multiple alignment programme BioEdit version 7.0.5.3 (Hall 1999) using default settings.

2.3.2 Phylogenetic analyses

2.3.2.1 Molecular phylogeny estimation

To estimate a robust phylogenetic tree, three tree-based phylogenetic reconstruction methods (parsimony (MP) (Swofford 2001), maximum likelihood (ML) (Felsenstein 1988) and Bayesian analysis (BI) (Ronquist & Huelsenbeck 2001) were used for the analyses of the partial mitochondrial DNA (cyt-*b* and COI) datasets. Insertions and deletions were treated as missing data in all tree construction methods. To test if the sequences from the two genes could be combined in a single analysis, the partition homogeneity test in PAUP* version 4.0b10 (Swofford 2001) was utilized to test for incongruence. The significance was estimated at 1000 homogeneity replicates and the heuristic search option was utilized without branch-swapping.

2.3.2.1.1 Parsimony, maximum likelihood and Bayesian inference

Parsimony methods are based on character values for each species instead of the distance between sequences (Kvist 2000). MP analyses were performed with the heuristic algorithm (100 and 10 random addition replications) (Felsenstein 1985). Tree bisection and reconnection (TBR) tree searches were done for all parsimony analyses with bootstrap (1000 pseudoreplications) analysis to examine levels of the relative support for inferred monophyletic groups (Felsenstein 1985).

In contrast to MP, ML trees are reconstructed to find the most likely relationship for each set of gene sequences. The procedure is highly dependent on the order of sequences and many intermediate trees are constructed, from which a majority rule consensus tree is built in the consensus step (Schmidt & Rand 2001). ML employs standardized statistical methods for probabilistic evolution models by considering the changes that are more likely along the long branches than the short branches (Kvist 2000).

Both MP and ML were performed in PAUP* version 4.0b10 (Swofford 2001). Heuristic search procedures for the best ML tree were similar to MP searches. However, ML analyses incorporated MODELTEST version 3.7 (Posada & Crandall 1998) to select the most appropriate DNA substitution models among 56 different models. The Akaike Information Criterion (AIC) was employed to do comparisons between these alternative models of evolution and to determine the best fit substitution model for the data sets. The Ti:Tv ratio, base frequencies, proportion of invariable sites (I), and the α value of the gamma distribution (rate of variation among sites) were also determined (Posada & Crandall 1998). The models obtained (Table 2.4) were consequently used to determine the maximum likelihood relationships of species of *Caffrogobius* using generated sequences of *Bathygobius* species as an outgroup. This genus was selected as the out-group since the sister-group (or close relative) of *Caffrogobius* is unknown, but *Bathygobius* has been suggested as a possible sister group to *Caffrogobius* (Goren 1996). Additionally freshly collected specimens of *Bathygobius* made successful DNA extractions possible. Heuristic searches were done in PAUP to find the topology with the best likelihood score using the selected model, TBR with steepest descent option, and 100 random addition sequences of available species.

Bayesian Markov Chain Monte Carlo (MCMC) methods do not attempt to solve the maximum likelihood as defined in ML analyses. Instead, the resultant tree that is retained from BI maximizes the probability of the data. Bayesian analyses, on the other hand, can analyse each dataset using different heuristic searches and record the best score found by each heuristic search under various conditions (Ronquist & Huelsenbeck 2001). Bayesian inference analyses were performed using MrBayes version 3.1.2 (Ronquist & Huelsenbeck 2001). The MCMC process was set for four chains to run simultaneously for one million generations, with trees being sampled every 500 generations for a total of 20 000 trees in the initial sample. The first 25 % of the trees were then discarded as the burn-in period, and the remaining 15 000 trees were used to construct a majority rule consensus tree with posterior probabilities.

2.3.2.2 Morphological analyses

The relationships within the Gobiidae are not clear, specifically the relationships of the southern African endemic gobies, with the sister group of the genus *Caffrogobius* being unknown. Unfortunately, in this study it was not feasible to include all possible genera that have been reported as close relatives (Hoese & Winterbottom 1979a, b; Hoese & Gill 1993; Goren 1996).

The extreme morphological similarity between *C. saldanha* and *C. agulhensis* makes it very difficult to separate these two species, using the identification keys of Goren (1996). The number of rays in the second dorsal fin (D2) appears to be a key character to separate the two species (Goren 1996). To test if there is a statistically significant difference between the counts of this character in the two species, 36 randomly selected specimens, from voucher specimens of *C. agulhensis* and *C. saldanha* in the SAIAB collection facility were used to count the number of rays in the second dorsal fin (D2) (Table 3.1). A correlation test was conducted using software STATISTICA version 8 (StaSoft 2008). The Scatterplot, 2D–Regular was used to correlate the data (D2 of *C. agulhensis* vs D2 of *C. saldanha*) during the analyses. The regular scatterplot visualizes a relation between two variables, *X* and *Y* (e.g., weight and height). The two coordinates (*X* and *Y*), which determine the location of each point, correspond to specific values of the two variables. If the two variables are strongly related, then the data points form a systematic shape (e.g., a straight line or a clear curve). If the variables are not related, then the points form an irregular "cloud".

Additionally, in scrutinizing key characters that can separate the two species (*C. saldanha* and *C. agulhensis*) the character of papillae row **k** stands out. According to the identification key (Goren 1996); papillae row **k** is double in *C. saldanha* and single in *C. agulhensis*. Even with the help of the scanning microscope, it was difficult to see the double or single rows of papillae. To validate this crucial character, 20 specimens selected from this study and 20 additional specimens, from SAIAB fish collections facility, were examined. The specimens were prepared following the SEM protocol for

scanning electron microscopes. The sensory papillae, situated on the sides of the head, were used for analysis. One side was cut, depending on which side had the better visibility. SEM micrographs taken at the Rhodes University electron microscopy unit were used to examine cheek papillae characteristics. Acid blue staining methods were applied to the museum specimens to assist in the counting of visible papillae.

2.3.3 Intraspecific genetic structure analysis of *Caffrogobius gilchristi*

Genetic diversity within populations was estimated, using the software DnaSP 4.10 (Rozas *et al*. 2005), as follows: by haplotype (gene) diversity (*h*), i.e. the probability that two randomly chosen alleles are different in a sample; nucleotide diversity (π), i.e. the average number of nucleotide differences per site between two sequences; population comparison, and the mean number of pairwise differences (Nei 1987). The haplotype frequencies estimated from the pairwise differences in Network v.5.0 package (Polzin & Daneschmand 2003) were used to construct a minimum spanning network for a visual examination of genealogical relationships among haplotypes.

Hierarchical analysis of molecular variance (AMOVA) which produces estimates of variance components and F-statistic analogs reflecting the correlation of haplotype diversity at different levels of hierarchical subdivision (Excoffier *et al.* 1992), was performed to compare levels of genetic diversity within and among several possible population groupings of prison goby using ARLEQUIN 3.1.1 (Excoffier *et al.* 2006) with 1000 permutations. The groupings that will maximize values of Φ_{CT} and are found to be statistically significant, indicate the most parsimonious geographical subdivisions. Significance levels were obtained using a permutation approach with 1000 iterations. The following components of diversity were employed: “Within Populations” (WP), “Among Populations/Within Groups” (AP/WG), and “Among Groups” (AG) (Excoffier *et al*. 1992). The ‘grouping by genetic structure’ approach requires an *a priori* definition of group structure in order to group the population sets together to form different hierarchical levels in the analysis (Norton 2005). In this study four biogeographical regions were proposed, i.e. KwaZulu-Natal (KZN) (Umhlatuzi), Eastern Cape (EC) (Umzumvubu, Gonubie, Keiskamma, East Kleinemonde, Kowie, Gamtoos, Kromme),

Southern Cape (SC) (Keurbooms, Swartvlei, Groot Brak) and Western Cape (WC) (Gordon's bay) (Table 2.2). By using ARLEQUIN version 3.1.1, genetic differentiation between geographical subdivisions was analysed by comparing average numbers of pairwise differences between distribution ranges, average number of pairwise differences within populations, and the corrected average pairwise result.

To test the hypothesis of demographic expansion, the distribution of pairwise differences between haplotypes (mismatch distribution) with its parameters τ, neutrality tests, the Tajima's D-statistic (Tajima 1989), Fu's Fs test (Fu 1997) and R^2 test in order to depict the demographic history were computed in ARLEQUIN 3.1.1 and tested for significance with permutation tests (1000 replicates) for the sequences of both mtDNA genes (cyt-*b* and COI). Additionally the value of Ramos-Onsins & Rozas R_2 (Ramos-Onsins & Rozas 2002) was determined in DnaSP 4.10 (Rozas *et al*. 2005) since R_2 is a good statistical test for detecting population growth in small sample sizes (Ramos-Onsins & Rozas 2002).

According to Rand (1996), a mixture of tests of neutrality can serve as indicators of demographic parameters such as population growth or decline in the evolutionary history of taxa. Fu's F_S test was used to test for the possibility of population expansion for each sampling location and the combined dataset. A positive Tajima's *D* value may either indicate a selective sweep, a population expansion after a bottleneck, or a smaller founder event. Furthermore, positive values are consistent with models of positive and balancing selection, or an admixture of distinct and isolated populations (Rand 1996). On the other hand, a negative value of the *D* or Fu's test could be evidence against neutrality of mutations and when there is an excess of recent mutations in a species (Fu 1997). R_2 is very powerful in detecting population expansions from small samples (Ramos-Onsins & Rozas 2002). If a population expansion was detected, we estimate its age according to the following equation modified from Rogers & Harpending (1992). According to Rogers & Harpending (1992), the time expansion formula is: $T = \tau/2\mu$ (where μ = generation time x number of base pairs per sequence x mutation rate for marker used, and τ = expansion calculated from ARLEQUIN 3.1.1 version). The generation time used for *C. gilchristi* was adopted from Neethling *et al*. (2008).

To test for isolation by distance or biogeographic distribution, the spatial analysis of molecular variance (SAMOVA Dupanloup *et al*. 2002) principle was applied. The biogeographic groupings used for AMOVA, i.e. KZN, EC, SC and WC (see Table 2.2) were also applied for SAMOVA analysis. SAMOVA was run with the *K* value (number of populations) from two to five groupings.

Table 2.1: The locations, geographical coordinates and laboratory identifications codes of specimens used in the phylogenetic study of *Caffrogobius* species.

Taxon	Locality	GPS	Lab ID
Bathygobius sp1	Umhlatuze	28°46.86S/31°58.59E	BTGG07
Bathygobius sp2	Umhlatuze	28°46.86S/31°58.59E	BTGG08
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP01
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP02
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP03
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP04
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP05
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP06
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP07
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP08
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP09
C. agulhensis	Port Elizabeth	33°52.67S/25°38.22E	CAMP10
C. saldanha	Cape Recife	34°02.14S/25°38.66E	CAPE11
C. saldanha	Cape Recife	34°02.14S/25°38.66E	CAPE12
C. saldanha	Cape Recife	34°02.14S/25°38.66E	CAPE13
C. saldanha	Cape Recife	34°02.14S/25°38.66E	CAPE14
C. saldanha	Cape Recife	34°02.14S/25°38.66E	CAPE15
C. caffer	Gansbaai	28°22.52S/32°25.07E	CCGA10
C. caffer	Gansbaai	28°22.52S/32°25.07E	CCGA11
C. caffer	Gansbaai	28°22.52S/32°25.07E	CCGA12
C. gilchristi	Swartvlei	34°00.21S/22°48.31E	CGSW01
C. gilchristi	Umzumvumbu	31°34.69S/29°29.55E	CGUM01
C. gilchristi	Umhlatuze	28°46.86S/31°58.59E	CGUR04
C. nudiceps	Gordon's Bay	34°20.54S/18°49.28E	CNCA01
C. nudiceps	Keiskamma	33°11.09S/27°23.56E	CNKK10
C. saldanha	Langebaan	29°43.44S/31°05.21E	CSLB01
C. saldanha	Langebaan	29°43.44S/31°05.21E	CSLB02
C. natalensis	Gouritz	34°21.34S/21°53.34E	CTGZ02

Table 2.2: Localities where *Caffrogobius gilchristi* specimens were collected, estuarine ocean access and number of specimens successfully amplified per marker.

Geographical grouping	***Locality***	***Ocean access***	***Number of specimens***	
			cyt *b*	COI
KwaZulu-Natal (KZN)	Umhlatuze River	Temporal	3	3
Eastern Cape (EC)	Umzumvubu River	Temporal	3	4
	Gonubie River Estuary	Temporal		2
	Keiskamma River Estuary	Temporal	6	6
	East Kleinemonde River Estuary	Temporal		6
	Kowie River Estuary	Permanent	15	23
	Gamtoos River Estuary	Permanent	8	21
	Kromme River Estuary	Permanent	2	2
Southern Cape (SC)	Keurbooms River Estuary	Permanent	13	14
	Swartvlei River Estuary	Permanent	12	11
	Groot Braak River Estuary	Permanent		4
Western Cape (WC)	Gordon's Bay	Permanent		10
	TOTAL		63	109

Table 2.3: The primers used to amplify the partial genes COI (Ward *et al*. 2008) and cyt-*b* (Thacker & Hardman 2005).

Primer	**Sequence (5' – 3' direction)**	**Gene region**
FishF1*	TCAACCAACCACAAAGCATTGGCAC	COI
FishR1*	TAGACTTCTGGGTGGCCAAAGAATCA	COI
OsCytb-F1	CACCCATACTTCTCMTAYAAAGA	cyt *b*
Acytb-R1	TCCGGATTACAAGACCGGYGCTTT	cyt *b*

Table 2.4: Summary of the nucleotide composition and the selected model of evolution for each of the mitochondrial DNA genes sequenced as well as for the combined data set.

Data set	**Nucleotide composition (%)**				**AIC Model**
	A	C	G	T	
COI	0.2353	0.3823	0.1591	0.2382	TIM+I+G
Cyt-*b*	0.2204	0.3202	0.1839	0.2607	TrN+I+G
Combined data	0.2315	0.3452	0.1758	0.2474	TrN+I+G

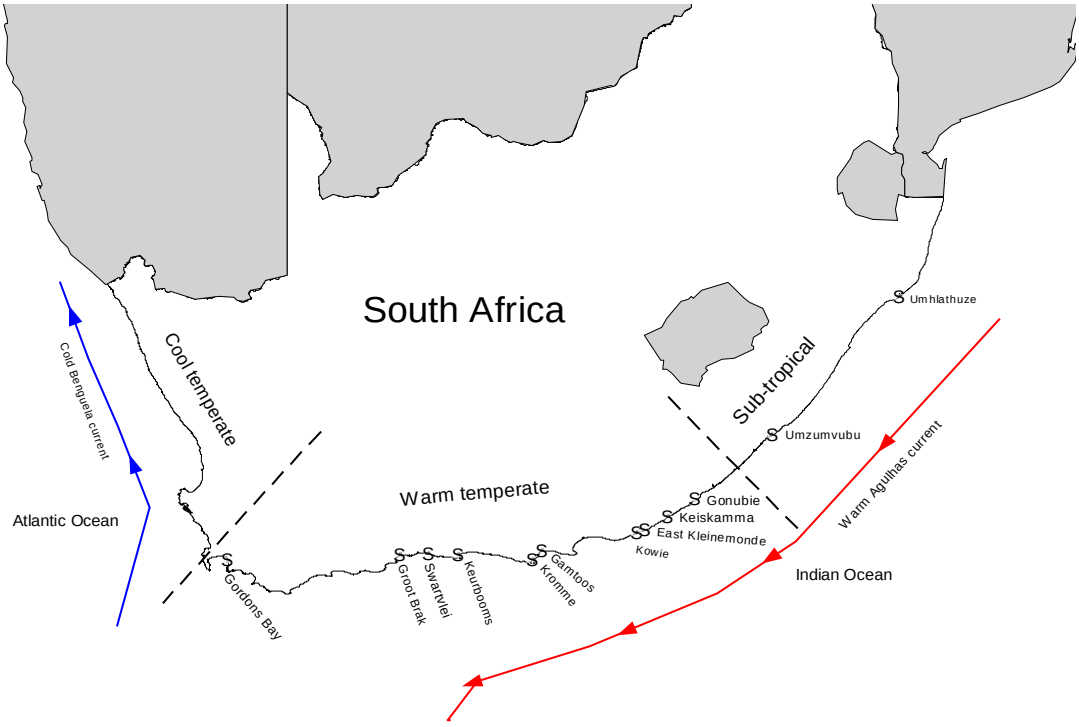

Figure 2.1: Map indicating the locations sampled for the study of phylogeography of *Caffrogobius gilchristi* and phylogeny of *Caffrogobius*. The blue line indicates the cold Benguela Current and the red line indicate the warm Agulhas Current. The dotted line separates the biogeographical regions, i.e. the cool temperate, warm temperate and the sub-tropical regions. Small circles along the coastline represent sampling locations (Gordon's Bay to Umhlatuze).

CHAPTER THREE
PHYLOGENETIC RELATIONSHIPS OF THE SOUTHERN AFRICAN ENDEMIC GOBIES OF THE GENUS *CAFFROGOBIUS* SMITT, 1900 (TELEOSTEI, GOBIIDAE) BASED ON MOLECULAR DATA

3.1 Introduction

In southern Africa, the Gobiidae represents a speciose family of about 51 genera, comprising 107 species, found primarily in inshore marine habitats, estuaries, coastal lakes and freshwater habitats (Smith & Heemstra 1986; Skelton 1994; Whitfield 1998). The gobies of the genus *Caffrogobius* are restricted to southern Africa except for one species described from Seychelles. They occur in shallow water habitats of estuaries and the coastal marine environment throughout the tropical, subtropical and temperate regions of southern Africa (Goren 1996; Whitfield 1998). Species of *Caffrogobius* appear morphologically similar to such an extent that observations of diagnostic features — such as the cheek papillae pattern and other features commonly used for gobioid taxonomy — are often not useful in identification of species. Goren (1996), in the only morphological review of the genus, identified seven nominal species (*Caffrogobius saldanha, C. agulhensis, C. gilchristi, C. natalensis*, *C. caffer*, *C. nudiceps* and *C. dubius*) and noted a high degree of similarity in scale counts, color, and sensory papillae patterns among some species of this genus. He further raised doubt with regards to character variation within *C. nudiceps*. Goren (1996) further noted the existence of only a single record of *C. dubius* in Seychelles, the validity of which required further verification.

The phylogenetic relationships among species of *Caffrogobius* is problematic since misidentifications among species of *Caffrogobius* are common in fish collections, which have led to a high number of synonyms (Goren 1996) and poor understanding of species relationships and biology. To date, there has been no appropriate outgroup recognized for the genus *Caffrogobius*, due to the lack of molecular and morphological phylogenetic studies for both family and genera levels (Akihito *et al*. 2000). The genus is presently characterised by a combination of the following characters: cephalic sensory system with cheek papillae in a transverse pattern (with the number of papillae of certain rows, as well as their arrangement, often varying among individuals of the same species); dorsal fin with segmented rays numbered 9-14, anal fin with segmented rays numbered 8-12; caudal fin with segmented rays

numbered 16-18; pectoral fin with segmented rays numbered 16-24; longitudinal scales along the body numbered 30-64; transverse rows of scales numbered 9-23, and vertebrae numbered 27 (11+16 or 10+17) (Hoese & Winterbottom 1979a,b; Hoese 1986; Goren 1996). The cephalic sensory system (papillae) has been recognised as a valid character for identification of gobies; hence the unique features used to separate species of the genus *Caffrogobius* are based on differences in this character. For example, *C. saldanha* and *C. agulhensis* can be distinguished by the characteristics of papillae in row **k**: having a double row and single row, respectively. Nevertheless, this character has some shortcomings that make it unreliable for separating these two species, for instance, some specimens of *C. samdanha* have single rows.

As previously stated, the relationship of *Caffrogobius* to other genera is not well understood: for instance, Miller (1978), Birdsong *et al*. (1988), Gill & Bradley (1992) and Goren (1996) have classified the genus together with several other closely-related genera. The sister genera of *Caffrogobius* can be traced to the following genera: *Gobius*, *Zebrus*, *Gammogobius* Bath, 1971, *Monishia*, *Bathygobius*, *Glossogobius*, *Gorogobius*, *Heteroleotris*, *Lesuerigobius*, *Nematogobius*, and *Corygalops* (Miller 1978; Birdsong *et al*. 1988; Gill & Bradley 1992; Goren 1996). Finding the sister group of *Caffrogobius* will help to elucidate phylogenetic relationships among species of *Caffrogobius*. The lack of phylogenetic studies on genera level might be due to high species diversity and morphological character reduction (Matsubara & Iwai 1959; Iwata *et al.* 2001; Larson 2001; Thacker 2003). Recently studies done to estimate the relationships amongst members of Gobioidei included only a few genera of the Gobiinae e.g. Akihito *et al*. (2000) no genus, Wang *et al*. (2001) six genera, Thacker (2003) 16 genera and Thacker (2009) five genera. Thus, a more intensive investigation of the phylogeny of the genera of the subfamily Gobiinae is necessary and urgently required (Wang *et al.* 2001). Although gobies are a prominent component in many regional fish faunas, not many studies of their phylogenetic relationships have been undertaken, because goby species are generally cryptic and difficult to identify. The presence of cryptic taxa can cause confusion since although they appear to be morphologically similar, they may be genetically distinct. The introduction of DNA techniques now assist with the delineation of cryptic forms and can provide further insights and a greater understanding of evolutionary history and life history styles. Recent systematic studies on related groups have made use of molecular data to separate cryptic species as well as morphologically-similar species that are difficult to identify (Close & Gouws 2007; Harold *et al* 2008). For example, new species of the genus *Gobiodon* were recently identified using molecular markers (Harold *et al*. 2008), while

allozyme markers were used to identify cryptic estuarine goby larva of *Pseudogobius olorum* (Close & Gouws 2007). Because of its rapid rate of sequence evolution, mitochondrial DNA (mtDNA) has been more successfully used in reconstructing phylogenetic relationships among closely related species than has been the case for nuclear DNA (Brown *et al.* 1979). It has also been observed that large data sets derived from a combination of genes have the best potential for resolving problematic issues in taxonomy (Muller 2004). Thacker (2003) estimated a molecular phylogeny of gobioid fish's families by using complete sequences of mitochondrial ND1, ND2 and COI genes. Thacker (2003) noted that the current classification of gobies reflects the uncertain knowledge of goby relationships and that, despite the availability of morphological data, no cladistic analysis of gobioid families or genera has previously been reported. A molecular phylogeny of basal gobioid fishes (Rhyacichthyidae, Odontobutidae, Xenisthmidae and Eleotridae) using complete sequences of the mtDNA ND1, ND2, COI and cyt-*b* genes by Thacker & Hardman (2005) later indicated that the gobies are a monophyletic group and that the basal taxon of this group is known from freshwater of the Indo-Pacific region, but with marine-dwelling lineages. In this study, the cyt-*b* and COI mtDNA genes were used as phylogenetic markers to investigate the phylogeny of *Caffrogobius* species.

The cytochrome-*b* gene is probably the best studied mtDNA gene in fishes (Avise *et al*. 1987 Meyer 1994). It has been used as the basis for a number of studies on the phylogenetics of gobies e.g. the phylogeography of the tidewater goby, *Eucyclogobius newberryi* Girard, 1856 (Dawson 2001), the genetics of Ponto-Caspian gobies (Stepien & Tumeo 2006), and the phylogeny of southern Africa *Glossogobius* species (Maake 2009). The COI gene is also being used to study the phylogeny of fishes (Thacker 2003; Thacker & Hardman 2005), as evidenced by its recent application in DNA barcoding projects (Ward *et al*. 2005).

In the present study, these two genes (cytochrome-*b* and the COI gene) were chosen for phylogenetic studies because they are the largest mtDNA-encoded protein-coding genes that are informative in phylogenetic studies of fishes (Avise et al. 1987; Meyer 1994; Stepien & Tumeo 2006) and are thus likely to provide a valid estimation of *Caffrogobius* phylogeny.

3.2 Material and methods

Details of the material and methods used in this study were given in Chapter Two.

3.3 Results

3.3.1 Phylogenetic analyses of relationships within *Caffrogobius* species

The estimated mtDNA phylogenetic trees were rooted using *Bathygobius* species as an outgroup taxon. After sequence amplification, *Caffrogobius* and *Bathygobius* sequences were edited and aligned prior to analyses. The aligned data set included sequences of 28 individuals per gene (cyt-*b* and COI respectively) of six species. Analyses were carried out for each gene region separately, followed by an analysis of the combined dataset since the partition homogeneity test indicated no conflicting phylogenetic signals ($p = 0.63$). No codon deletion or insertion needed to be invoked for aligning the sequences. The selection of a model of evolution for character partitions under the likelihood ratio test in PAUP* led to the implementation of a model using the standard Akaike Information Criterion (AIC).

3.3.2 Cytochrome oxidase subunit I

The COI region data analyses produced an aligned data set of 535 base pairs (bp) of which 368 were constant (68%), 14 were variable but parsimony uninformative (3%), and 153 were parsimony informative (29%). The data set had an unequal frequency of bases (A = 23.53%; C = 38.23%; G = 15.91%; T = 23.82%). Phylogenetic inferences under the parsimony optimality criterion resulted in two optimal trees. The trees were 251 steps long and characterised by the following indices: consistency index = 0.81, consistency index for parsimony informative characters only = 0.73 and retention index = 0.90. The estimated MP tree (Figure 3.1) comprised of four recognisable clades (A, B, C, and D). Clade D represented specimens of *C. natalensis*, which appears as a basal species of *Caffrogobius*. Clade C consisted of *C. gilchristi* and *C. caffer,* which are sister groups with 78% bootstrap support. There was a 99% bootstrap support for Clade A that was made up of *C. saldanha* and *C. agulhensis* specimens. Clade B conssting of *C. nudiceps* is a well supported node with 100% bootstrap support and these specimens form the sister group of Clade A (*C. saldanha* and *C. agulhensis)*, with 96% bootstrap support.

The MLtree (not shown) was identical and congruent to the MP topology that supported the monophyletic grouping of *Caffrogobius*. The optimal tree had a likelihood of lnL -1502.95. *Caffrogobius natalensis* retained the basal position in the ML tree to support the MP tree. The BI tree also agreed with the MP topology on the monophyly of the genus *Caffrogobius.*

However, the BI analysis proposed different phylogenetic relationships between the species of this genus (Figure 3.2) suggesting that *C. gilchristi* is the basal group in the genus, but still supporting that *C. gilchristi* is a close relative of *C. caffer*. The clade consisting of *C. saldanha* and *C. agulhensis* is well supported with >98% posterior probability.

3.3.3 Cytochrome-*b*

Partial cyt-*b* gene sequences of 281 nucleotides in length were obtained for each of the 28 specimens. The cyt-*b* data set had 198 constant characters (71%), 12 variable parsimony-uninformative characters (4%), and 71 parsimony-informative characters (25%). The data set had an unequal frequency of bases (A = 22.04%; C = 32.02%; G = 18.39%; T = 26.07%). The MP tree (Figure 3.3) (TL = 159, CI = 0.77, CI for parsimony informative characters = 0.69 and RI = 0.89) was congruent with the ML tree (lnL -786.62, not shown), but was different from the BI tree (Figure 3.4) in which the *C. caffer* and *C. gilchristi* are sister groups rather than *C. caffer* being basal to *C. gilchristi*. The genus *Caffrogobius* also appears to be monophyletic with a 100% support. Clade A had 100% bootstrap support and 98% posterior probability and *Caffrogobius natalensis* was the basal species in these phylogenies (Figure 3.3 and Figure 3.4). These topologies differ from the ones estimated using COI. A notable difference was the relationship of *C. gilchristi* and *C. caffer*, which are still closely related, but not sister taxa as in the MP tree (Figure 3.1) and with *C. caffer* more basal (Figure 3.3) rather than *C. gilchristi* (Figure 3.2).

3.3.4 Combined analyses

The partition homogeneity test indicated no conflicting phylogenetic signals between the two genes (p = 0.63). The data sets of cyt-*b* (281 bp) and COI (535 bp) mitochondrial genes were combined (816 total characters) in the estimation of the *Caffrogobius* species relationships in an attempt to provide better phylogenetic resolutions from the three methods used. The combined data set of 816 characters of the two mitochondrial DNA genes (COI and cyt-*b*) had 583 constant characters (71%), 30 variable parsimony-uninformative (4%) and 203 parsimony-informative (25%) characters. The combined data set had an unequal frequency of bases (A = 23.15%, C = 34.52%, G = 17.58% and T = 24.75%). The phylogenetic analyses of the combined sequences using all three phylogenetic methods (MP, ML and BI) were highly resolved with the exception of the terminal group of *C. saldanha* and *C. agulhensis*. The parsimony phylogenetic analysis yielded 23 trees with a tree length of 415 and consistency index of 0.78, retention index of 0.89, and a re-scaled consistency index of 0.70.

The 50% majority rule consensus topology (Figure 3.5) is more resolved than those of the individual genes. The genus *Caffrogobius* is monophyletic with 100% bootstrap support, as was the case with the individual genes. *Caffrogobius caffer* and *C. gilchristi* are sister groups with *C. natalensis* being basal. The ML analysis (see Table 2.4) is characterised by an unconstrained likelihood score of –ln L = 2405.24. The ML topology (Figure 3.6) is congruent with the MP topology. The BI (see Table 2.4) tree (Figure 3.7) is also congruent to the previous two with 100% posterior probability for the monophyly of the genus as well as Clade A. Thus all estimated topologies using MP, ML and BI are congruent. All three trees support *C. natalensis* as the basal group, as well as the sister grouping of *C. gilchristi* and *C. caffer* and *C. nudiceps* the sister groups of Clade A (comprising of *C. saldanha*/*agulhensis*).

3.3.5 Morphological observations

There is a high degree of morphological similarity among species (as also noticed by Goren 1996). The basal taxon, *C. natalensis*, differs from all the other members of the genus, except *C. gilchristi*, in the absence of a skinny flap on the anterior nostril (Goren 1996). In Figure 3.7, Clade C contains species that are different with respect to habitat, with *C. caffer* being a marine coastal species while *C. gilchristi* is an estuarine-dependent species. Morphologically, *C. caffer* differs from *C. gilchristi* by the absence of predorsal scales and the paired arrangement of the dark bars on the body (Goren 1996). The main distinguishing characteristic of *C. nudiceps* (Clade B in Figure 3.7) is a wide yellowish band on the proximal part of upper pectoral fin rays and the presence of one to five predorsal scales, although the mid-predorsal area is naked in some some specimens (Goren 1996). The individuals forming Clade A (Figure 3.7) consisted of 17 individuals that had a black spot at the rear of the first dorsal fin, an elongated flap anteriorly on the nostrils and papillae of moderate size on the cheek. The cephalic papillae of row **k**, which according to Goren (1996), is supposed to be a double row in *C. saldanha* and a single row in *C. agulhensis* could not be used to differentiate these specimens, since no clear pattern could be distinguished among the 40 examined specimens. The papillae are of moderate size on the cheeks of *C. saldanha* and *C. agulhensis* specimens (Clade A members) and minute on the other southern African members of the genus. It is the main decisive character for both species morphologically (Goren 1996).

3.4 Discussion and conclusion

This study represents the first molecular investigation about the phylogenetic relationships of the gobies of the genus *Caffrogobius*, using mitochondrial DNA markers. Mitochondrial DNA has become a common source of markers in studies of closely-related taxa because of its haploid nature and lack of recombination (Avise 2004). In the present study, 816 characters from two partial mitochondrial DNA genes, cytochrome-*b* (cyt-*b*) and cytochrome oxidase subunit I (COI), provided a strong phylogenetic signal both when analysed separately and in combination. Both genes are regularly used to infer phylogenetic relationships in gobies and other fishes (Akihito *et al*. 2000; Thacker & Hardman 2005; Maake 2009).

The COI region amplified in this study yielded 535 characters. Monophyly of *Caffrogobius* was strongly supported (100% MP bootstrap support and 100% posterior probability). Clade A (*C. saldanha* and *C. agulhensis*) is also fairly strongly supported (MP 99% bootstrap support and BI 98% posterior probability). *Caffrogobius saldanha* and *C. agulhensis* are both marine species that are morphologically very similar both with a black spot on the dorsal fin and elongated flap on the nostrils (Goren 1996). Goren (1996) reported that *C. agulhensis* differs from other members of the genus by possessing double horizontal lines of black dots on the dorsal fins, while it differs from *C. caffer*, *C. natalensis*, *C. nudiceps* and *C. gilchristi* by possessing anterior nostrils with an elongated flap. *Caffrogobius saldanha* differs from these species by the possession of a black spot at the rear of the first dorsal fin and apparently it differs from *C. agulhensis* by having a double row of papillae **k**. However, the overlapping counts make it difficult to separate some individuals of these species (Table 3.1) and thus the current analysis suggest that there is a need of alternative characters for proper species identification between *C. agulhensis* and *C. saldanha*. *Caffrogobius nudiceps* is the sister group of these two species in all estimated topologies. The majority of the estimated topologies (all but Figure 3.2), strongly support estuarine-dependent *C. natalensis* as the ancestral taxon for the investigated *Caffrogobius* species.

The three tree construction methods used in the analyses of the two individual marker genes and a combination of the two genes resulted in highly congruent topologies, with comparable nodal support. In the combined data set, the BI analysis indicated a 100% posterior probability for monophyly of the genus (Figure 3.7). Since posterior probabilities are known to overestimate confidence in a given clade (in line with the general statistical conventions)

only posterior probability of ≥ 95% are regarded as indicative of significance (Huelsenbeck & Ronquist 2001).

The molecular phylogenetic trees (Figure 3.1 to Figure 3.7) presented in this study are proposed as a working hypothesis for genealogical relationships among the *Caffrogobius* species. In all the resultant trees from the analyses, it is apparent that the genus is monophyletic. According to the morphological characteristics of these species, the monophyletic result can be expected, as most species in the ingroup have a high degree of similar and overlapping of meristic counts (Goren 1996). Among the six species examined, *C. natalensis* diverged first (except in the topology represented in Figure 3.2.). This cannot be explained in terms of habitat preference since C. *natalensis, C. gilchristi* and *C. nudiceps* are estuarine species with *C. caffer* (sister group of *C. gilchristi*) a marine species. The monophyletic group of *C. caffer* and *C. gilchristi* diverged next followed by *C. nudiceps* and the group consisting of *C. saldanha* and *C. agulhensis* as a terminal group, (except in the analysis represented in Figure 3.2). Among other well-supported results from our analysis is the resolution of Clade A (99 - 100% bootstrap support, 99-100% posterior probability Figure 3.1, 3.3, 3.4, 3.5, 3.7). Members of this clade has a high degree of similarity in terms of habitat preferences and some morphological features e.g. moderately large papillae on cheek, anterior nostrils with an elongated flap and overlapping scale counts along the body (Goren 1996). It is therefore difficult to separate the two species on the basis of morphological characteristics and additionally they are inseparable when using genetic characteristics (Fig. 3.1 – 3.7). It is therefore seems like *C. agulhensis* and *C. saldanha* may represent a single species. However, further study with a larger sample size of individuals representing the two species is necessary to confirm this result.

The six species included in the present study comprise three estuarine-dependent species (*C. gilchristi*, *C. nudiceps*, and *C. natalensis*) and three marine species (*C saldanha*, *C. agulhensis* and *C. caffer*). The basal taxon, *C. natalensis*, is represented by an estuarine dependent species without a marine phase. Clade C, *C. gilchristi* and *C. caffer*, consists of an estuarine-dependent species (*C. gilchristi*) with a marine larval phase and marine species (*C. caffer*). Clade B, *C. nudiceps*, even though little is known about the biology and ecology of this species, it is likely that it shares a similar life-history style with *C. gilchristi*, i.e. estuarine spawning followed by a marine larval phase and postlarvae entering nearby estuaries (Whitfield 1998). However, *C. nudiceps* is not as estuarine-dependent as *C. gilchristi* since it

is often recorded in the littoral marine zone (Whitfield 1998). Clade A, *C. saldanha*/*C. agulhensis*, again consists of marine species. Thus, no clear pattern emerged, through the estimated phylogenies, in terms of the occurrence of these species in the different habitat types and the evolution of their preference for estuarine or marine habitats. Intense sampling of the *Caffrogobius* species, and the incorporation of different sequence data in combination with morphological characters, may help to increase our future knowledge and understanding.

The goal of the study was to test Goren's (1996) morphological concept of the species composition of *Caffrogobius* using mtDNA genes (COI and cyt-*b*) and to generate a phylogenetic hypothesis for interrelationships within this genus. Although South African gobiid fishes inhabit a vast range of habitats, from freshwater to estuarine and marine environments, limited attention has been given to the systematics of this group (Goren 1996; Maake 2009). Knowledge about the relationships within gobiid fishes may increase the understanding of the evolution of this group (Harold *et al*. 2008).

Table 3.1: Number of second dorsal fin rays overlap within the *Caffrogobius* species, *C. agulhensis* and *C. saldanha*. N = total number of random specimens counted; D2 = second dorsal fin ray.

	Number of second dorsal fin rays					
D2 count	8	9	10	11	12	N
C. agulhensis	5	6	10	13	2	36
C. saldanha	1	7	25	3	-	36

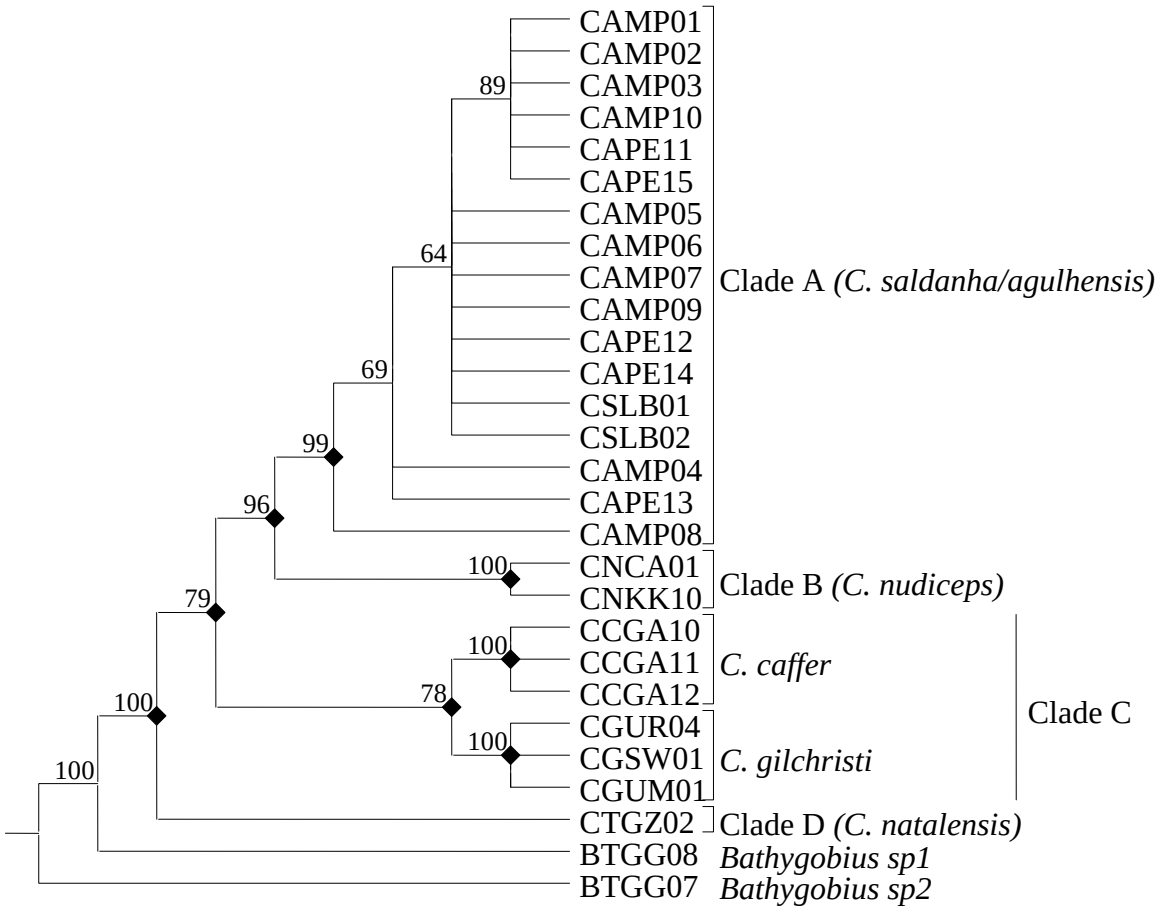

Figure 3.1: Parsimony estimation of phylogenetic relationships among Caffrogobius species using the partial mtDNA COI gene sequences with bootstrap support values *Caffrogobius* species using maximum parsimony with bootstrap support values (64 - 100%) indicated (Specimen annotations according to Table 2.1; Tree length (TL) = 251, Consistency index (CI) = 0.805, Retention index (RI) = 0.901, Rescaled consistency index (RC) = 0.725).

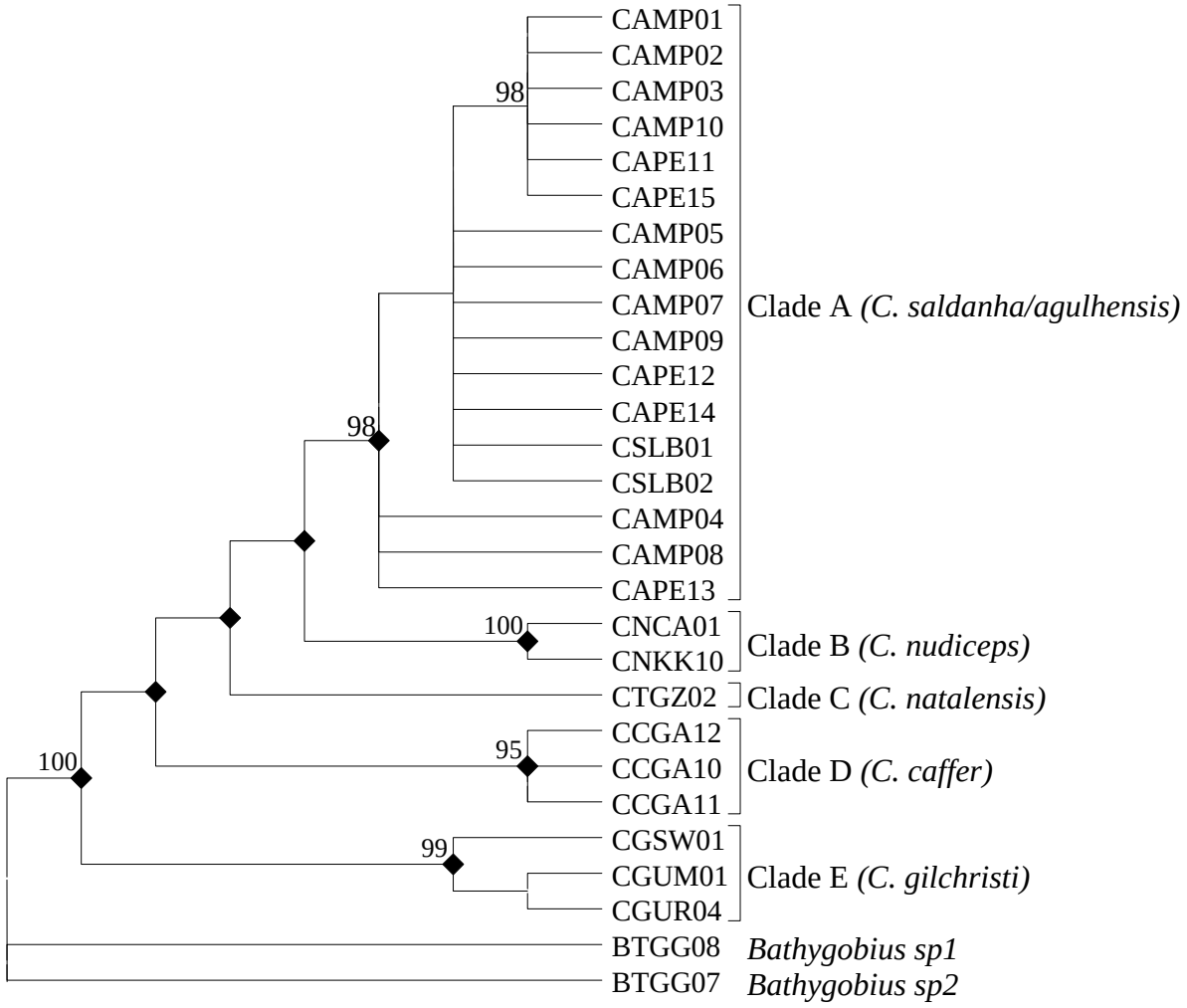

Figure 3.2: Bayesian inference with phylogenetic reconstruction of the relationships among *Caffrogobius* species using mtDNA partial COI gene sequences with posterior probability values (specimen annotations according to Table 2.1).

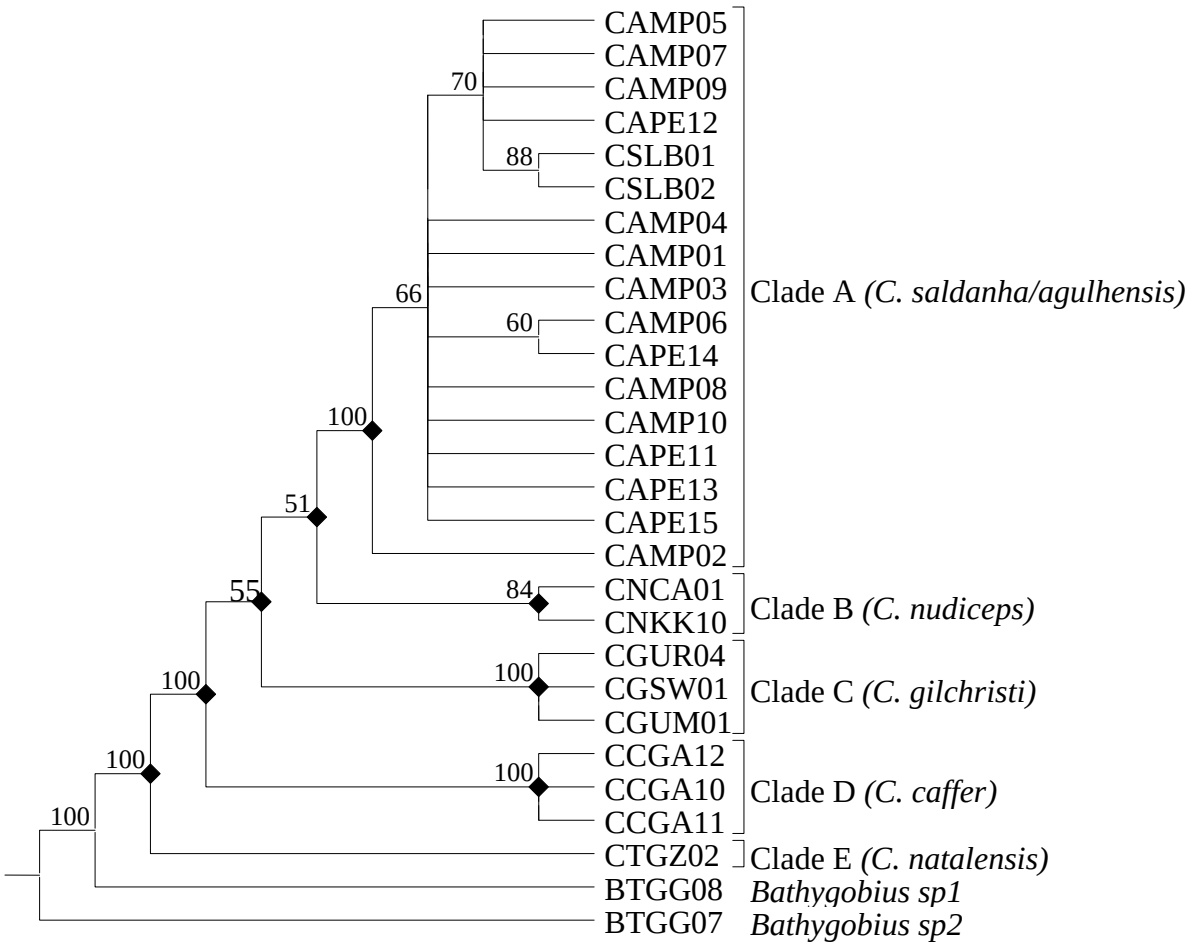

Figure 3.3: Parsimony estimation of phylogenetic reconstruction of *Caffrogobius* using mtDNA cyt-*b* gene sequences with applicable bootstrap values (51 – 100%). Specimen annotations according to Table 2.1, Tree length (TL) = 159, Consistency index (CI) = 0.761, Retention index (RI) = 0.896, Reconsistency index (RC) = 0.0.682).

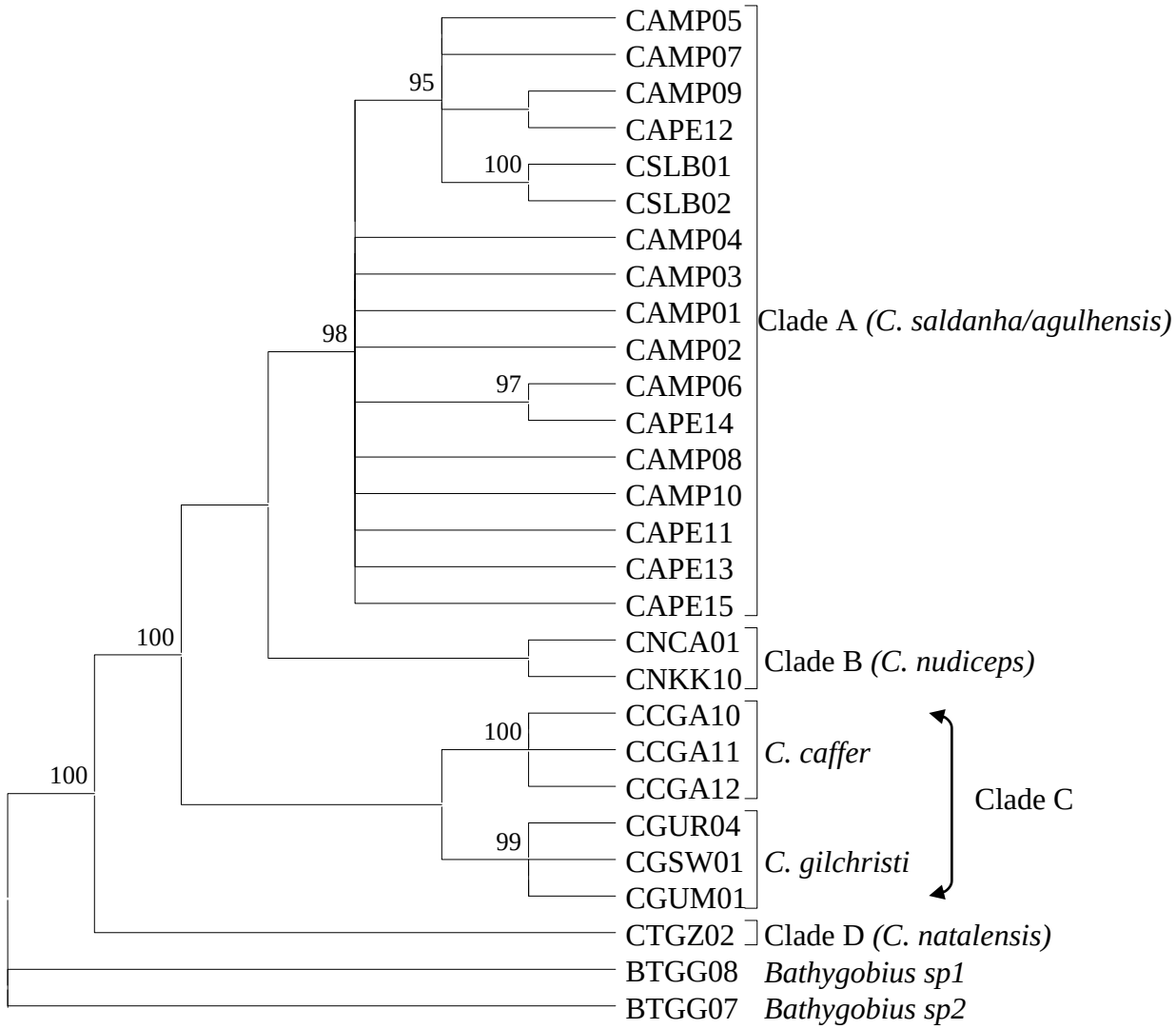

Figure 3.4: Bayesian inference with phylogenetic reconstruction of the relationships among *Caffrogobius* species using mtDNA partial cyt-*b* gene sequences with applicable posterior probability values indivated (specimen annotations according to Table 2.1).

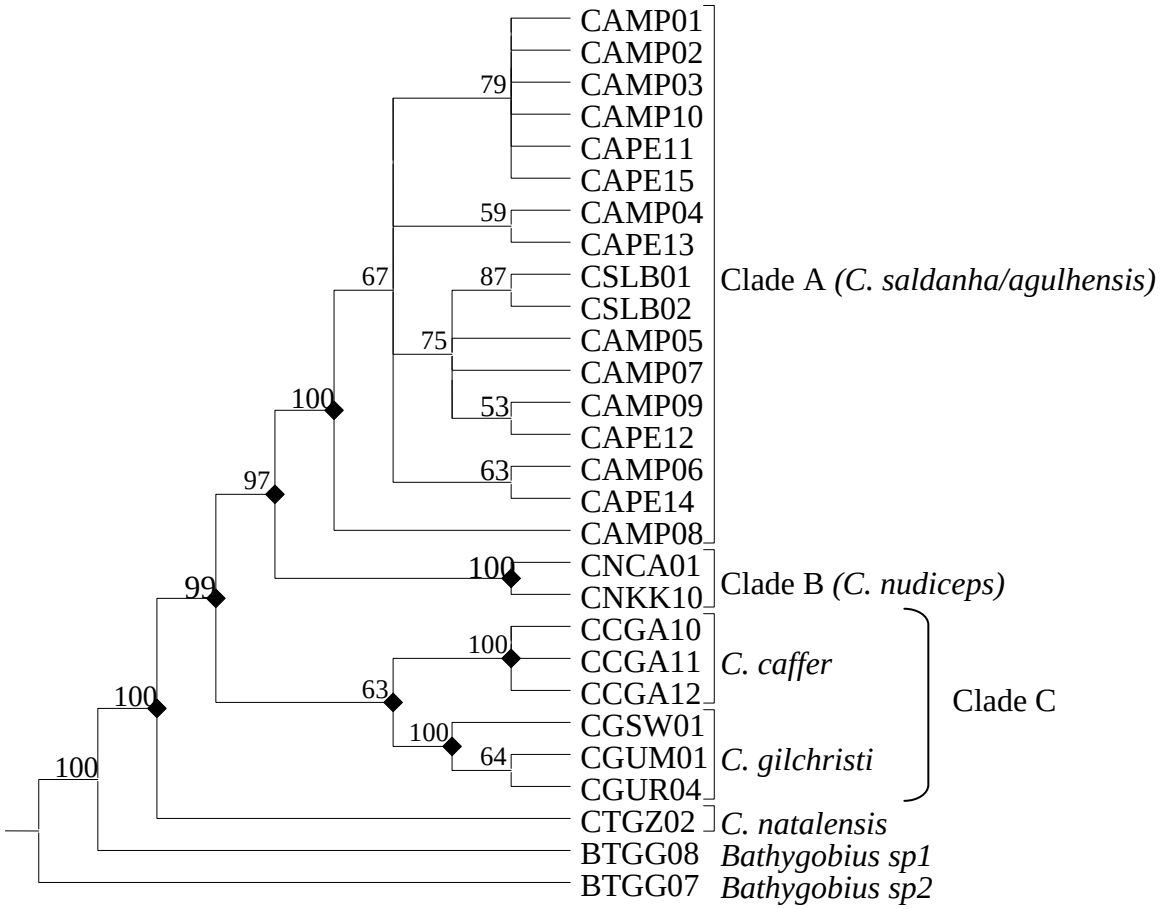

Figure 3.5: 50% majority rule consensus tree of *Caffrogobius* species using the combined mtDNA (COI and cyt-*b*) sequences, with bootstrap values (%) shown (specimen annotations according to Table 2.1; Tree length (TL) = 415, Consistency index (CI) = 0.778, Retention index (RI) = 0.893, Rescaled consistency index (RC) = 0.695).

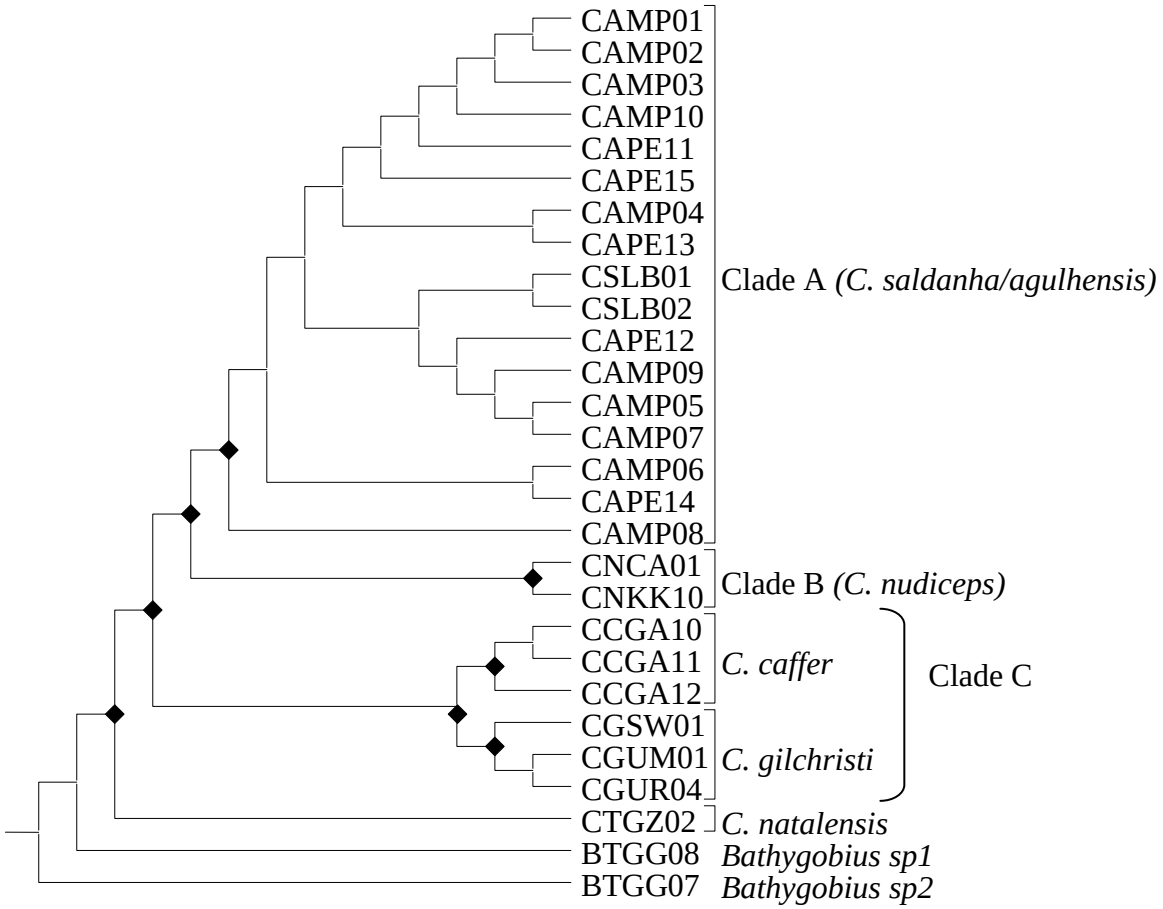

Figure 3.6: Maximum likelihood phylogenetic reconstruction (lnL = 2405.24) of *Caffrogobius* species using mtDNA combined dataset (COI and cyt-*b*) specimen annotations according to Table 2.1.

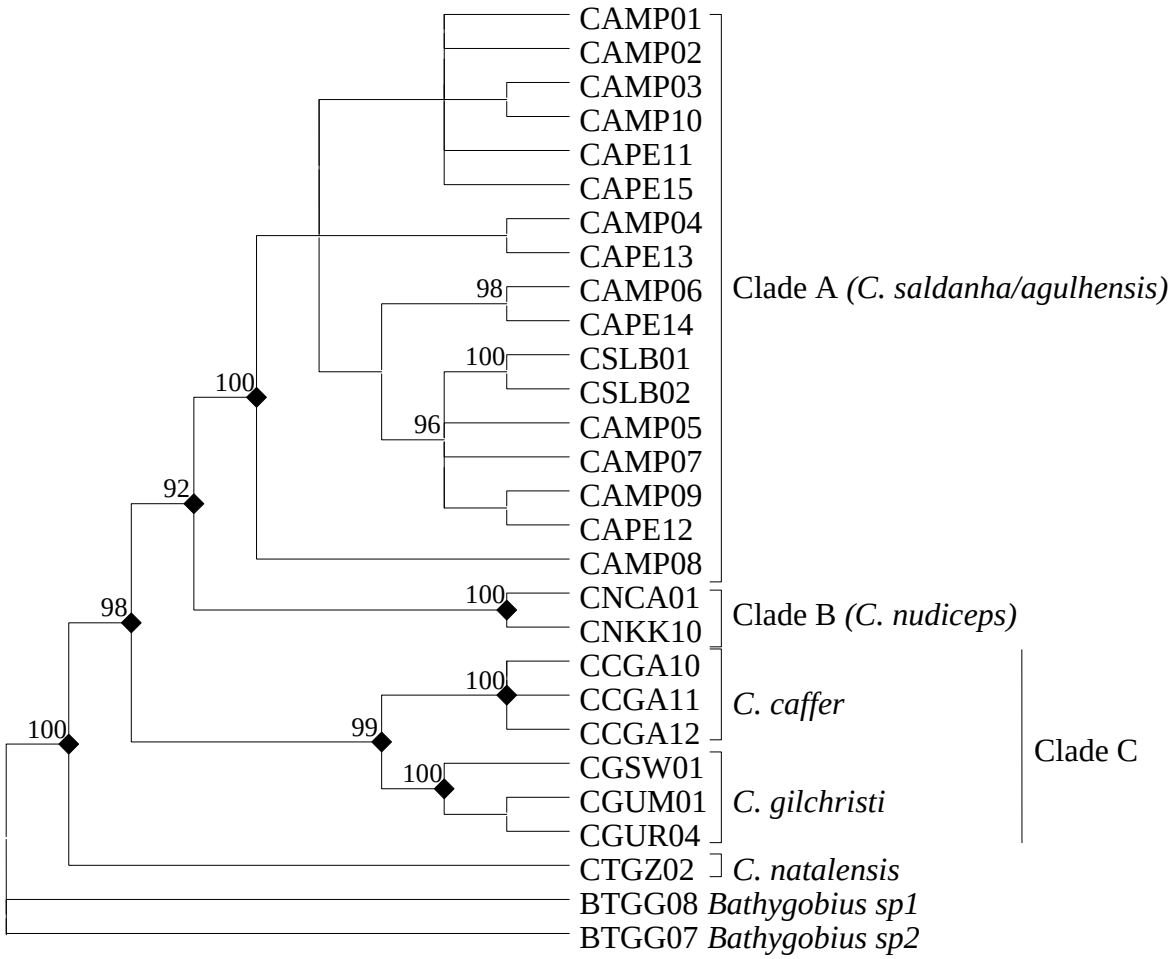

Figure 3.7: Bayesian inference with phylogenetic reconstruction of *Caffrogobius* using partial mtDNA combined dataset (COI and cyt-*b*) sequences, with applicable posterior probability values indicated (specimen annotations according to Table 2.1).

CHAPTER FOUR
INTRASPECIFIC GENETIC VARIATION IN THE PRISON GOBY, *CAFFROGOBIUS GILCHRISTI* (TELEOSTEI, GOBIIDAE) FROM THE EASTERN AND SOUTHERN CAPE ESTUARIES

4.1 Introduction

The prison goby *Caffrogobius gilchristi* Boulenger, 1900 is an estuarine-dependent species endemic to southern African coastal waters from the Olifants System on the Atlantic Coast to Durban Bay in the Indian Ocean (Whitfield 1998). The species is commonly found in tidal pools as well as middle and upper reaches of certain warm-temperate estuaries, especially those that are permanently open to the sea. These habitats are characterised by a high degree of instability and are influenced by terrestrial events, such as rainfall and soil erosion, and marine factors such as tides, wave action and currents (Goren 1996; Whitfield 1998). The prison goby matures at a length of approximately 93.2 mm standard length (Goren 1996) with breeding occurring during the spring and early summer (September to January) (Whitfield 1998). It is well documented that mass hatching of larvae usually coincides with the nocturnal high tide, which influences its marine dispersal capabilities. *Caffrogobius gilchristi* occurs in a salinity range of 0-42%, which allows it to survive in high saline environments as well as the low salinities that occur during flooding of estuaries (Whitfield 1998). According to studies done in the Bot and Swartvlei estuaries, the diet of this goby consists primarily of amphipods, isopods, brachyurans, insect larvae, anomurans, macrurans, polychaetes and small fishes (Whitifeld 1998).

While a number of biological and ecological studies of the prison goby have been undertaken (e.g. Strydom 1998; Whitfield 1998; Strydom & Whitfield 2000; Strydom & Neira 2006), little is known about its population genetic structure in the region. Various factors, such as larval behavior, circulation patterns, and distribution of suitable habitats, affect the genetic structure of aquatic fauna (Grosberg & Cunningham 2001). It is generally accepted that species with pelagic larvae have high dispersal capabilities and are expected to show no genetic structure within open marine environments (Palumbi 1994). Knowledge of population structure within marine species is critical for conservation and/or management purposes. It is well documented that genetic divergence and speciation in marine organisms could result in population structuring if dispersal potential is very limited (Palumbi 1994; Norton 2005;

Teske *et al*. 2007a,c; Neethling *et al.* 2008; Maake 2009). Palumbi (1994) discussed five directional or temporal limits to dispersal that can result in populations becoming isolated over time. These include the following invisible barriers to dispersal (e.g. current patterns, oceanic circulation or tectonic plate effects), diffusion effects in which the density of dispersing propagules, such as larvae, decreases with increasing distance from the source of origin; behavioral characteristics such as philopatry, selection pressures and life history strategies of organisms.

Over the past five years, several research projects that focused on the genetic structure of populations of marine and estuarine organisms along the southern African coast have indicated that there is considerable population genetic structure (Teske *et al.* 2006; Zardi *et al*. 2007). These studies included fishes (Norton 2005; von der Heyden *et al.* 2007; Maake 2009), and other marine organisms such as shrimps (Teske *et al*. 2007a). Such research represents initial attempts to understand intraspecific genetic variation of marine and estuarine organisms and have suggested that there are still questions regarding the biogeography as well as the phylogeography of many marine organisms in southern Africa. Norton (2005) studied the population structure of two estuarine fish species, *Atherina breviceps* Cuvier & Valenciennes, 1835 and *Gilchristella aestuaria* Gilchrist, 1913 along the southern African east and west coasts. *Atherina breviceps* demonstrated substantial gene flow while *G. estuaria* demonstrated a more structured population and considerably less gene flow. Klopper (2005) studied intraspecific genetic variation in two commercially important species, namely *Argyrosomus japonicus* Temminck & Schlegel, 1843 and *Pomadasys commersonni* Lacepéde, 1801. His results indicated high levels of genetic diversity within both species, with no significant population structure detected, indicating that certain widespread species from different regions along the southern African coast are highly connected. Furthermore, Neethling *et al.* (2008) investigated the evidence for panmixia in the southern African endemic *Caffrogobius caffer,* despite barriers to gene flow, there was no population subdivision found. Although results indicated a very limited gene flow amongst the populations, the analyses suggested a recent population divergence of the species. All the above-mentioned studies attempted to test the null hypothesis, i.e. that marine or estuarine organisms have single populations throughout their range along the southern African coastline, i.e. interconnected gene flow. To date, these southern African studies have come up with different results based on different markers. Oosthuizen *et al.* (2004) reported that *Octopus vulgaris* Cuvier, 1797 has a single population around the South African coastline,

and thus the null hypothesis was accepted while Teske *et al.* (2006) and Zardi *et al.* (2007) rejected this hypothesis, indicating considerable phylogeographic divisions in the populations of some marine invertebrates. If the above-mentioned null hypothesis (single population) is left untested, this may lead to incorrect interpretations of population structures of many marine and estuarine species.

According to Teske *et al.* (2006), there are some possible scenarios that could explain the lack of genetic diversity in marine/estuarine species: (a) southern African populations have undergone recent genetic population expansion (b) mitochondrial DNA evolves more slowly in fishes than in other marine organisms, such as invertebrates; (c) the low amount of genetic diversity is a sampling artifact; (d) there has been a contamination problem during amplification, and (e) to a large extent, the degree of genetic structure, or the absence of structure, in marine organisms depends on the mode of dispersal of the larval or adult stages of the organism, for example the larval stages of the *C. caffer* (Neethling *et al.* 2008). The population structure of the estuarine-dependant *C. gilchristi* will therefore be explored using sequences of two mitochondrial DNA regions (COI and cyt-*b*) in populations along the eastern and southern Cape coastline of South Africa.

4.2 Material and methods

Details of the material and methods used in this study are given in Chapter Two.

4.3 Results

4.3.1 Cytochrome Oxidase Subunit I

A 524 bp fragment of the COI gene was amplified from 109 *C. gilchristi* individuals collected from 12 estuaries (Table 2.2, Fig. 2.1). Of the 32 (6%) polymorphic sites, six (19%) were parsimony informative while eight (25%) were singleton variable sites. These polymorphic sites defined 14 unique haplotypes amongst the individuals that were analysed (Table 4.1). The total haplotype (gene) diversity (*h*) was 0.673 with a nucleotide diversity (π) of about 0.0019. The most common haplotype (H1) was shared by 56 individuals, with most of these originating from the Kowie (KO), Gamtoos (GM), Keurbooms (KB), and Swartvlei (SW) estuaries. Haplotype H2 was represented by 25 individuals, with most individuals originating from Gordon's Bay (GB), Gamtoos (GM) and Kowie (KO) estuaries. Haplotype

H4 was found in 13 individuals, most of which originated from the Gamtoos (GM) and Swartvlei (SW) estuaries; haplotype H3 was shared by four individuals and haplotype H8 was shared by two individuals, while the rest of the haplotypes were rare or unique to a particular individual (Table 4.2). The minimum spanning network (Figure 4.1) indicated that there was no genetic population structure and no relationship between genealogy and geographic location. There were only a four haplotypes that were connected by more than one mutational step namely, H2 to H8 and H2 to H12. There was no apparent geographic pattern within the minimum spanning network.

Neutrality tests of population differentiation amongst all locations were performed. The value for Tajima's D equivalent was -1.701, which is not significant ($0.10 > P > 0.05$), suggesting a recent population expansion in these areas. Other neutrality tests, such as the Fu and Li's D statistic (-3.065) and Fu's Fs statistic (-9.003), were statistically significant ($P < 0.05$). According to Fu (1997), population demography analysis that indicates only the Fu and Li's tests as significant, suggests population growth as the likely cause of any recent population expansion. The mismatch distribution (Figure 4.2) was unimodal, also supporting a model of a recent population expansion. The calculated raggedness index *r* was 0.095 (Figure 4.2), which represents a relatively low indication of population expansion. Thus all neutrality tests support recent population expansion and growth in *C gilchristi*. The more powerful The Ramos-Onsins and Rozas R_2 value for small population sizes ($R_2 = 0.0366$) is however not small enough to indicate that the populations have undergone historical growth or that the marker is under selection (Ramos-Onsins & Rozas 2002).

No correlation ($r^2 = 0.1165$) exists between the fixation index obtained for different individuals and the geographic distances among localities (Figure 4.3). This maybe an indication of gene flow among subpopulations.

A slow-evolving mutation rate of 3.6% per million years (generally used for teleosts) and the fast-evolving rate of 11-13% (used for white sturgeon) (Neethling *et al*. 2008) were selected for this study for the COI and cyt-*b* markers. The population expansion time of divergence was estimated between 19 400 (11% mutation rate) and 59 400 (3.6% mutation rate) for a 1-year generation time and between 6 500 (11% mutation rate) and 19 800 (3.6% mutation rate) for a 3-year generation time.

The estimated components of genetic structure and gene flow suggested a high level of gene flow among localities. An analysis of molecular variance (AMOVA) (Excoffier *et al*. 2006) was performed to assess the genetic structure and differentiation among *C. gilchristi* locations, according to the assigned groupings in Table 4.3. Hierarchical AMOVA revealed that regional grouping of estuaries into three, four, and even five groups yielded no consistent patterns of differences in genetic structure within populations (Table 4.3). The results obtained through spatial analysis of molecular variance (SAMOVA) gave less resolution in terms of groupings than those obtained through AMOVA and are therefore not presented.

4.3.2 Cytochrome-*b*

A 425 bp fragment of the cytochrome-*b* region was obtained from 63 *C. gilchristi* individuals collected from eight estuaries (Table 2.2). Of these, 406 (96%) sites were invariable (monomorphic) with 19 (4%) sites being variable (polymorphic) and made up of 12 (63%) singleton variable sites and seven (37%) parsimony informative sites. Haplotype (*h*) and nucleotide diversities (π) were 0.743 and 0.003 respectively (Table 4.4). The estimated cyt-*b* haplotype diversity values within each locality (with the exception of the Kowie (KO) and Gamtoos (GM) estuaries) were relatively high for the species (Table 4.4). The most common haplotype (H2) was shared among 31 individuals, from five localities i.e. the Kowie (KO) (10), Keurbooms (KB) (8), Gamtoos (GM) (6), Swartvlei (SW) (6) and Keiskamma (KK) (1) estuaries (Table 4.5). Haplotype H5 was found in six individuals and was also shared amongst four localities, namely the Kowie (KO) (3), Keurbooms (KB) (2) and Umzumvubu (UZ) (1) estuaries. Haplotype H6 was found in six individuals: from the Umhlatuze (UM) (2), Swartvlei (SW) (1), Kromme (KR) (1), Keiskamma (KK) (1) and Keurbooms (KB) (1) estuaries (Table 4.5). No genetic structure was observed among haplotypes from these localities.

A minimum spanning tree was constructed connecting all haplotypes on the basis of the least number of substitutions (Figure 4.4). The resulting network does not suggest sub-structuring, with most haplotypes being connected by a single mutation step, except for H12 and H14 that are connected by two mutational steps (Figure 4.4). Further analysis to test for the possibility of isolation by distance (IBD) was conducted. The mismatch distribution was unimodal, and a sign for a sudden population expansion (Figure 4.5). The calculated raggedness index, r = 0.068, was very low, also suggesting sudden population expansion. The evidence of recent population expansion was further supported by the following test results: neutrality test of

Fu's Fs (-19.271) and Fu's and Li's F statistic (-3.376), with statistical significance at $P < 0.02$; Fu and Li's D statistic (-3.252) with statistical significance at $P < 0.05$ and Tajima's D statistic (-2.092) with statistical significance ($P < 0.05$).

No correlation ($r^2 = 0.234$) exists between the fixation index obtained for different individuals and the geographic distances among localities (Figure 4.6) which maybe an indication of gene flow among subpopulations.

AMOVA estimated that a significant percentage of sequence variation existed within population (83.18%), followed by that among populations within groups (7.38%) (Table 4.3). Similar to the results of COI spatial analysis of molecular variance (SAMOVA) gave less resolution in terms of groupings than those of AMOVA and are therefore not presented. An estimate of the time since population expansion (t) was calculated, based on results from previous studies on correlating mutation in DNA sequence and time among fishes (Craig *et al*. 2006; Zhang *et al*. 2006; Neethling *et al.* 2008). Craig *et al*. (2006) obtained mutation rate estimates for cyt-*b* for some reef fishes of around 1–3% Myr^{-1}, which was similar to the estimated mtDNA clock of 2% Myr^{-1} proposed by (Kocher & Stepien 1997). Thus the present study, following Craig *et al*. (2006), estimates the range of mutation rates for cyt-*b* to be around 1–3% Myr^{-1}. By using 1-year generation time (according to Neethling *et al.* 2008), the present study estimates the time of population expansion to be between 136 000 (11% mutation rate) and 41 700 (3.6% mutation rate) years ago for *C. gilchrist,* while this increased to between 59 400 (11% mutation rate) and 139 000 (3.6% mutation rate) years ago, when using the three-year generation time.

4.4 Discussion and conclusion

This is the second in-depth study to determine the genetic variation within species of the genus *Caffrogobius,* following that of Neethling *et al*. (2008) on the species *Caffrogobius caffer.* As stated before, it is generally expected and accepted that very little intraspecific population structure can be expected in large populations with a continuous distribution, covering large geographic areas in marine and estuarine environments and that has a high level of adult dispersal and larval life (Avise *et al*. 1987; Palumbi 1992). Similar to the findings by Neethling *et al*. (2008) and contradictory to the results by Maake (2009), the mitochondrial DNA analyses using cyt-*b* and COI supported the null hypothesis that the *C.*

gilchristi has a single population throughout its range. Therefore, establishing ESUs seem to be unnecessary but it is important to do a thorough examination of their morphological characterics, their ecological distribution as well as using nuclear genes (Crandall *et al*. 2000).

For the COI analysis, the results indicated that there is a high level of haplotype diversity (0.673) and a low level of nucleotide diversity (0.0019), suggesting rapid population growth from an ancestral population with a small effective population size (Avise 2000). The highest haplotype diversities were recorded from the Gonubie (GO) and Keiskamma (KK) estuaries, whilst the lowest were from the Kromme (KR), Kowie (KO) and Groot Braak (GR) estuaries, implying a high number and frequency of different alleles at a locus at Gonubie and Keiskamma estuaries than at the others. The highest nucleotide diversity was estimated at the Keiskamma (KK) and Keurbooms (KB) sites, while the lowest was recorded at the Kromme (KR), Grootbraak (GR) and Umhlatuze (UM) sites (Table 4.2), implying higher sequence divergence among individuals at Keiskamma (KK) and no sequence divergence among individuals at Kromme estuaries. In the haplotype network (Figure 4.1), the torso is comprised of four haplotypes (H1, H2, H7, and H10) and is shared amongst all localities, with different percentage contributions.

Results indicated that cyt-*b* sequences generated for *C. gilchristi* specimens have a high level of haplotype diversity (h = 0.743) and a low level of nucleotide diversity (π = 0.003), again suggesting rapid population growth from an ancestral population with a small effective population size (Avise 2000). The combination of high haplotype diversity and low nucleotide diversity is a typical signature of population expansion after a bottleneck or a founder effect, which is influenced by a lack of variability, or little variability, in population structure. These highest levels of haplotype diversity were noted at the Umzumvubu (UZ), Keiskamma (KK), and Kromme (KR) estuaries, and the lowest at the Kowie (KO) and Gamtoos (GM) estuaries. The highest nucleotide diversity was noted at KR and KK estuaries, and the lowest at the Umzumvubu (UM) and Gamtoos (GM) estuaries (Table 4.4). These levels of genetic diversity are comparable to those of *C. caffer* (Neethling *et al*. 2008) which were 0.935-0.982 haplotype diversity and 0.008-0.011 nucleotide diversity. In the haplotype network (Figure 4.4), the torso is comprised of four haplotypes (H2, H5, H6, and H20) and is shared amongst all localities, with different percentage contributions. In *C. gilchristi*, the unimodal mismatch distribution and significant Fu's F_S (-19.271) both suggest that the

species has undergone population expansion similar to *C. caffer* (see Neethling *et al*. 2008). For the cyt-*b* gene, the population expansion time divergence has been estimated to be 41 700 (11% mutation rate) and 136 000 (3.6% mutation rate) years ago, considering a 1-year generation time, as the species is recorded as spawning once per year. Estimating time divergence at a 3-year generation time suggests a population expansion time of 59 400 (11% mutation rate) and 139 000 (3.6% mutation rate) years ago. The analyses suggested that the *C. gilchristi* lack genetic structure as compared to other studies on closely related species e.g. *Glossogobius callidus* Smith, 1937 (see Maake 2009) and that the total population expanded during the Pleistocene period (275 000 to 15 000 years ago depending on the mutation rate and generation time). This study supports the notion of a single population structure in this estuarine-dependent species.

4.4.1 Historical demography, lack of population structure and fishery management

The overall findings, according to the high haplotype diversity and the low nucleotide diversity from both genes, suggest a bottleneck followed by population expansion or growth (Avise 2000). This is well supported by the significant Fu's F_s ($P < 0.02$) and Tajima's D tests. The estimated gene flow (from haplotype data from COI and cyt-*b*) resulted in a high number of migrants (2.89 and 2.81 respectively). Previous geological and climatic events have played a major role in estuarine and marine biogeography in southern Africa. Low sea levels were associated with Pleistocene glaciations (2.4 – 10 000 Mya) that resulted in the formation of the complex geological landscape of the southern Cape coastline (Lutjehams 2005). This geological landscape does not, however seem to limit gene flow amongst estuarine and marine species.

According to Zhang *et al*. (2006), larval dispersal in marine species has long been assumed to be widespread due to the lack of obvious geographical barriers to migration in oceanic environments, which leads to no genetic differentiation amongst populations. This is the case for both the present study and the study on *C. caffer* (Neethling *et al*. 2008) as well as some other estuarine and marine species (Klopper 2005; Norton 2005). This was not, however, the case for many studies on South African invertebrates, in which evidence of phylogeographic structure was found (Teske *et al*. 2007b; Zardi *et al*. 2007).

A recent estuarine study of population genetics on the endangered Knysna seahorse, *Hippocampus capensis*, implied that the one-population hypothesis of homogenous genetic

structure was not applicable to all fish species (Teske *et al*. 2003). These authors demonstrated the value of population genetic research for assessing the conservation status of populations as a basis for planning marine protected areas. Various factors affect the structure of aquatic animal populations, including larval behavior, circulation patterns, distribution of suitable habitats, and the geological and climatic history of these habitats (Grosberg & Cunningham 2001). Although general aspects of the reproductive biology of the estuarine-dependent *Caffrogobius gilchristi* are known (Whitfield 1998), there is little or no data on the behavior of their larvae in the sea that would explain the finding of a single population encompassing their entire geographical range.

The population of *C. gilchristi* is distributed across South Africa's three coastal biogeographic zones, the cool temperate west coast, the warm temperate south coast and the subtropical east coast (Goren 1996; Whitfield 1998). In this study, the eastern and southern Cape areas are of particular interest, being intermediate (or transitional) between the subtropical and the cool temperate zones. The present research covered estuaries from the cool temperate Gordon's Bay in the west to the subtropical Umhlatuze Estuary in the east. The oceanography of the region between Port Elizabeth and Mossel Bay is extremely complex due to the interaction between wind-driven currents and the Agulhas Current (Lutjehams 2005). The resulting conditions may interfere with the movement of larvae into, or out of, affected areas within this region. However, the fact that no population structure was observed implies no interference in movement of larvae due to abiotic conditions.

Are there any reasons to believe that *C. gilchristi* populations along South Africa's coast are homogeneous, or whether geology and historical demography have had any genetic impacts on these populations? According to the present analyses, genetic diversity is mostly uniform throughout its distribution along the southern African coast. It is highly unlikely that adult *C. gilchristi* can disperse for long distances as they are estuarine-dependant; therefore the most possible dispersal could be via their pelagic marine larval phase. The results of this study therefore indicate that dispersal against apparent oceanographic barriers is possible for some estuarine species.

The current results indicate that, for the purpose of conservation management, *C. gilchristi* along the southern African coast can be treated as a single population. It is, however, also recommended that more detailed studies be undertaken on this species, incorporating

sampling that covers a larger extent of the distribution range. Future studies should also make use of a higher number of genes. Since genes have different mutation rates, studying a greater variety of molecular characters could help to reveal the true nature of intraspecific genetic variation in this species. Due to the high rate of coastal development, it is imperative that a greater level of support is given to research covering the following aspects: population genetics research into estuarine-dependent species; careful monitoring of populations in South Africa's estuarine environment; and assessments of data obtained from periodic physical and genetic surveys. Knowledge generated from such research is crucial for the conservation and management of marine and estuarine biodiversity in South Africa. The present study highlights the importance of understanding the mobility of either the adults (males or females) or larvae, which is fundamental for understanding population structure as well as life-history characteristics (Neethling *et al*. 2008).

Table 4.1: Haplotype frequencies from 12 locations where *Caffrogobius gilchristi* were collected, inferred from the mtDNA gene COI. (H = haplotype, EK = Kleinemonde, GO= Gonubie, GM = Gamtoos, GB = Gordon's bay, GR = Groot Braak, KB = Keurbooms, KK = Keiskamma, KO = Kowie, KR = Kromme, SW = Swartvlei, UM = Umhlatuze and UZ = Umzumvubu).

H	EK	GO	GM	GB	GR	KB	KK	KO	KR	SW	UM	UZ	Total
H_1	2	-	12	3	3	9	1	18	2	5	-	1	56
H_2	3	1	4	6	-	1	1	4	-	1	2	2	25
H_3	1	1	1	-	-	-	-	-	-	-	1	-	4
H_4	-	-	4	1	1	2	1	-	-	4	-	-	13
H_5	-	-	1	-	-	-	-	-	-	-	-	-	1
H_6	-	-	-	1	-	-	-	-	-	-	-	-	1
H_7	-	-	-	-	-	1	-	-	-	-	-	-	1
H_8	-	-	-	-	-	1	1	-	-	-	-	-	2
H_9	-	-	-	-	-	-	1	-	-	-	-	-	1
H_10	-	-	-	-	-	-	1	-	-	-	-	-	1
H_11	-	-	-	-	-	-	-	1	-	-	-	-	1
H_12	-	-	-	-	-	-	-	1	-	-	-	-	1
H_13	-	-	-	-	-	-	-	-	-	1	-	-	1
H_14	-	-	-	-	-	-	-	-	-	-	-	1	1

Table 4.2: Sample size and descriptive genetic variability estimates based on mtDNA COI sequences from the 12 locations where *Caffrogobius gilchristi* specimens were collected (N = number of individuals/sample size; HL = haplotypes per locality; π = nucleotide diversity; *h* = haplotype (gene) diversity. Localities are: EK = Kleinemonde, GO= Gonubie, GM = Gamtoos, GB = Gordon's bay, GR = Groot Braak, KB = Keurbooms, KK = Keiskamma, KO = Kowie, KR = Kromme, SW = Swartvlei, UM = Umhlatuze and UZ = Umzumvubu).

Population	**N**	**HL**	**HL/N**	**Polymorphic sites**	**Nucleotide diversity (π)**	**Haplotype diversity (*h*)**
UM	3	2	0.67	1	0.00127	0.667
UZ	4	3	0.75	2	0.00191	0.833
GO	2	2	1.00	1	0.00191	1.000
KK	6	6	1.00	6	0.00433	1.000
EK	6	3	0.50	2	0.00165	0.733
KO	24	4	0.17	4	0.00113	0.423
GM	22	14	0.23	4	0.00164	0.664
KR	2	1	0.50	0	0.00000	0.000
KB	14	5	0.36	5	0.00201	0.593
SW	11	4	0.36	3	0.00167	0.709
GR	4	2	0.50	1	0.00095	0.500
GB	11	1	0.36	3	0.00167	0.672
TOTAL	109	47	0.13	32	0.00188	0.673

Table 4.3: *Caffrogobius gilchristi* results for the *a priori* population structures defined in AMOVA (Excoffier *et al*. 1992) using the program ARLEQUIN v2.0 (Excoffier *et al*. 2006) from the mtDNA dataset (cyt-*b* and COI). COI groupings consisted of the following groups i.e. 3 groups (UM, UZ); (GO; KK, KO, EK, GM, KR); and (SW, KB, GB, GR), 4 groups (UM, UZ); (KB, SW, GR); (GO, EK, KO, GM); and (GB) and 5 groups (UM, UZ); (GO, KK, KO, EK); (KR, GM); (KB, SW); and (GB,GR). Cyt-*b* groupings consisted of the following groups i.e. 3 groups (UM, UZ); (KK, KO, KR, GM); (SW, KB); and 4 groups (UM, UZ); (KK, KO); (GM, KR); and (SW, KB). (Localities: EK = Kleinemonde, GO= Gonubie, GM = Gamtoos, GB = Gordon's bay, GR = Groot Braak, KB = Keurbooms, KK = Keiskamma, KO = Kowie, KR = Kromme, SW = Swartvlei, UM = Umhlatuze and UZ = Umzumvubu).

	Percentage of variation (COI and cyt-*b*)		
Source of variation	3 groups	4 groups	5 groups
	COI / cyt-*b*	COI / cyt-*b*	COI / 4cyt-*b*
Among groups	3.82 / 9.44	8.48 / 5.37	0.54 / -
Among population within groups	9.48 / 7.38	6.22 / 9.84	11.56 / -
Within populations	86.69 / 83.18	85.30 / 84.79	87.90 / -

Table 4.4: Sample size, polymorphic sites and genetic variability estimates based on mtDNA cyt-*b* sequences from eight populations where *Caffrogobius gilchristi* individuals were collected (N = number of individuals/sample size; HL = haplotype per locality; π = nucleotide diversity; h = haplotype diversity).

Population	**N**	**HL**	**HL/N**	**Polymorphic sites**	**Nulceotide diversity (π)**	**Haplotype diversity (h)**
Umhlatuze	3	2	0.67	1	0.00157	0.667
Umzumvumbu	4	4	1.00	4	0.00472	1.000
Keiskamma	6	6	1.00	7	0.00550	1.000
Kowie	15	4	0.27	4	0.00193	0.543
Gamtoos	8	3	0.38	2	0.00118	0.464
Kromme	2	2	1.00	5	0.01179	1.000
Keurbooms	13	5	0.38	4	0.00175	0.628
Swartvlei	12	7	0.58	2	0.00293	0.773
TOTAL	63	31	0.32	29	0.00301	0.743

Table 4.5: The haplotype frequency of the partial mitochondrial DNA gene cyt-*b* of *Caffrogobius gilchristi* from eight different localities. (Localities are: GM = Gamtoos, KB = Keurbooms, KK = Keiskamma, KO = Kowie, KR = Kromme, SW = Swartvlei, UM = Umhlatuze and UZ = Umzumvubu).

Haplotypes	**GM**	**KB**	**KK**	**KO**	**KR**	**SW**	**UM**	**UZ**	**Total**
Hap_1	1	-	-	-	-	-	-	-	1
Hap_2	6	8	1	10	-	6	-	-	31
Hap_3	1	-	-	-	-	-	-	-	1
Hap_4	-	1	-	-	-	-	-	-	1
Hap_5	-	2	-	3	-	-	-	1	6
Hap_6	-	1	1	-	1	1	2	-	6
Hap_7	-	1	-	-	-	1	-	-	2
Hap_8	-	-	1	-	-	1	-	-	2
Hap_9	-	-	1	-	-	-	-	1	2
Hap_10	-	-	1	-	-	-	-	-	1
Hap_11	-	-	1	-	-	-	-	-	1
Hap_12	-	-	-	1	-	-	-	-	1
Hap_13	-	-	-	1	-	-	-	-	1
Hap_14	-	-	-	-	1	-	-	-	1
Hap_15	-	-	-	-	-	1	-	-	1
Hap_16	-	-	-	-	-	1	-	-	1
Hap_17	-	-	-	-	-	1	-	-	1
Hap_18	-	-	-	-	-	-	1	-	1
Hap_19	-	-	-	-	-	-	-	1	1
Hap_20	-	-	-	-	-	-	-	1	1

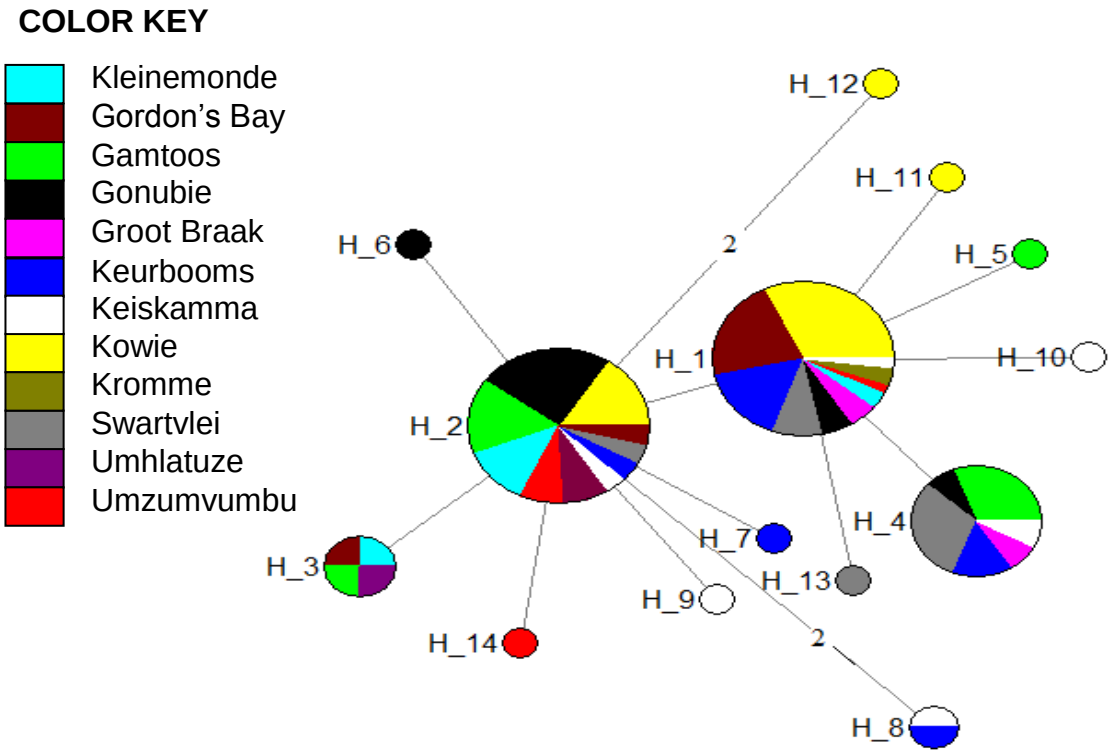

Figure 4.1: The haplotype minimum spanning network constructed with the statistical parsimony method for the partial mtDNA Cytochrome Oxidase Subunit I (COI) gene of *Caffrogobius gilchristi*. Each numerical number '2' represents two mutational changes that interconnect two haplotypes, while other lines are just one mutational step. The size of each circle is proportional to the frequency of each haplotype. Localities that share haplotypes are shaded, based on the corresponding colour key.

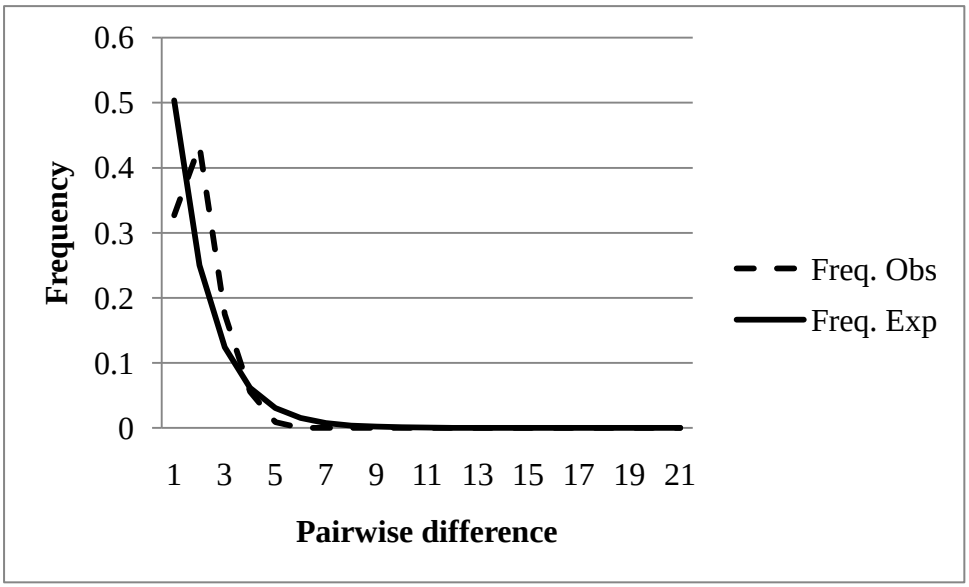

Figure 4.2: The unimodal mismatch distribution of the Cytochrome oxidase subunit I (COI) partial sequences (r = 0.095).

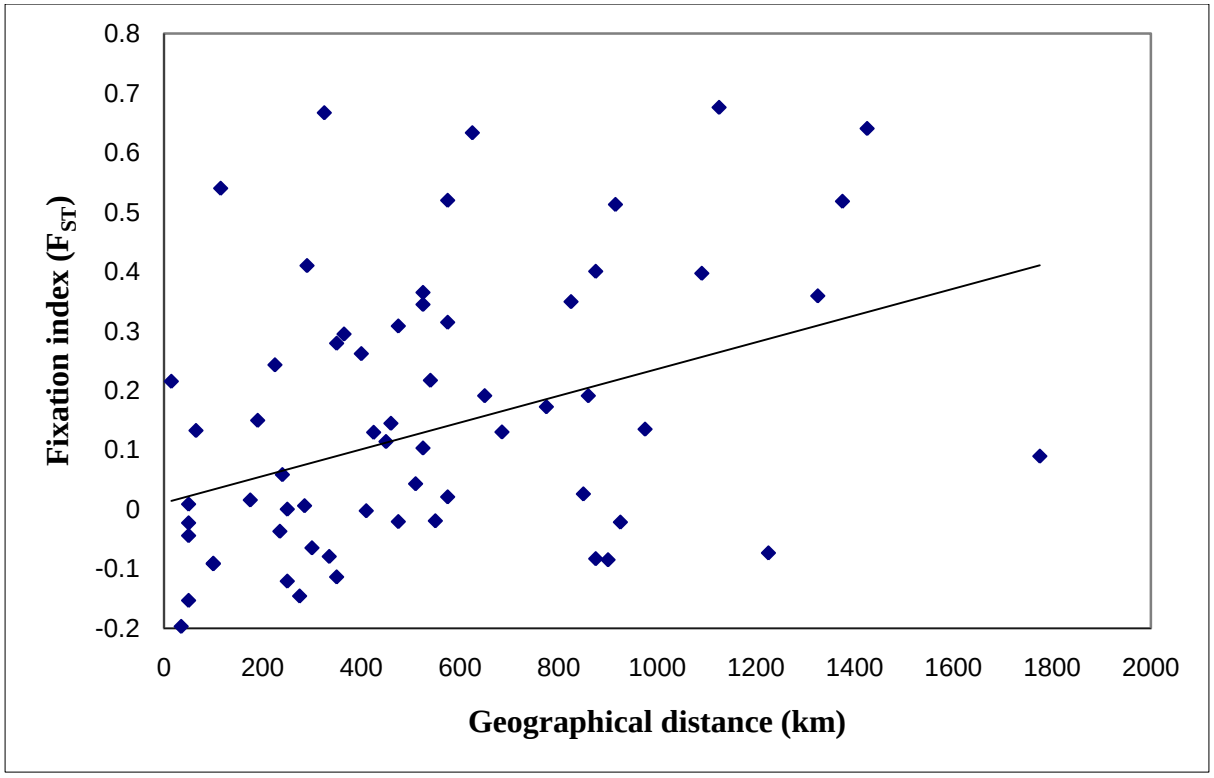

Figure 4.3: Fixation index (F_{ST}) inferred from COI sequences for *Caffrogobius gilchristi* individuals versus geographical distance among the 12 different localities within the eastern and southern Cape region (r^2 = 0.1165).

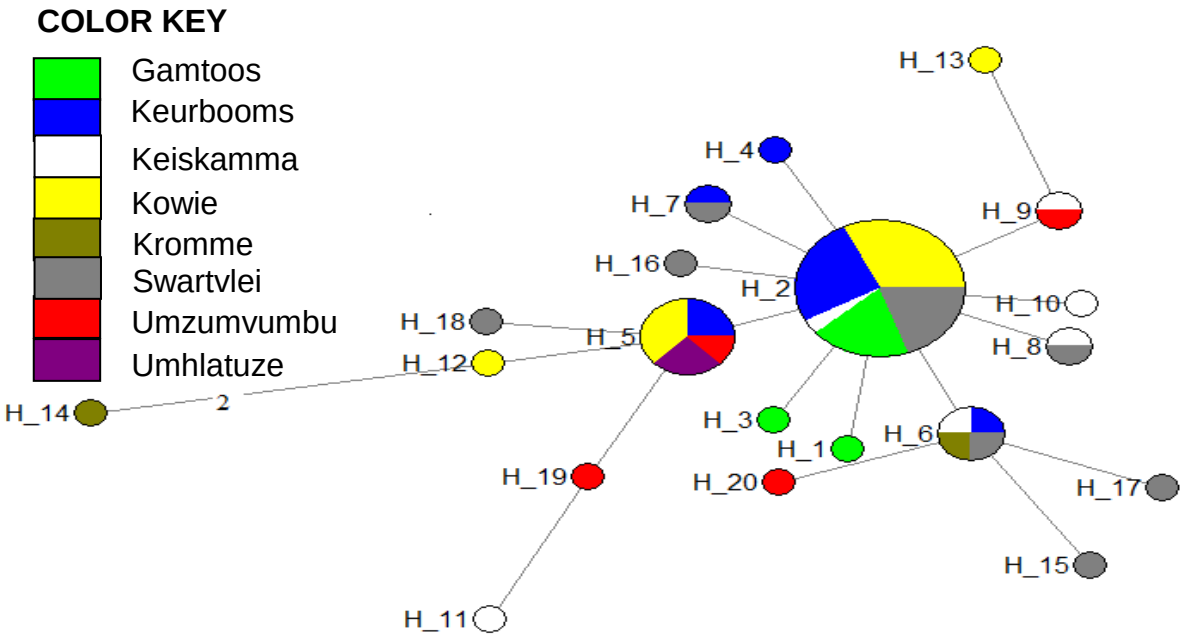

Figure 4.4: Haplotype network for *Caffrogobius gilchristi* individuals obtained from partial cyt-*b* mtDNA gene sequences. The size of circles is proportional to the frequency of each haplotype; the number '2' between H12 and H14 represents two mutational steps. Localities are indicated by the color key.

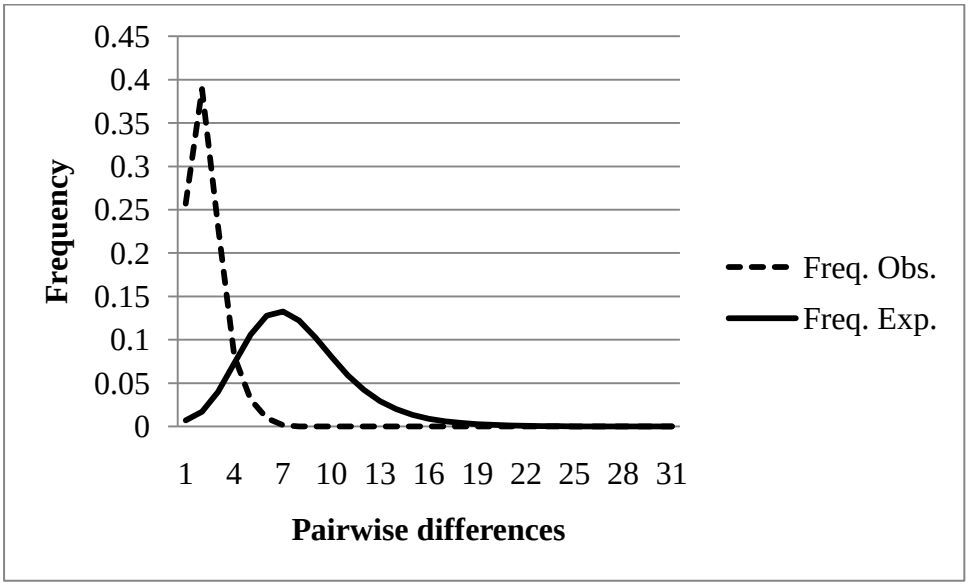

Figure 4.5: The unimodal mismatch distribution of the *C. gilchristi* cytochrome-*b* (cyt-*b*) partial sequences (r = 0.068).

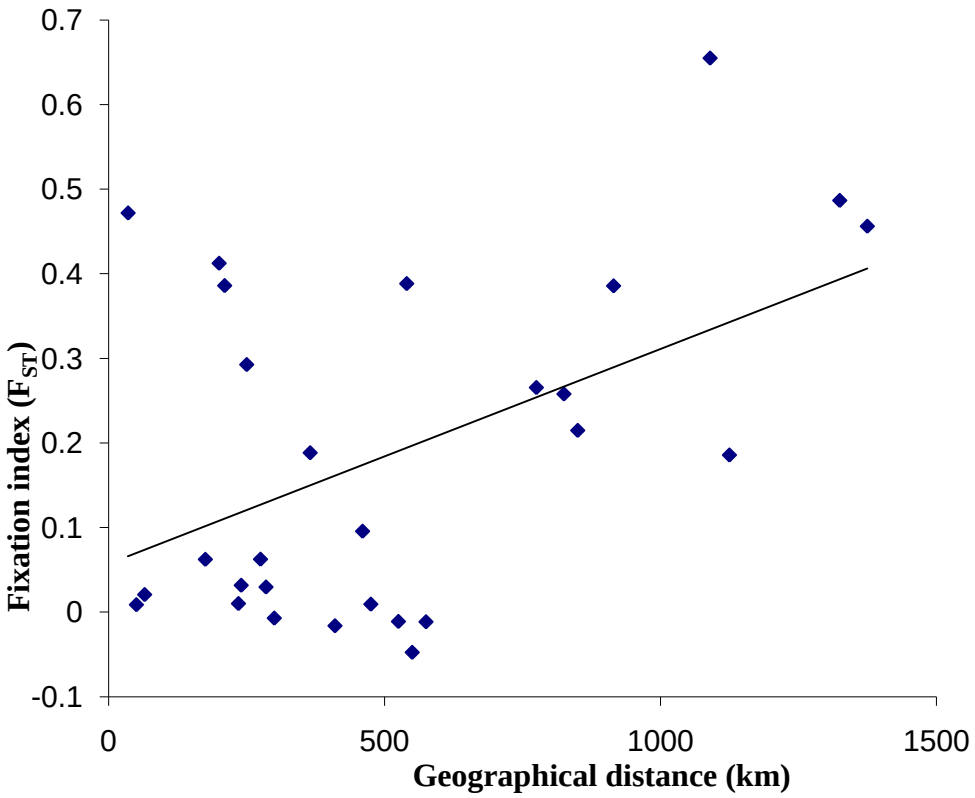

Figure 4.6: Fixation index (F_{ST}) inferred from cyt-*b* sequences for *Caffrogobius gilchristi* individuals versus geographical distance among the eight different localities within the eastern and southern Cape region ($r^2 = 0.234$).

CHAPTER FIVE
GENERAL DISCUSSION AND CONCLUSION

This study estimated the molecular phylogeny of the gobioid genus *Caffrogobius* and intraspecific genetic variation of one of its estuarine-dependent species, *Caffrogobius gilchristi*.

5.1 Molecular phylogeny of the genus *Caffrogobius*

The results of the phylogenetic analyses of sequence data obtained from two partial mitochondrial genes suggested that *Caffrogobius* is a monophyletic group (Figure 3.7) composed of three well supported clades (A–C). Clade A is represented by *C. saldanha* specimens and *C. agulhensis* specimens. Both species are marine inhabitants. These two species of clade A share similar morphological characters and are very difficult to distinguish. Since they make up the same clade with no distinction among the two in any of the estimated phylogenies it is suggested that they may be the same species. More elaborate sampling of these two species may provide better evidence of their species status. Clade B is comprised of *C. nudiceps* specimens and appear to be the sister group of Clade A. Clade C is composed of sister groups represented by *C. caffer*, a marine species, and *C. gilchristi*, an estuarine-dependent species with a marine larval phase. Future studies with higher sampling numbers may shed light on this sister grouping from different habitats and apparent different life histories. The basal taxon is represented by *C. natalensis*, an estuarine-dependent species (Figure 3.7).

Major problems, recognised from Maugé (1986) and Goren (1996), relate to the characters used for the identification of *Caffrogobius* spp. Goren (1996) noted similarities in colour and overlap of standard meristic counts. The molecular phylogenies estimated in this study (Fig. 3.7) provide evidence of these morphological similarities e.g. the unresolved Clade A.

Using DNA analysis, the study intended to test Goren's (1996) morphological concept of the species composition of *Caffrogobius* and generate a phylogenetic hypothesis for interrelationships within this genus. Thacker (2009) studied the phylogeny of Gobioidei based on DNA sequence data. The author found that the phylogeny of Gobiidae (using five genera belonging to the Gobiinae) exhibits a complex plot of lineages through time,

consistent with an elevation of speciation rate early in the clade's history and high species diversity in Gobiidae. Harold *et al*. (2008) studied phylogenetic relationships of Indo-Pacific coral gobies of the genus *Gobiodon*, based on morphological and molecular data. His findings indicated that a combination of sequence and morphological data provided the best resolved topology from which geographical deductions can be made. Although the present study focused on molecular phylogeny, with observed less emphasis on morphological aspects, considerable diversity has been observed in body shape and pigmentation among the species of this genus. Results from the present study indicate that the estuarine-dependent species, *C. natalensis*, appears as the basal group, suggesting that the genus might have evolved from an estuarine ancestor. The clade comprising of *C. gilchristi* and *C. caffer*, suggests a strongly supported relationship between estuarine and marine species. However, is should be kept in mind that *C. gilchristi* has marine larvae that may explain this relationship. Another sister grouping consisting of marine and estuarine species are formed by the *C. saldanha/C. agulhensis* and *C. nudiceps* clades respectively. Thus phylogenetic observations indicate a possible habitat shift, and an interaction between estuarine and marine species. However further study with more elaborate sampling of all species, including a morphological analysis, is necessary to be able to explain the current findings in more detail.

Thacker (2009) reported that the origin of Gobiidae indicated a habitat switch occurring from freshwater to marine environments. The invasion into a new complex, extensive environment may have opened the way to the tremendous diversification in marine Gobiidae, compared to that of brackish and freshwater species. In order to support these estimated relationships (see e.g. Figure 3.7), it is vital to conduct more extensive sampling over a larger distribution area and to use more molecular markers. Furthermore, the estimated relationships of the different species of *Caffrogobius* may change once the sister group of *Caffrogobius* is known.

5.2 Intraspecific genetic variation of the prison goby, *Caffrogobius gilchristi*

The prison goby, *Caffrogobius gilchristi*, is an estuarine-dependent species endemic to southern Africa, distributed from the Olifants River on the west coast to Durban on the east coast (Whitfield 1998). It inhabits shallow water, in the upper and middle reaches of the estuaries and tidal pools. In the present study, the intraspecific genetic structure in the prison goby in eastern and southern Cape estuaries was studied using partial mitochondrial DNA gene regions (cyt-*b* and COI). The DNA sequences of the cyt-*b* and COI genes used in this

study have shown sufficient variability for the purpose of population studies as seen in previous studies (Gysels *et al*. 2004; Craig *et al*. 2006; Maake 2009).

Haplotype distributions and AMOVA all suggest a single panmictic population (i.e. no significant population structure detected) within the prison goby's geographical range. Thus, Algoa Bay does not present barrier to the gene flow of this species, in contrast to reports for other species (Teske *et al*. 2003; Zardi *et al*. 2007). In addition, an investigation of phylogeographic patterns in *Glossogobius callidus* Smith, 1937 (Maake 2009) revealed two highly divergent lineages, which were associated with coastal biogeographic provinces, or gene flow barriers. The Eastern Cape lineage, which was phylogeographically structured, consisted of specimens that were all estuarine, while the KZN lineage, which was almost genetically homogenous, consisted of specimens from freshwater riverine systems. The *G. callidus* adults, in contrast to *C. gilchristi*, access freshwater and estuarine environments, suggesting a high potential dispersal rate which was not found. With regards to the *C. gilchristi*, the pelagic larvae may be responsible for the observed lack of population structuring due to high migration rates. The larvae of the prison goby are carried out to sea by the ebb tide. After about a month at sea they migrate back into an estuary as post-larvae. By contrast, larvae of *G. callidus* have been found in the lower reaches of estuaries (Strydom *et al*. 2003) and there is no record of them going out to sea. Hence, *G. callidus* was found to have genetical subdivision (Maake 2009), but *C. gilchristi* has no obvious subdivision. Although general aspects of the reproductive biology of the *C. gilchristi* are known, there is little or no data on the behaviour of their larvae in the sea to support and explain this notion of a single population encompassing their entire geographical range (Whitfield 1998).

Neethling *et al*. (2008) studied the phylogeography of the marine goby, *C. caffer*, a close relative to *C. gilchristi*. The study of *C. caffer* also demonstrated no genetic structuring according to population structure and gene flow analyses. The combination of higher haplotype diversity and lower nucleotide diversity suggested a signature of population expansion (Avise 2000). The study by Neethling *et al.* (2008) suggests that the species' pelagic larvae, or geologically-recent events, might be responsible for the high gene flow. The Agulhas Current plays an important role in the migration patterns of *C. caffer* and may influence dispersal capabilities of the pelagic larvae. In comparison to the current study of the prison goby, the species has been reported to release the bulk of their larvae into the open ocean during the tidal cycle (Whitfield 1998). This would increase the dispersal success of its

pelagic larvae. Strydom *et al*. (2003) reported that larvae of *Caffrogobius* species dominate the estuary-resident fishes, even though the specimen has not been identified to species level. This study and that of Neethling *et al*. (2008) showed that, in these popoulations of *C. gilchristi* and *C. caffer*, gene flow occurs across Cape Agulhas, even though some previous studies had suggested that Cape Agulhas could be a barrier to gene flow (Evans *et al*. 2004; Teske *et al* 2007a, b; von der Heyden *et al*. 2007; Zardi *et al*. 2007).

Oceanography can play a major role in structuring genetic subdivisions within populations. The genetic structure of marine and estuarine fauna of the eastern and southern Cape has been shaped by numerous biogeographic events driven by geological and climatic history of southern Africa. Larvae of the prison goby are highly mobile and probably disperse using sea currents, hence the high gene flow that was observed in this study. In the western part of the Port Elizabeth - Mossel Bay area, surface water tends to drift to the northwest whereas in the eastern part there is often a clockwise movement of water i.e. on-shore and to the east (Lutjehams & Ballegooyen 1988; Lutjehams & Ansorge 2001; Lutjehams 2005). The clockwise movement of water onshore would, however, mostly result in low larval dispersal, which was not the case in this study. This might imply that the larvae migrate further into the sea than anticipated, resulting in high larval dispersal.

Since a single population exists amongst the localities analysed for *C. gilchristi*, this means that conservation should not be focused on individual estuaries, but on a wider geographical scale. The results show that successful larval dispersal across documented oceanographic barriers, such as Algoa Bay and Cape Agulhas, is possible for some estuarine-dependent species with marine larvae. It is reasonable to assume that estuaries further from a particular spawning site would receive fewer recruits than estuaries closer to the spawning event. This is not a conclusive result because individuals analysed were a mixture of juveniles and adults from various cohorts collected over an extended period. More research, with a larger sample size from more localities needs to be carried out in order to get a clearer picture.

REFERENCES:

AKIHITO, P., IWATA, A., KOBAYASHI, T., IKEO, K., IMANISHI, T. ONO, H., UMEHARA, Y., HAMAMATSU, C., SUGIYAMA, K., IKEDA, Y., SAKAMOTO, K., FUMIHITO, A., OHNO, S. & GOJOBORO, T. 2000. Evolutionary aspects of gobioid fishes based upon a phylogenetic analysis of mitochondrial cytochrome-*b* genes. *Gene* **259**: 5-15.

AVISE, J.C. 2004. Molecular Markers, Natural History and Evolution. Sinauer Associates, Chapman & Hall, New York.

AVISE, J.C. 2004. What is the field of biogeography, and where is it going? *Taxon* **53**: 893-898.

AVISE, J.C. 2000. In: *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, MA.

AVISE, J.C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society London Biological Sciences* **312**: 325-342.

AVISE, J.C., BALL, M. & ARNOLD, J. 1988. Current versus historical population sizes in vertebrate species with high gene flow: A comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution* **5**: 331-344.

AVISE, J.C. & WALKER, D. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London Biological Sciences* **265**: 457-463.

AVISE, J.C., WALKER, D. & JOHNS, G.C. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London Biological Sciences* **265**: 1707-1712.

AVISE, J.C., ARNOLD, J., BALL, R.M., BERMINGHAM, E., LAMB, T., NEIGEL, J.E., REEB, C.A. & SAUNDERS, N.C. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecological Systematics* **18**: 489-522.

BIRDSONG, R.S., MURDY, E.O. & PEZOLD, F.L. 1988. A study of vertebral column and median fin osteology in Gobioid relationships. *Bulletin of Marine Science* **42**: 174-214.

BROWN, W.M., GEORGE Jr., M. & WILSON, A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Procedures of the Natural Academy of Science* USA **76**: 1967-1971.

BROWN, W.M., PRAGER, E.M., WANG, A. & WILSON, A.C. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *Journal of Molecular Evolution* **18**: 225-239.

CATERINO, M.S. & SPERLING, F.A.H. 1999. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Molecular Phylogenetics and Evolution* **11**: 122–137.

CLOSE, P.G. & GOUWS, G. 2007. Cryptic estuarine gobiid fish larva (Pisces: Gobiidae) identified as *Pseudogobius olurum* using allozyme markers. *Journal of Fish Biology* **71**: 288-293.

CRAIG, M.T., HASTINGS, P.A., PONDELLA II, D.J., ROBERTSON, D.R. & ROSALES-CASIAN, J.A. 2006. Phylogeography of the flag cabrilla *Epinephelus labriformes* (Serranidae): implications for the biogeography of the Tropical Eastern Pacific and the early stages of speciation in a marine shore fish. *Journal of Biogeography* **33**: 969-979.

CRANDALL, K.A., BININDA-EMONDS, O.R.P., MACE, G.M. & WAYNE, R.K. 2000. Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**: 290-295.

DAWSON, M.N. 2001. Phylogeography in coastal marine animals: a solution from California? *Journal of Biogeography* **28**: 723-736.

DOADRIO, I., CARMONA, J.A. & MACHORDOM, A. 2002. Haplotype diversity and phylogenetic relationships among the Iberian barbels (*Barbus*, Cyprinidae) reveal two evolutionary lineages. *The American Genetic Association* **93**:140-147.

DOBZHANSKY, T., AYALA, F.J., STEBBINS, G.L. & VALENTINE, J.W. 1977. *Evolution*. W.H. Freeman and Company, San Franscisco.

DOUGHERTY, J.D., MOORE, W.S. & RAM, J.L. 1995. Mitochondrial DNA analysis of round goby (*Neogobius melanostomus*) and tubenose goby (*Proterorhinus maximus*) in the Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Science* **53**: 474-480.

DURAN, S. 2003. Phylogeography, gene flow and population structure of *Crambe crambe* (Porifera: Poecilosclerida). Ph.D. thesis, University of Barcelona, Barcelona, Spain.

DUPANLOUP, I., SCHNEIDER, S. & EXCOFFIER, L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**: 2571-2581.

EDKINS, M.T., TESKE, P.R., PAPADOPOULOS, I. & GRIFFITHS, C.L. 2007. Morphological and genetic analyses suggest that southern African crown crabs, *Hymenosoma orbiculare*, represent five distinct species. *Crustaceana* **80**: 667–683

ESCHMEYER, W.N. & FONG, J.D. 2008. Species of fishes by family/subfamily. On-line version dated 07 August 2008. http:// research.calacademy.org /research /ichthyology /catalog /SpeciesByFamily.html

EVANS, B.S., SWEIJD, N.A., BOWIE, R.C.K., COOK, P.A. & ELLIOT, N.G. 2004. Population genetic structure of the perlemoen, *Haliotis midae*, in South Africa: evidence of range expansion and founder events. *Marine Ecology Progress Series* **270**: 163–172.

EXCOFFIER, L., LAVAL, G. & SCHNEIDER, S. 2006. Arlequin ver. 3.1: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**: 47-50.

EXCOFFIER, L., SMOUSE, P.E. & QUATTRO, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

FELSENTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

FELSENTEIN, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review in Genetics* **22**: 521-565.

FINNE, K.L. 2001. Phylogeographic structure of the Atlantic pupfish, *Cyprinodon variegates* (Cyprinodontidae), along the eastern coast of North America: Evidence from mitochondrial nucleotide sequences. M.Sc. thesis, Virginia Polytechnic Institute and State University Blacksburg, VA.

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

GERARD, K., ROB, C., CHEVALIER, N., THOMASSIN, B., CHENUIL, A. & FERAL, J-P. 2008. Assessment of three mitochondrial loci variability for the crown-of-thorns starfish: A first insight into *Acanthaster* phylogeography. *Comptes Rendus Biologies* **331**: 137-143.

GILL, H.S. & BRADLEY, J.S. 1992. Validation of the use of cephalic lateral-line papillae pattern for postulating relationships among gobioid genera. *Zoological Journal of Linnaean Society* **106**: 97-114.

GOREN, M. 1996. A review of the Southern African Gobiid fish genus *Caffrogobius* Smitt, 1900. *J.L.B. Smith Institute of Ichthyology Special Publication no.* **57:** 1-28.

GRANT, W.S. 1993. Population genetics in southern Africa: what are the questions? *Suid-Afrikaanse Tydskrif vir Wetenskap* **89**: 113-18.

GROSBERG, R.K. & CUNNINGHAM, C.W. 2001. Genetic structure in the sea: from populations to communities. In: *Marine Community Ecology*, (eds) M. D. Bertness, S. Gaines, and & Hay, M. E. Sinauer Associates, Sunderland, MA.

GRUENTHAL, K.M. & BURTON, R.S. 2008. Genetic structure of natural populations of the California black abalone (*Haliotis cracherodii* Leach, 1814), a candidate for endangered species status. *Journal of Experimental Marine Biology and Ecology* **355:** 47-58.

GYSELS, E.S., HELLEMANS, B., PAMPOULIE, C. & VOLCKAERT, A.M. 2004. Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Biology* **13**: 403-417.

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

HAPONSKI, A.E. & STEPIEN, C.A. 2008. Molecular, morphological, and biogeographic resolution of cryptic taxa in the Greenside Darter *Etheostoma blenniodes* complex. *Molecular Phylogenetics and Evolution* **49**: 69-83.

HARADA, S., JEON, S.-R., KINOSHITA, I., TANAKA, M. & NISHIDA, M. 2002. Phylogenetic relationship of four species of floating gobies (*Gymnogobius*) as inferred from partial mitochondrial cytochrome *b* gene sequences. *Ichthyological Research* **49**: 324-332.

HAROLD, A.S., WINTERBOTTOM, R., MUNDAY, P.L. & CHAPMAN, R.W. 2008. Phylogenetic relationships of Indo-Pacific Coral gobies of the genus *Gobiodon*

(Teleostei: Gobiidae), based on morphological and molecular data. *Bulletin of Marine Science* **82**: 119-136.

HARRISON, T.D. 2002. Preliminary assessment of the biogeography of fishes in South African estuaries. *Marine and Freshwater Research* **53**: 479–490.

HARRISON, T.D. 2004. Physico-chemical characteristics of South African estuaries in relation to the zoogeography of the region. *Estuarine, Coastal and Shelf Science* **61**: 73-87.

HEWITT, G.M. 2004. Genetic consequences of climatic oscillations in the Quatenary. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 183-195.

HILLIS, D.M. 1987. Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics* **18**: 23-42.

HILLIS, D.M. & MORITZ, C. 1990. *Molecular systematics*. Sinauer Associates, Sunderland. MA.

HILLIS, D.M., MORITZ, C. & MABLE, B.K. 1996. *Molecular systematics*. Second edition. Sinauer. Sunderland, MA.

HOESE, D.F. 1984. Gobioidei: relationships. In: *Ontogeny and Systematics of Fishes* (eds) Moser HG, Richards WJ, Cohen DM, Fahay MP, Kendall Jr AW, Richardson SL. *In: American Society Ichthyology and Herpetology Special Publication No. 1*. pp. 588-591, Allen Press, Lawrence, KS, In Anonymous.

HOESE, D.F. 1986. Gobiidae. In: *Smith's Sea Fishes*. Xx 1047, (eds) M. M. Smith, P. C. Heemstra) pp. 774-807, Macmillan South Africa, Johannesburg.

HOESE, D.F. & GILL, A.C. 1993. Phylogenetic relationships of Eleotride fishes (Perciformes: Gobioidae). *Bulletin of Marine Science* **52**: 415-440.

HOESE, D.F. & WINTERBOTTOM, R. 1979a. A new genus and four new species of fishes from the Indo-Western Pacific (Pisces; Perciformes; Gobiidae), with comments on relationships. *Royal Ontorio Museum Life Science Occasional Papers* **37:** 1-17.

HOESE, D.F. & WINTERBOTTOM, R. 1979b. A new species of *Lioteres* (Pisces: Gobiidae) from Kwazulu, with a revised checklist of South African Gobies and comments on the generic relationships and endemism of Western Indian Ocean Gobioids. *Royal Ontorio Museum Life Science Occasional Papers* **31:** 1-13.

HUANG, J., XU, Q., SUN, Z.J., TANG, G.L. & SU, Z.Y. 2007. Identifying earthworms through DNA barcodes. *Pedobiologia* **51:** 301-309.

HUELSENBECK, J.P. & RONQUIST, F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754-755.

IRWIN, D.M., KOCHER, T.D. & WILSON, A.C. 1991. Evolution of the cytochrome *b* Gene of Mammals. *Journal of Molecular Evolution* **32**: 128-144.

IWATA, A., HOSOYA, S. & LARSON, H.K. 2001. *Paedogobius kimurai*, a new genus and species of goby (Teleostei: Gobioidei) from the west Pacific. *Records of the Australian Museum* **53**: 103-112.

KLOPPER, A.W. 2005. Intraspecific genetic variation in the percoid teleosts *Argyrosomus japonicus* (Temminck and Schlegel, 1843) and *Pomadasys commersonnii* (Lacepede, 1801) as inferred from the mitochondrial control region. M.Sc. thesis, University of Pretoria, South Africa.

KOCHER, T.D. & STEPIEN, C.A. 1997. *Molecular Sytematics of Fishes*. Academic Press, USA.

KVIST, L. 2000. Phylogeny and Phylogeography of European parids. Academic Dissertation, University of Oulu, Finland.

LARSON, H.K. 2001. A revision of the gobiid fish genus *Mugilogobius* (Teleostei: Gobioidei), and its systematic placement. *Records of the West Australian Museum Supplement* No. **62**.

LUNT, D.H., ZHANG, D.-X., SZYMURA, J.M. & HEWITT, G.M. 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molecular Biology* **5**: 153–165.

LUTJEHARMS, J.R.E. 2005. The coastal oceans of south-eastern Africa. In: *The Sea* (eds). ROBINSON, A.R. and BRINK, K, Pp. 781-832. Chicago University Press, Chicago.

LUTJEHARMS, J.R.E. 1988. Examples of extreme circulation events at the Agulhas retroflection. *South African Journal of Science* **84**: 584-586.

LUTJEHARMS, J.R.E. & ANSORGE, I.J. 2001. The Agulhas Return Current. *Journal of Marine Systems* **30**: 115-138.

LUTJEHARMS, J.R.E. & COOPER, J. 1996. Interbasin leakage through Agulhas filaments. *Deep-Sea Research* **43**: 213-238.

LUTJEHARMS, J.R.E. & GORDON, A.L. 1987. Shedding of an Agulhas Ring observed at sea. *Nature* **325**: 138-140.

LUTJEHARMS, J.R.E. & VAN BALLEGOOYEN, R.C. 1984. Topographic control in the Agulhas Current systems. *Deep-Sea Research* **31**: 1321-1337.

LUTJEHARMS, J.R.E. & VAN BALLEGOOYEN, R.C. 1988. Anomalous upstream retroflection in the Agulhas Current. *Science* **240**: 1770-1772.

LUTJEHARMS, J.R.E., VALENTINE, H.R., VAN BALLEGOOYEN, R.C. 1993. On the subtropical convergence in the South Atlantic Ocean. *South African Journal of Science* **89**: 552-559.

MAAKE, P.A. 2009. Molecular systematics of the South African gobiid fishes of the genus *Glossogobius* Gill, 1862 (Teleostei: Gobiidae). M.Sc. dissertation, University of Limpopo, South Africa.

MARTINS, C., WASKO, A.P., OLIVIERA, C. & FORESTI, F. 2003. Mitochondrial DNA variation in wild populations of *Leporinus elongatus* from the Parana River basin. *Genetics and Molecular Biology* **26**: 1-10.

MATSUBARA, K. & IWAI, T. 1959. Description of a new sandfish, *Kraemeria sexradiata*, from Japan, with special reference to its osteology. *Journal of Washington Academy of Science* **49**: 27-32.

MAUGE, A.L. 1986. Gobiidae. In: *Checklist of the Freshwater Fishes of Africa* (eds) Daget J, Gosse JP & Thys van den Audenaerde DFE (CLOFFA) **2**: 358-388.

MEYER, A. 1993. Evolution of mitochondrial DNA in fishes. In: *Molecular biology frontiers, biochemistry and molecular biology of fishes* (eds). Hochachka, IW and Mommsen TP. Volume **2**, Elsvier Science Publishers, Amsterdam.

MEYER, A. 1994. Shortcomings of the cytochrome-*b* gene as a molecular marker. *Trends in Ecology and Evolution* **9**: 278-280.

MILLER, P. J. 1978. The status of the west African fish *Gobius nigricinctus* with reference to New World autochthones and an Old World colour-analogue. *Zoological Journal of the Linnean Society* **64**: 27-39.

MOORE, W.S. 1995. Inferring phylogenies from DNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**: 718-726.

MOORE, W.S. & DEFILLIPSS, V.R. 1997 Taxonomic resolution based on cytochrome *b* DNA. *In Avian Molecular Evolution and Systematics* (ed) Mindell D.P. Academic Press, San Diego, California.

MORITZ, C., DOWLING, T.E. & BROWN, W.M. 1987 Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual Review of Ecological Systematics* **18**: 269-292.

MULLER, K. 2004. PRAP-computation of Bremer support for large data sets. *Molecular Phylogenetics and Evolution* **31**: 780-782.

NEETHLING, M., MATHEE, C.A., BOWIE, R.C.K. & VON DER HEYDEN, S. 2008. Evidence for panmixia despite barriers to gene flow in the southern African endemic, *Caffrogobius caffer* (Teleostei: Gobiidae). *BioMed Central Evolutionary Biology* **8**: 325-334.

NEI, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.

NETO, M.A.F. 1998. Molecular systematic and population genetics of marine vertebrates from Brazil. Ph.D. thesis, Memorial University of Newfoundland, Newfoundland, Canada.

NORTON, O.B. 2005. The population structure of two estuarine fish species, *Atherina breviceps* (Pisces: Atherinidae) and *Gilchristella aestuaria* (Pisces: Clupeidae), along the southern African coastline. M.Sc. thesis, Rhodes University, South Africa.

OOSTHUIZEN, A., JIWAJI, M. & SHAW, P. 2004. Genetic analysis of the *Octopus vulgaris* population on the coast of South Africa. *South African Journal of Science* **100**: 603-607.

PALUMBI, S.R. 1992. Marine speciation on a small planet. *Trends in Ecology and Evolution* **7**: 114-118.

PALUMBI, S.R. 1994. Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecological Systematics* **25**: 547-572.

PAGE, T.J. 2006. An evolutionary history of the freshwater shrimp family Atyidae in Australia. Ph.D. thesis, Griffith University, Brisbane, Australia.

POLZIN, T. & DANESCHMAND, S.V. 2003. On Steiner trees and minimum spanning trees in hypergraphs. *Operations Reserach Letters* **31**: 12-20.

POSADA, D. & CRANDALL, K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.

PRUETT, L.C., SAILLANT, E. & GOLD, R.J. 2005. Historical population demography of red snapper (*Lutjanus campechanus*) from the northern Gulf of Mexico based on analysis of sequences of mitochondrial DNA. *Marine Biology* **147**: 593-602.

QI, D., GUO, S., ZHAO, X., YANG, J. & TANG, W. 2007. Genetic diversity and historical structure of *Schizopygopsis pylzovi* (Teleostei: Cyprinidae) in the Qinghai-Tibetan Plateau. *Freshwater Biology* **52**: 1090-1104.

RAMOS-ONSINS, S. & ROZAS, J. Statistical properties of a new neutrality test against population growth. *Molecular Phylogenetics and Evolution* **19**: 2092-2100.

RAND, D.M. 1996. Neutrality tests of molecular markers and the connection between DNA polymorphism, demography, and conservation biology. *Conservation Biology* **10**: 665-671.

RIDLEY, M. 1993. *Evolution*. Blackwell Science Inc., Oxford.

RODRIGUEZ, D. 2007. Crocodilian evolution, systematics and population genetics: Recovery and ecological interactions of the American crocodile (*Crocodylus acutus*). Ph.D. thesis, Texas Tech University, Texas, USA.

ROE, A.D. & SPERLING, F.A.H. 2007. Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding. *Molecular Phylogenetics and Evolution* **44**: 325-345.

ROGERS, A.R. & HARPENDING, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-569.

RONQUIST, F. & HUELSENBECK, J.P. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17**: 754-755.

ROZAS, J., SACHEZ-DELBARRIO, J.C., MESSEGUER, X. & ROZAS, R. 2005. DNA Sequence Polymorphism version 4.10.5 for 32-bit Windows. Universitat de Barcelona.

SARASTE, M. 1990. Structural features of cytochrome oxidase. *Quarterly Review of Biophysics* **23**: 331–366.

SCHMIDT, P.S. & RAND, D.M. 2001. Adaptive maintenance of genetic polymorphism in an intertidal barnacle habitat and life stage specific survivorship of MPI genotypes. *Evolution* **55**: 1336-1344.

SIMON, C., FRATI, F., BECKENBACH, A., CRESPI, B., LIU, H. & FLOOK, P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**: 651–701.

SKELTON, P.H. 1994. *Baseline survey of fishes in the Keurbooms/Bietou and Piesary Rivers, southern Cape Unpublished report.* J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

SMITH, M.M. & HEEMSTRA, P.C. 1986. *Smith's Sea Fishes*. MacMillan, Johannesburg, South Africa.

STAHLS, G. & NYBLOM, K. 2000. Phylogenetic analysis of the genus *Chleilosia* (Diptera, Syrphidae) using mitochondrial COI sequence data. *Molecular Phylogenetics and Evolution* **15**: 235-241.

STASOFT, INC. 2008. *Electronic Statistics Textbook*. Tulsa, OK: StaSoft. WEB: http://www.stasoft.com/textbook/.

STEFANNI, S. & THORLEY, L.J. 2003. Mitochondrial DNA phylogeography reveals the existence of an evolutionary significant unit of the sand goby *Pomatoschistus minutus* in the Adriatic (Eastern Mediterranean). *Molecular Phylogenetics and Evolution* **28**: 601-609.

STEPIEN, C.A. & ROSENBLATT, R.H. 1991. Patterns of gene flow and genetic divergence in the Northeastern Pacific Clinidae (Teleostei: Blennioidei), based on allozyme and morphological data. *Copeia* **4**: 873-896.

STEPIEN, C.A. & TUMEO, M.A. 2006. Invasion genetics of Pont-Caspian gobies in the Great Lakes: a 'cryptic' species, absence of founder effects, and comparative risk analysis. *Biological Invasions* **8**: 61-78.

STRYDOM, N.A. 1998. Fish larval dynamics in the mouth region of the Gamtoos Estuary. M.Sc. Thesis, University of Port Elizabeth, Port Elizabeth, South Africa.

STRYDOM, N.A. & NEIRA, F.J. 2006. Description and ecology of larvae of *Glossogobius callidus* and *Redigobius dewaali* (Gobiidae) from temperate South African estuaries. *African Zoology* **41**: 240-251.

STRYDOM, N.A. & WHITFIELD, A.K. 2000. The effects of a single freshwater release into the Kromme Estuary. 4: Larval fish response. *Water SA* **26**: 319-328

STRYDOM, N.A., WHITFIELD, A.K. & WOOLDRIDGE, T.H. 2003. The role of estuarine type in characterizing early stage fish assemblages in warm temperate estuaries, South Africa. *African Zoology* **38**: 219-243.

SWOFFORD, D.L. 2001. PAUP*: phylogenetic analysis using parsimony (*and Other Methods). Version 4, Sinauer Associates, Sunderland, MA.

TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.

TESKE, P.R., CHERRY, M.J. & MATHEE, C.A. 2003. Population genetics of the endangered Knysna seahorse, *Hippocampus capensis*. *Molecular Ecology* **12**: 1703-1715.

TESKE, P.R., CHERRY, M.I. & MATTHEE, C.A. 2004. The evolutionary history of seahorses (Syngnathidae: *Hippocampus*): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. *Molecular Phylogenetics and Evolution* **30**: 273–286

TESKE, P.R., McQUAID, C.D., FRONEMAN, P.W., BARKER, N.P. 2006. Impacts of marine biogeographic boundaries on phylogeographic patterns of three South African estuarine crustaceans. *Marine Ecology Progress Series* **314**: 283–293

TESKE, P.R., BARKER, N.P., OOSTHUIZEN, A. & PAPADOPOULOS, I. 2007a. Phylogeographic structure of *Octopus vulgaris* in South Africa revisited: identification of a second lineage near Durban harbour. *Marine Biology* **151**: 2119-2122.

TESKE, P.R., FRONEMAN, P.W., BARKER, N.P. & McQUAID, C.D. 2007b. Phylogeographic structure of the caridean shrimp *Palaemon peringueyi* in South Africa: further evidence for intraspecific genetic units associated with marine biogeographic provinces. *African Journal of Marine Science* **29**: 253-258.

TESKE, P.R., PAPADOPOULOS, I., ZARDI, G.I., McQUAID, C.D., EDKINS, M.T., GRIFFITHS, C.L. & BARKER, N.P. 2007c. Implications of life history for genetic structure and migration rates of southern African coastal invertebrates: planktonic, abbreviated and direct development. *Marine Biology* **152**: 697-711.

THACKER, C.E. 2003. Molecular phylogeny of the gobioid fishes (Teleostei: Perciformes: Gobioidei). *Molecular Phylogenetics and Evolution* **26:** 354-368.

THACKER, C.E. 2009. Phylogeny of Gobioidei and placement within Acanthomorpha, with a new classification and investigation of diversification and character evoliution. *Copeia* **1**: 93-104.

THACKER, C.E. & HARDMAN, M.A. 2005. Molecular phylogeny of basal gobioid fishes: Rhyacichthyidae, Odontobutidae, Xenisthmidae, Eleotridae (Teleostei: Perciformes: Gobioidei). *Molecular Phylogenetics and Evolution* **37**: 858-871.

TRONTELJ, P., MACHINO, Y. & SKET, B. 2005. Phylogenetic and phylogeographic relationships in the crayfish genus *Austropotamobius* inferred from mitochondrial COI gene sequences. *Molecular Phylogenetics and Evolution* **34**: 212-226.

VANDEWOESTIJNE, S., BAGUETTE, M., BRAKEFIELD, P.M. & SACCHERI, I.J. 2004. Phylogeography of *Aglais urticae* (Lepidoptera) based on DNA sequences of the mitochondrial COI gene and control region. *Molecular Phylogenetics and Evolution* **31**: 630-646.

VAN NIEKERK, L. 2007. In: A Review of Information on Temporary Open/Closed Estuaries in the Warm and Cool Temperate Biogeographic Regions of South Africa, with Particular Emphasis on the Influence of River Flow on These Systems. Eds: WHITFIELD, A.K. & BATE. WRC Report No 1581/1/07.

VAN TASSELL, J.L., MILLER, P.J. & BRITO, A. 1988. A revision of *Vanneaugobius* (Teleostei: Gobiidae), with description of a new species. *Journal of Natural History* **22**: 545-567.

VERSPOOR, E., STRADMEYER, L. & NIELSEN, J.L. 2007. *The Atlantic Salmon Genetics, Conservation and Management*. Blackwell Publishing Ltd, UK.

VON DER HEYDEN, S., LIPINSKI, M.R. & MATTHEE, C.A. 2007. Mitochondrial DNA analyses of the Cape hakes reveal an expanding population for *Merluccius capensis* and population structuring for mature fish in *Merluccius paradoxus*. *Molecular Phylogenetics and Evolution* **42**: 517-527.

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

WANG, X., WANG, J., HE, S. & MAYDEN, R.L. 2007. The complete mitochondrial genome of the Chinese hook snout carp *Opsariichthys bidens* (Actinopterygii: Cypriniformes) and an alternative pattern of mitogenomic evolution in vertebrate. *Gene* **399**: 11-19.

WHITFIELD, A.K. 1998. Biology and Ecology of Fishes in Southern African Estuaries. *Ichthyological Monograph of the J.L.B Smith Institute of Ichthyology* No. **2.**

WHITFIELD, A.K. & BATE, G. 2007. *A Review of Information on Temporary Open/Closed Estuaries in the Warm and Cool Temperate Biogeographic Regions of South Africa, with Particular Emphasis on the Influence of River Flow on These Systems*. Water Research Commission Report No **1581**/1/07.

WOOLDRIDGE, T.H. 1991. Exchange of two species of decapod larvae across an estuarine mouth inlet and implications of anthropogenic changes in the frequency and duration of mouth closure. *South African Journal of Zoology* **87**: 519–525.

ZARDI, G.I., McQUAID, C.D., TESKE, P.R., BARKER, N.P. 2007. Unexpected genetic structure of mussel populations in South Africa: indigenous *Perna perna* and invasive *Mytilus galloprovincialis*. *Marine Ecology Progress Series* **337**: 135–144.

ZHANG, J., CAI, Z. & HUANG, L. 2006. Population genetic structure of crimson snappe *Lutjanus erythropterus* in East Africa, revealed by analysis of the mitochondrial control region. *Journal of Marine Science* **63**: 693-704.