Combining DNA barcoding and morphology to identify larval fishes from the nearshore environment off the south-east coast of South Africa

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

MASTERS OF SCIENCE

at

RHODES UNIVERSITY

By

Zinzi Sinazo Somana

June 2019

DECLARATION

I declare that the dissertation here submitted to the Rhodes University in Grahamstown, for the degree of Master of Science in Zoology and Entomology, has not been submitted in whole or part for any degree or examination at any university. I am aware of the University's policy on plagiarism. The work contained herein is my own and the material (ideas, phrases and illustrations) presented from other authors has been duly acknowledged.

Name: Zinzi Sinazo Somana

Student Number: g17S4592

Table of Contents

DECLARATION i

ABSTRACT vi

ACKNOWLEDGEMENTS viii

CHAPTER ONE 1

1.1 Background to the study 1

1.2 Statement of the problem 4

1.3 Rationale/justification of the study 5

1.4 Aims and Objectives 5

CHAPTER TWO 6

Identification of larval fishes from the south-east coast of South Africa using morphological and DNA barcoding techniques 6

2.1 Introduction 6

2.2 Materials and Methods 9

2.2.1 Field work 9

2.2.2 Laboratory work 10

2.2.3 Morphological identification 10

2.2.4 DNA extraction, PCR amplification and sequencing 11

2.2.5 Data Analysis 12

2.3 Results 14

2.3.1 Morphological identification 14

2.3.2 Molecular identification 15

2.3 Discussion 30

CHAPTER THREE 36

The use of DNA barcoding for the identification of the early life history stages of a cryptic benthic fish species 36

3.1 Introduction 36

3.2 Material and Methods 40

3.2.1 DNA extraction, PCR amplification and sequencing 40

3.2.2 Sequence alignment and analysis 40

3.2.3 Phylogenetic analysis 41

3.3 Results 42

3.3.1 *COI* sequences 42

3.3.2 *16S* sequences 43

3.4 Discussion 46

CHAPTER FOUR 50

4.1 General Discussion 50

4.2 Conclusion and Recommendations 54
References 56

List of Tables

Table 2.1: Total number of specimens collected and identified to family level using morphometric characters.. 15

Table 2.2: Sequence match for successfully amplified egg and larval specimens. The range of similarity match (%) for each species is shown for BOLD and NCBI databases. Number of specimens indicated in brackets. 16

Table 2.3: Sequences without closest match together with those identified to family level on BOLD. Indicted in Column 2 are the similarity values from GenBank. 20

Table 2.4: K2P genetic distances among the close matches and the sequences categorised as unknowns. Highlighted is the lowest genetic distance between FBIP0120 and Blenniidae *sp.* 21

Table 2.5: K2P values between Gobiidae species sourced from BOLD and NCBI together with the larval sequences derived from this study. The lowest and highest genetic distances among sequences are highlighted in bold.. 22

Table 2.6: Photographs and identifications of individual specimens, based on traditional morphology and DNA barcoding techniques.. 23

Table 3.1: Top matching sequences for the *COI* gene fragment from BOLD and NCBI databases. 42

Table 3.2: K2P genetic distances between sequences from BOLD (Tripterygiidae), GenBank (*P. profundus*) and the unidentified specimens calculated in MEGA based on *COI* sequences. 44

List of Figures

Figure 2.1: Sampling stations (marked with black dots) within the shallow nearshore region of the Eastern Cape along the south-east coast of South Africa. 10

Figure 2.2: An example of the morphometric measurements collected during morphological identification based on photograph of a Soleidae larva. BD = body depth, PAL = preanal length and TL = total body length... 11

Figure 2.3: Examples of damaged fish larvae that could not be morphologically identified, but were used for DNA barcoding.. 14

Figure 2.4: Photographs of specimens with successful sequences that failed to find matches on the BOLD online platform and were identified through morphological characteristics. (A) Soleidae (FBIP0143); (B) Gobiidae (FBIP0087); (C) Callionymidae (FBIP0365) and (D) Blenniidae (FBIP0398)... 17

Figure 2.5: Neighbor-joining tree of the *COI* sequences showing the divergence patterns and distance relationships between sequences. Bootstrap values are represented by number at the nodes. Black sample names represent the species names for larval samples of the current study. Red sample names represent unpublished sequence data from BOLD. Blue sample names are extracted from public data on BOLD. Green samples are the probability matches to from NCBI. Samples without a closest match on BOLD and NCBI are presented as an FBIP number. 18

Figure 2.6: Maximum likelihood phylogenetic tree of the *COI* sequences showing the divergence patterns and distance relationships between sequences. Bootstrap values are represented by number at the nodes. Black sample names represent the unidentified specimen of the current study. Red sample names represent a bony fish sequence data from BOLD. Green samples represent the sequences of fish species found in South Africa. 19

Figure 3.1: Neighbor-joining tree based on K2P genetic distances of the *COI* sequences for an egg and larval fishes (in black) as well as accessioned sequences of Tripterygiidae larvae from BOLD (in green) and a Clinidae adult from GenBank (in red). The scale bar represents 0.05 sequence divergence. 45

ABSTRACT

The early life history stages of most marine fish species are undescribed. The problem is, most of these fishes have pelagic larvae which are minute, delicate forms. Linking the larval stage to an adult counterpart is extremely challenging as larvae are morphologically different from the adults. Historically, larval fish identification relied solely on distinguishing morphological characteristics and meristic measurements, which has resulted in taxonomic confusion and misidentification. The introduction of the deoxyribonucleic acid (DNA) barcoding technique as an alternative approach has been successful in positively identifying larval fishes. The correct identification of larval specimens is the key to a better understanding of larval ecology, which underpins the success of any adult fish population. This study aimed to positively identify larval fishes of the south-east coast of South Africa using morphological characteristics and DNA barcoding.

Larval and eggs specimens for this study were collected from the shallow nearshore waters of the south-east coast of South Africa. A total of 177 larval specimens were used for morphological analysis. Body shape, gut shape, pigmentation and morphometric measurements (such as body depth, preanal length and total body length) were used to identify each specimen to the family level. In addition, a fragment of mitochondrial cytochrome *c* oxidase subunit 1 gene (*COI*) was adopted for sequencing to identify larval fish specimens and fish eggs. Sequences generated from this study were compared to those in the Barcode of Life Database (BOLD). When there were no close matches to a sequence, the GenBank nucleic acid sequence database, maintained by the National Center for Biotechnology Information (NCBI), was used as an alternative.

A total of 18 different families were identified through morphology. Seventy-seven of the 177 larval specimens were not subjected to morphological identification due to physical damage. The majority of larvae identified using morphological characteristics belonged to either the Sparidae, Tripterygiidae or Gobiesocidae fish families. Through DNA barcoding, 12 fish families, 16 genera and 18 different species were identified. Ten DNA barcodes (categorised as ‘no match’) from 10 different larval specimens were not identified through any of the online databases. Therefore, the 2% threshold value was used to identify members of the same species. The K2P genetic distance relationships were calculated among the no match sequences and downloaded probability matches from NCBI. This resulted in two unknown specimens assigned to the Blenniidae and Gobiidae. All other taxa were identified to species level, except specimens representing the Gobiidae and Tripterygiidae families. Based on the K2P genetic distances Gobiidae representatives were categorised as members of the *Caffrogobius* genus.

Twenty-eight barcodes represented specimens from the Tripterygiidae. DNA barcode data from *COI* was analysed using the standard phylogenetic procedures in MEGA6 to examine relationships and differentiation among sequences. These could not be identified to the lowest taxonomic rank due to limited sequence data to compare them with. The sequence data from these specimens gave different results in the two online databases. BOLD results were to family level (Tripterygiidae) and NCBI to the species level (Clinidae: *Pavoclinus profundus*).

Results in this study confirmed the efficiency of the DNA barcoding technique in species level identification of fish larvae. The evidence from genetic barcodes of the Tripterygiidae specimens, supported by morphological characteristics, suggests the need for thorough research to identify the individuals to the species level. The fact that this study identified taxonomically problematic Gobiidae and Tripterygiidae specimens suggests that studies similar to this may highlight additional diversity and help to resolve the taxonomy of other species in these families. However, the lack of reference sequence data from the adult specimens, and especially those with cryptic diversity, were both shortcomings for the positive identification of larvae. With that being said, it shows the necessity for more research to be conducted on barcoding of larvae in general as to accommodate all kinds of species from biodiversity to economic perspectives.

／

CHAPTER ONE

General Introduction

1.1 Background to the study

The primary task of taxonomy is to identify, delineate species, and to improve available information to further assign specimens to nominal species (Mallet & Willmott, 2003; Seberg *et al.,* 2003; Godfray, 2007). A set of individuals or a population that has an ability to interbreed in the wild, leading to the production of offspring that have the capability to reproduce, and which are recognised in terms of physical differences, can be referred to as a biological species (Mayr, 2000). Nonetheless, taxonomists often disagree on a species' status when the distinguishing physical differences are small or overlap (Pinceel *et al.,* 2005; Horne *et al.,* 2008).

Different concepts have been proposed to provide guidelines in defining species of a particular taxon (Winston, 1999). In order to provide boundaries for species delineation and description however, one has to adopt either one or more of these species' concepts (Aldhebiani, 2018). For example, the biological species concept emphasises reproductive isolation (Mayr, 1942; Coyne & Orr, 2004). Mayr (2000) defined biological species as the set of individuals or populations that interbreed in the wild, resulting in the production of offspring that also have the capacity to reproduce. The Ecological Species Concept highlights the occupation of a distinct niche or adaptive zone (Van Valen, 1976; Andersson, 1990), meaning organisms are adapted to the habitats they occupy and resources they exploit (Ridley, 2004). Only species that occupy satisfactorily different niches can coexist, however. Should their niche be similar or relatively so, the effective competitor could drive the lesser one to extinction (Ridley, 2004). The Evolutionary Species Concept is concerned with speciation (Bernardi, 2013). A species is defined as a single lineage that maintains its identity from other lineages with its own evolutionary tendencies (Wiley, 1978). The Evolutionary Species Concept does not focus on species as a class of ecological entities and, for that reason, the concept is capable of dealing with species as behavioural, phonetic, physiological, ecological, epigenetic, genetic, temporal and spatial entities (Coyne & Orr, 2004). The Phylogenetic Species Concept stresses that population lineages are different (Cracraft, 1983; Nixon & Wheeler, 1990) and monophyletic (Rosen, 1979; Donoghue, 1985). Species are defined as a group of organisms that are diagnosable and distinct (recognised by one or more diagnostic characters) and with a genealogical pattern of ancestry (Cracraft, 1983). In contrast to the above-mentioned definitions of species, most

taxonomists have adopted the Morphological Species Concept, which characterises a species by body shape and other structural features (Mayr, 1982). Dobzhansky (1950) included gene pool in the definition and defined a species as a community of individuals that can reproduce sexually, cross-fertilise and share the same gene pool. According to Cladridge *et al.* (1997), the Morphological Species Concept is not a concept, but rather, a method of description because the Morphological Species Concept does not treat species as historical entities that form lineages, meaning that a species definition will change as the species changes through time (Cladridge *et al.,* 1997). Although different species concepts have been recommended for larval fish delineation (Bowen *et al.,* 2006; Overdyk *et al.,* 2016), most have applied similar boundaries for species identification. Considering the various definitions of concepts, the Morphological Species Concept and the Phylogenetic Species Concept were adopted in the current study.

The differences in morphological features are therefore used to determine whether a certain individual belongs to the same or a different species (Mayr, 1982). The Morphological Species Concept may lead to the recognition of new species when two or more phenotypes within a species are encountered. When morphologically similar species, or two populations that have recently evolved from a common ancestor, are encountered, the Morphological Species Concept could, however, be misleading (Ridley, 2004) because overlap in phenetic similarities in stages of development of clear cryptic taxa could result in incorrect identity (Hyde *et al.,* 2005; Gleason & Burton, 2012). The Morphological Species Concept does not discriminate the species in reproductive terms as the biological species concept does, but the morphological inconsistences recognised among species do reflect the biological effects of isolation, common interbreeding and genetic divergence (Stuessy, 1972; Hull, 1970).

On the other hand, the Phylogenetic Species Concept emphasises the study of stages in the evolutionary history of an organism (Mayden, 1997; Coyne & Orr, 2004). Cracraft (1983) defines a species as a group of organisms which are distinct and have a genealogical pattern of ancestry and descent. In the Phylogenetic Species Concept, one or more characters are used to distinguish distinct groups (Cracraft, 1989). The distinguishing characters can be morphological and/or molecular, that is, DNA sequences (Cracraft, 1989; Coyne & Orr, 2004). When morphological characters fail to delineate species, sequence-based analysis can be used (Zhang *et al.,* 2013). Sequence-based analysis can be done by looking at the pairwise differences between all sequences of a barcode data set understudied, which involves observation of the gap between *COI* sequences. Based on the studies conducted prior the current study, the DNA barcoding gap is regarded as efficient when the COI sequences sampled within the same species are normally more alike than the ones extracted

from different species (Puillandre *et al.,* 2011). However, a slight difference in colour, body shape, or change in a single nucleotide in the DNA sequence will result in a different species, tremendously and illogically escalating the number of species recognised (Coyne & Orr, 2004).

Most marine fishes begin their lives as microscopic eggs, hatch as planktonic larvae, and then disperse in the water column (Moser *et al.,* 1984; Leis, 1991; Greer *et al.,* 2016). The planktonic larval stage increases the chances for long-distance dispersal (Riginos *et al.*, 2011). The advantages of having a planktonic larval stage include the ability to easily colonise new habitats, a wide distribution of individuals, potential for increased population connectivity and the elimination of competition for food and space between larvae and conspecific adults (Sale, 1991; Strathmann *et al.,* 2002). The time between hatching and full attainment of adult morphological features can take from several days or weeks to months, depending on the species (Webb, 1999). Defining the larval period in fishes is not straightforward and several definitions have been reported (Barrington, 1961; 1968; Balon, 1979; 1981; 1986; 1990; 1999; Crawford & Balon, 1996; Urho, 2002; Penaz, 2001). The three clear, early life history stages of most fishes include the egg, larva, and juvenile (Moser & Ahlstrom, 1970; Kendall *et al.,* 1984; Leis & Trnski, 1989), with the larval stage further subdivided into the yolk-sac, preflexion, flexion and postflexion stages (Kendall *et al.,* 1984).

The yolk-sac developmental stage commences with hatching, with the yolk-sac categorised as an embryonic organ that plays a vital role in nourishment of the larva after hatching (Leis & Carson-Ewart, 2000; Urho, 2002). The preflexion stage starts once both hatching and yolk-sac absorption are complete and ends at the start of the upward flexion of the notochord (Kendall *et al.,* 1984; Leis & Carson-Ewart, 2000). The flexion stage, which is when the notochord tip reaches its final position at a 45-degree angle from the notochord axis and the caudal fin rays form concurrently, is also associated with changes in body shape and feeding behaviour (Blaxter, 1986). Postflexion commences with the full completion of notochord flexion and is the beginning of the transformation to the juvenile stage (Leis & Carson-Ewart, 2000; Miller & Kendall, 2009). Such complex and variable larval phases are characterised by the presence of embryonic and/or specialised organs that are lost as development continues (Leis *et al.*, 2015).

1.2 Statement of the problem

Ichthyoplankton assemblages are typically composed of individuals belonging to multiple species, and studies assessing these groupings are used to evaluate habitat use and trophic ecology, amongst others (Boehlert & Mundy, 1993; Richardson *et al.,* 2007). A taxonomic resolution of ichthyoplankton and identification to species level could further help in locating spawning areas (Govoni *et al.,* 2003; Serafy *et al.,* 2003), understanding the distribution of rare or cryptic species (Richardson & Cowen, 2004), and in estimating population size (Ralston *et al.,* 2003). According to Simon and Lyons (1995), ichthyoplankton assemblages are important indicators of the state of health of aquatic ecosystems, as the pelagic larval stages regulate the dynamics of local populations and further influence the demographics of adjacent communities through immigration and emigration (Sale, 1991; Caley *et al.,* 1996). A dispersive larval phase also implies that the offspring can settle in any habitat where conditions are suitable for development and final recruitment (Johnson, 2005), which may differ from the location of the original population (Sinclair, 1988; Jones *et al.,* 2005; Almany *et al.,* 2007). Due to this potential spatial variability in the occurrence of the different life stages, as well as the ontogenetic morphological changes during fish development (Kendall *et al.,* 1984; Bingpeng *et al.*, 2018), it is hard to track and match larvae to their corresponding adults (Hare *et al.,* 1999, Pineda *et al.,* 2007; Valdez-Moreno, 2010).

Body shape, pigmentation, and morphometrics (measurements) are commonly used to identify a fish larva (Matarese *et al.,* 2011; Ko *et al.*, 2013; Azmir *et al.,* 2017). Morphometric characters do, however, change as the larva develops from preflexion, through postflexion to the juvenile stage (Zhang & Hanner, 2011; Ko *et al.,* 2013). In addition, larvae of marine fish species of certain fish families also have similar morphological characteristics with overlapping morphometric measurements (Victor, 2009). Therefore, species identification from egg to the larval stage can be challenging and time consuming (Neira *et al.*, 2015).

The DNA barcoding approach helps to distinguish similar looking species independently of developmental stage or physical condition (Hebert *et al.,* 2003; Hubert *et al.,* 2008). This approach has been used successfully in the positive identification of the early life history stages of marine fishes, such as eggs (Shao *et al.*, 2002; Burrows *et al.*, 2018), preflexion through to postflexion larval stages (Valdez-Moreno *et al.*, 2010; Ko *et al.*, 2013; Friedheim, 2016), as well as juveniles (Pegg *et al.*, 2006; Gleason & Burton, 2012; Ardura *et al.*, 2016).

1.3 Rationale/justification of the study

The use of integrative taxonomy, or integrated concepts, is the best practice to combine a set of unifying properties within a concept for species identification in order to incorporate the mechanisms of speciation in nature (De-Queiroz, 2007). In cases where the morphological identification of biological species is impossible, DNA sequences of highly conserved genes can be used (Collin & Cruickshank, 2013). An alternative approach in species identification, DNA barcoding, has therefore become popular (Hebert *et al.*, 2003) as a technique that could help resolve taxonomy. DNA barcoding is a well-recognised taxonomic method which makes use of a short standardised genetic marker to enable identification of specific species (Hebert *et al.*, 2003; Packer *et al.,* 2009). DNA barcoding also supports current taxonomic research by providing detailed information for identification of species in nature (Hajibabaei *et al.,* 2007; Hebert *et al.,* 2004; Saunders, 2005; Ward *et al.,* 2005).

1.4 Aims and Objectives

The aim of the study was to classify and identify ichthyoplankton occurring in the nearshore of the south-east coast of South Africa. The main objectives were: (1) to use DNA barcoding combined with traditional morphological techniques to identify fish eggs and larvae and (2) using Tripterygiidae larvae as an example of the merits and demerits of morphological techniques and DNA barcoding. Chapter 1 of this study provides a broad overview of species identification based on different species' concepts and the use of integrative taxonomy to describe the early life stages of fishes. Chapter 2 includes a detailed analysis of morphological identification and DNA barcoding of the ichthyoplankton collected from the south-east coast of South Africa. Difficulties associated with positive identification of the early life stages of fishes are further highlighted in Chapter 3, using an example of Tripterygiidae larvae which were identified to the family level using both the morphological and DNA barcoding approach. Chapter 4, the final chapter, provides an overall discussion with concluding remarks for this study.

CHAPTER TWO

Identification of larval fishes from the south-east coast of South Africa using morphological and DNA barcoding techniques

2.1 Introduction

The South African coastline is characterised by a remarkably diverse marine environment (Jooste *et al.,* 2018), due to its expansive length (~3650 km) and high degree of variability in oceanographic conditions and inshore habitats (Awad *et al.,* 2002; Griffiths *et al.,* 2010). In addition, the high marine species diversity is largely influenced by the fact that South Africa is located at the confluence of three oceans, namely the Indian, Atlantic and Southern Oceans (Griffiths *et al.,* 2010). This also contributes to the high endemism and diversity in the fish species occurring in this coastal region (Turpie *et al.,* 2000).

The majority of marine fishes are broadcast spawners, meaning that eggs and sperm are released into the water column where fertilisation and development into pelagic ichthyoplankton occurs (Leis, 2015). The pelagic larvae are morphologically distinct from the fully-developed juvenile and adult stages, given that they lack morphological structures and features, such as scales and fins (Leis, 1989; Kendall & Matarese, 1994; Webb, 1999). Similarly, fish eggs often lack diagnostic morphological characteristics (Shao *et al.,* 2002). Additionally, larvae frequently present specialised organs for pelagic existence which are lost as development proceeds (Leis & Carson-Ewart, 2000; Leis, 2015). Due to the small size of fish eggs and larvae, the characteristics useful for morphological identification can only be observed under a stereoscopic microscope (Shao *et al.,* 2002). Furthermore, fish larvae undergo rapid morphological and pigmentary changes through their development (Leis, 2015), and show geographic and individual variability (Baldwin, 2013). These factors make morphological identification of fish eggs and larvae very difficult.

Based on studies between 1900 and the early 2000s, taxonomic descriptions and identification of fish eggs and larvae were mostly (and exclusively for South Africa) based on morphological characteristics (Brownell, 1979; Connell, 2010). Early stages in ontogeny are difficult to identify because of a lack of distinguishing characters that allow a distinct separation to species, especially those that are closely related (Turgeon *et al.,* 1999). Overlapping morphological characteristics at the family level also increase chances of species identification errors (Overdyk *et al.*, 2016). Another challenge for egg and larval morphological identification is their fragile state; eggs and larvae are

often badly damaged during sampling and fixation, which may result in a loss of key diagnostic features, such as pigmentation patterns and morphometric characterstics (Valdez-Moreno *et al.*, 2010).

The lack of experienced and trained people in the morphological identification of fish eggs and larvae has been noted as one of the main causes of misidentification of fish larvae (Ko *et al.*, 2013; Puncher *et al.,* 2015). As an example, several researchers from five different laboratories in Taiwan were tasked with identifying fish larvae based on morphological characters (Ko *et al.,* 2013). The results indicated an accuracy of ~80% for identification at the family level, half of that (~40%) was observed at the genus level and only ~10% accuracy was observed at the species level (Ko *et al.,* 2013). This shows how difficult it is to identify larvae to the species level. Advancements in knowledge and technology, such as molecular techniques, have, however, resulted in improved taxonomic identification of fish eggs and larvae (Ko *et al.,* 2013). There is a need therefore, for molecular genetic approaches to support the traditional morphological techniques to resolve larval fish species identification (Zhang & Hanner, 2012; Friedheim, 2016).

Molecular DNA barcoding involves examining a short gene sequence of DNA that can be used for identification, irrespective of the developmental stage or somatic damage, provided that the DNA can be successfully extracted from the sample for PCR amplification (Hubert *et al.,* 2008). A fragment of the cytochrome *c* oxidase subunit I (*COI*) mitochondrial gene region is commonly used for DNA barcoding and the identification of species (Pegg *et al.,* 2006; Valdez-Moreno *et al.*, 2010; Ko *et al.,* 2013). DNA barcoding has emerged as an effective method for identification when morphological identification is problematic (Hebert *et al.,* 2003) and especially when the morphological characters are insufficient to separate two individuals (Becker *et al.,* 2015, Friedheim, 2016). Molecular DNA barcoding has been used broadly in the identification of fish eggs and larvae in southern Africa (Ardura *et al.,* 2016; Steinke *et al.*, 2016; Ockhuis *et al.,* 2017) and South Africa (Steinke *et al.*, 2016; Connell, 2012). There are, however, very few studies (e.g., Ko *et al.,* 2013; Becker *et al.,* 2015; Puncher *et al.,* 2015; Friedheim, 2016; Hulley *et al.,* 2018) that have compared the effectiveness and efficiency of DNA barcoding and morphological identification techniques.

Correct identification of fish eggs and larvae is important, as misidentification can lead to misinterpretation of the biology and ecology of a fish species (Powles & Markle, 1984). This misidentification also includes misunderstanding larval fish biodiversity and incorrect information on the early life history of fishes, such as spawning, hatching, dispersal, and the use of nursery areas

(Hutching *et al.,* 2002). Ultimately, taxonomic uncertainty of early life history stages might result in biased stock assessments and poor fisheries management measures (Armstrong *et al.*, 2001; Fox *et al.*, 2005). Hence, it is necessary to also use molecular genetic techniques to authenticate uncertain morphological larval identifications.

The aim of this study is to increase taxonomic resolution in identifying the early life history stages of fishes occurring in the shallow nearshore of the Eastern Cape, South Africa. The objective of this study is to provide the effective and the best approach in species identification of fish larvae occurring in the shallow nearshore of the Eastern Cape by combining morphological techniques and DNA barcoding. The hypothesis was that, the identification of fish larvae would be more accurate using both traditional morphological characters and DNA barcoding.

2.2 Materials and Methods

2.2.1 Field work

Samples for this study were obtained from collections performed in 2016 during the African Coelacanth Ecosystem Programme (ACEP) project, "Pathways of larval dispersal: the roles of alongshore and cross shore transport". Samples were collected from the shallow nearshore region from Kini Bay (34° 06' 3.32" S, 25° 36' 35.78" E), in the west, to Kenton-on-Sea (33° 37' 7.19" S, 26° 52' 19.19" E), in the east, Eastern Cape, South Africa (Figure 2.1). The sampling stations were situated 900 m – 2.5 km from the shore, at depths ranging from the surface to 50 m. Samples were collected using a bongo net. The bongo net consisted of a pair of plankton nets each with a total length of 2.5 m and canvas cod-ends. Each net had a diameter of 52 cm and a mesh size of 500 µm. A General Oceanics Incorporated flowmeter was suspended at the centre of the opening of each plankton net to calculate the volume of seawater that filtered through the net. The bongo net was hauled through the water at two different depths (surface and bottom) using the research vessel, uKwabelana [National Research Foundation – South African Institute for Aquatic Biodiversity (NRF-SAIAB)], at a speed of ~2 knots for 3 minutes. Plankton samples were stored on site in 250 ml jars and preserved using absolute ethanol (99.9%). A total number of 226 plankton samples were collected during the operation from which 440 fish specimens were sorted, which included 411 fish larvae and 29 eggs.

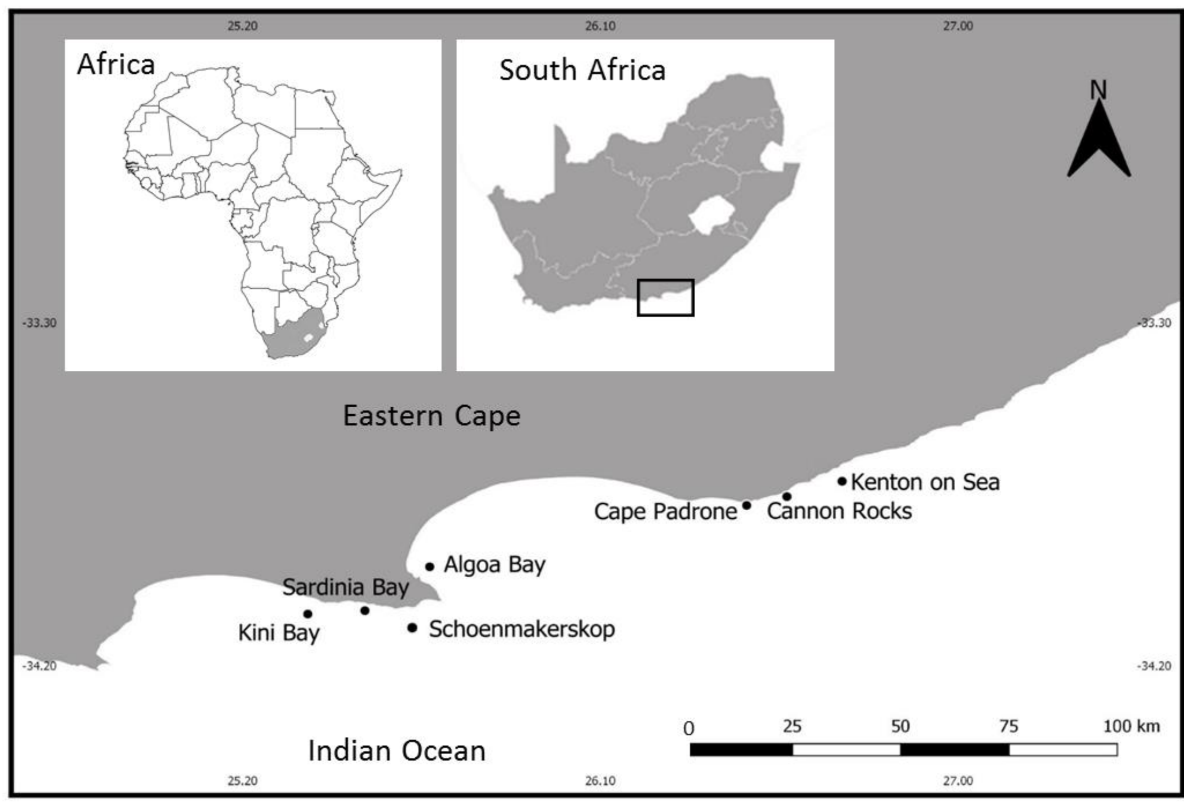

Figure 2.1: Sampling stations (marked with black dots) within the shallow nearshore region of the Eastern Cape along the south-east coast of South Africa.

2.2.2 Laboratory work

Fish larvae were removed from the plankton samples at the NRF-SAIAB laboratory and sorted under a Leica Zoom 2000™ model z45v compound microscope. Unique identification numbers, starting with the acronym FBIP (Foundational Biodiversity Information Programme), were assigned to each larval fish specimen. Specimens were individually preserved in absolute ethanol in a 1.5 ml Eppendorf tube. Photographs of each specimen were captured using an Olympus SZX16 model SZX2-ILLT stereo microscope with a built-in U-TVIXC camera, at the Electron Microscope Unit (EMU) at Rhodes University, Grahamstown, South Africa.

2.2.3 Morphological identification

Morphological identification was done using some of the main characteristics for identification of fish larvae. Those characteristics included body shape, gut shape, pigmentation, presence/lack of a gas bladder and head spination (Leis & Carson-Ewart, 2000). The morphometric measurements (Figure 2.2) of each larval fish specimen were taken to the nearest 0.1 mm using Stream Motion software. Larval specimens were identified to family level following Neira *et al.* (1998) and Leis &

Carson-Ewart (2000); to the genus level using www.fisheggsandlarvae.com (Connell 2007), and where possible, to the species level using Mwaluma *et al.* (2014) and Strydom *et al.* (2015). Eggs were not identified morphologically as they become cloudy immediately after being preserved in ethanol (Murphy & Willis, 1996; Burrows *et al.,* 2018).

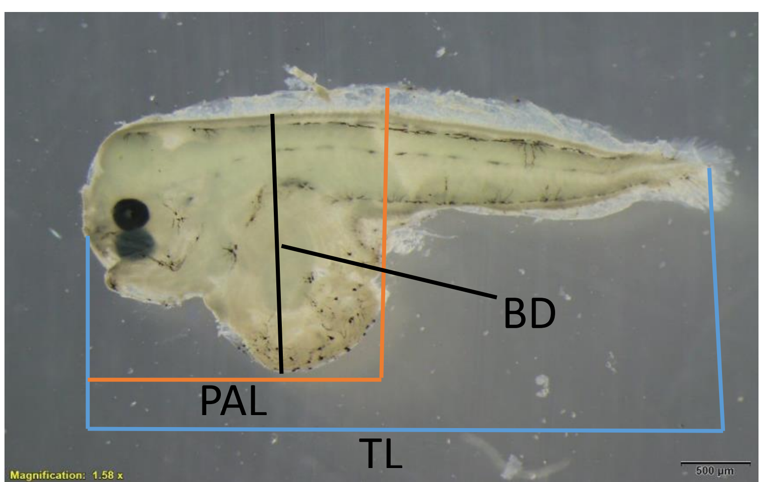

Figure 2.2: An example of the morphometric measurements collected during morphological identification based on photograph of a Soleidae larva. BD = body depth, PAL = preanal length and TL = total body length.

2.2.4 DNA extraction, PCR amplification and sequencing

All genetic analyses were carried out at the NRF-SAIAB Aquatic Genomics Research Platform, Grahamstown, South Africa. DNA was extracted from each specimen using a "salting out" genomic DNA purification protocol (Sunnucks & Hales, 1996). All larval and egg specimens were very small (<5 mm), so the entire specimen was used for the DNA extraction. A NanoDrop 1000 spectrophotometer (Thermo Scientific, USA) was used to measure the concentration and quality of the DNA extracts.

A fragment of the mitochondrial cytochrome *c* oxidase subunit 1 gene (*COI*) was amplified using polymerase chain reaction (PCR), with the forward primer VF_2t_1 and the reverse primer VR_1t_1 (Ivanova *et al.,* 2007). Each PCR amplification was performed in a total volume of 20 µl, with the following constituents: 10 µl of ReadyMix Taq (KapaTaq: Kapa Biosystems, South Africa); 4 µl of molecular graded water; 0.5 µl of each of the forward and reverse primers (10 µM); and 2 or 5 µl of DNA template. When the DNA concentration was less than 100 ng/µl, 5 µl of DNA template was used in the amplification, and 2 µl of DNA was used for extracts above 100 ng/µl. The PCR thermal conditions followed those of Overdyk *et al.* (2016), with alterations in denaturation and annealing temperatures, which were 95°C for 1 minute and 48°C for 40 seconds, respectively.

The degenerate universal metazoan primers dgHCO2198 (10µM) and dgLCO1490 (10µM) (Meyer, 2003) were used to amplify samples that were not successfully amplified using the primers above. The PCR reaction volumes were the same as those described above. The thermal cycling regime comprised an initial denaturing at 95°C for 2 minutes, followed by 35 cycles of denaturing at 95°C for 30 seconds, annealing at 48°C for 40 seconds and an extension at 72°C for 1 minute, followed by the final extension at 72°C for 10 minutes. In all of the amplifications, a negative control (without DNA template) was used to rule out contamination. The PCR products were electrophoresed in a 1% agarose gel stained with ethidium bromide (0.05 µg/ml) and observed under ultraviolet (UV) light to verify the amplification of a product and to check for contamination. A 100 bp molecular ladder was used to verify the size of the amplification band. The presence of a band of the desired size (600–1000 bp) in individual samples and a lack of bands in the negative control signified a successful amplification. The amplified samples were then purified using an Exonuclease 1-Shrimp Alkaline Phosphatase (ExoSap10; Thermofisher Scientific, USA) protocol (Werle *et al.,* 1994). Purified samples were sequenced using the BigDye Version 3.1 Cycle Sequencing Kit (Applied Biosystems, USA), following the manufacturer's instructions. The products were then precipitated using an ethanol-EDTA precipitation method (Sambrook & Russell, 2001). After precipitation, the products were re-suspended in Hi-Di™ Formamide (Applied Biosystems, USA) and analysed on an ABI-Hitachi Genetic Analyser 3500 (Applied Biosystems, USA) at the NRF-SAIAB.

2.2.5 Data Analysis

All sequences were observed in Chromas Lite v.2.1 (Technylesium Pty, Ltd) to verify the quality of sequences. Sequences were edited manually using SeqMan Pro (DNASTAR® Lasergene® 7.2). The Barcode of Life Data System (BOLD) (Ratnasingham & Hebert, 2007; https://www.boldsystems.org/) was used for the comparison of sequences and species identification. The standard nucleotide Basic Local Alignment Search Tool (BLAST) (https://www.ncbi.nlm.nih.gov/) (Altschul *et al.,* 1990) on the National Center for Biotechnology Information (NCBI) database (GenBank) was used as an alternative database tool (Friedheim, 2016) when there was no sequence match in the BOLD database, and for those specimens that were not identified to genus or species level. Some inconsistencies in the results, such as the same sequences being assigned to a different species when comparing the sequences through BLAST searches in NCBI and searches on BOLD, were encountered. GenBank does not conform to the same data principles that were implemented when BOLD was created (Puncher *et al.,* 2015), such as linking the DNA barcoding to a voucher specimen in the form of a photograph. In light of the potential issues raised above, sequences from the NCBI database were therefore used with caution. The

specimen designations to the family, genus and species levels for BOLD comparison were considered positive when the probability values were >97% (Meyer & Paulay, 2005). Sequences identified to the family level or with no results after comparing the generated data on BOLD were categorised as unknowns/ no match.

Accessible and closely matching sequences from the BOLD database were downloaded, and the closest matching sequences from the NCBI database were extracted for those specimens that could not be matched to any species on BOLD. These were aligned, together with the sequences obtained for this study, using MEGA software version 6.0 (Kumar *et al.,* 2013). Kimura two-parameter (Kimura, 1980) sequence divergences were calculated among the sequences generated in this study and the sequences downloaded from BOLD and NCBI. The K2P model was used to estimate the evolutionary distances in fish barcoding studies (Ward *et al.*, 2005; Ward, 2009; Zhang & Hanner, 2012; Ko *et al.*, 2013; Ardura *et al.*, 2016), which enables comparison of the results and sequence divergences to other studies. To identify the unknown (no match) sequences in this study, the divergence thresholds proposed by Hebert *et al.* (2004) were applied. A barcoding study based on marine fish (Ward *et al.,* 2005) established different ranges in sequence divergences that categorise individuals of the same species. A threshold divergence of not more than 2 % characterises members of the same species, and a genetic distance falling beyond 2 % but within 3-8 % was considered as representative of the same genus (Ward *et al.,* 2005). Sequences differentiated by longer branch lengths and K2P distances between 10-20% shows that the unknown larvae belong to a different genus, but possibly the same family. This was done to investigate the levels of variation among the unknown larval samples and the matching sequences from the online data bases. A neighbor-joining tree (Saitou & Nei, 1987) was reconstructed in MEGA to provide a graphic representation of the pattern of divergence between different samples. A maximum likelihood tree was also reconstructed on MEGA to explain the evolutionary history among the sequences of the vouchered species from the BOLD database that are commonly found on the south-east coast and that are representatives of the individuals identified in this study; and the sequences, and the unidentified barcodes. One thousand bootstrap replicates were used to assess the level of support of each clade (Felsenstein, 1985).

2.3 Results

2.3.1 Morphological identification

The preflexion larval development stage dominated (98%) the collection of 411 larvae. For this specific study, a subsample of 177 larval specimens were used for morphological analysis, with the remainder retained for a separate ecological project. Out of the total 177 larval specimens, 100 (56.5%) were identified to the family level based on morphology due to limited identification keys of fish larvae within South African waters. These 100 morphologically-identified specimens represented 18 different fish families (Table 2.1). Specimens from the Sparidae fish family dominated the collection (n=24), followed by the Tripterygiidae (n=19) and then Gobiesocidae (n=15). The remaining fish families were represented by fewer than 10 specimens each (Table 2.1). The remaining 77 (44.5%) larvae of the total 177 specimens, could not be identified since they were damaged and the key features for morphological identification were lost (Figure 2.3). These were barcoded, however.

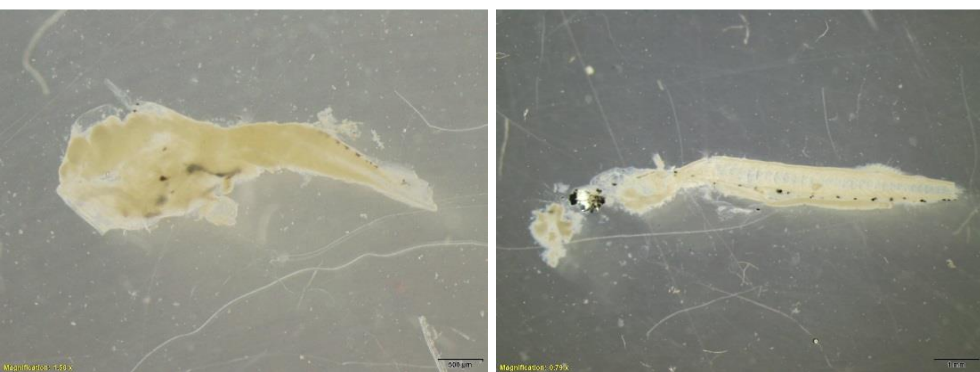

Figure 2.3: Examples of damaged fish larvae that could not be morphologically identified, but were used for DNA barcoding.

Table 2.1: Total number of specimens collected and identified to family level using morphometric characters.

Family	No. of specimens	Body length range (mm)	Body depth range(mm)	Preanal length range (mm)
Ambassidae	2	4.3-6.3	1.1-2.2	2.3-3.7
Blenniidae	2	5.3-5.8	1.3-2.3	2.3-2.0
Callanthiidae	2	6.4-10.9	1.9-2.9	3.3-5.2
Callionymidae	4	2.8-3.6	1.0-1.4	2.1-2.4
Centrolophidae	3	4.3-7.1	0.9-1.9	2.4-3.9
Cynoglossidae	3	6.9-1.3	1.1-3.5	2.5-5.1
Gobiesocidae	15	2.7-1.5	0.7-1.7	1.2-4.3
Gobiidae	5	3.1-4.6	0.7-1.1	1.5-2.0
Phycidae	1	8.6	2.8	5.2
Samaridae	3	2.9-4.3	0.6-1.4	0.8-1.5
Scorpaenidae	2	3.3-7.5	0.9-3.2	1.8-5.1
Soleidae	8	3.3-11.1	0.7-5.1	1.5-5.2
Sparidae	24	3.4-7.1	0.9-2.9	1.2-1.9
Syngnathidae	4	11.2-17.9	0.9-1.3	5.2-7.9
Terapontidae	1	6.9	1.7	2.9
Tetraodontidae	1	5.8	1.9	4.1
Triglidae	1	7.5	2.0	3.8
Tripterygiidae	19	3.4-1.2	0.7-3.0	1.5-5.0

2.3.2 Molecular identification

A total of 177 larvae and 17 fish eggs were used for DNA barcoding. Of the total 194 larval and egg specimens from which DNA was extracted, 116 larval specimens and 15 eggs failed to amplify by PCR with both primer sets. *COI* sequences were thus obtained for 61 larval specimens and two fish eggs. Twelve fish families, 16 genera and 18 different species were identified through DNA barcoding (Table 2.2). Ten larval sequences could not be matched to species level with any sequence on the BOLD or NCBI online reference libraries. Twenty-eight of the total barcoded specimens are excluded from this chapter, but are discussed further in Chapter 3.

Six Sparidae and five Callionymidae sequences were identified to species level with a similarity match of >98% on the BOLD database. Only one sequence of each of the Carangidae, Cheilodactylidae, Dussumieriidae, Engraulidae, Gobiidae and Haemulidae were identified to species level.

Table 2.2: Sequence match for successfully amplified egg and larval specimens. The range of similarity match (%) for each species is showed for BOLD and NCBI databases. Number of specimens indicated in brackets.

Family	DNA barcoding ID	BOLD Similarity (%)	NCBI Similarity (%)
Blenniidae*	Blenniidae sp	nm (1)	96.89
Blenniidae	*Parablennius pilicornis*	98 - 99.5	-
Blenniidae*	*Parablennius cornutus*	nm (1)	93.99
Callionymidae	*Paracallionymus costatus*	99.5 – 99.7	-
Callionymidae	*Callionymus marleyi*	98.6 – 99.1	-
Callionymidae*	*Callionymus marleyi*	nm (1)	94.29
Carangidae	*Trachurus trachurus*	98.4	-
Centrolophidae*	*Hyperoglyphe japonica*	nm (2)	80.50 – 82.26
Cheilodactylidae	*Cheilodactylus pixi*	99.4	-
Dussumieriidae	*Etrumeus whiteheadi*	99.8	-
Engraulidae	*Engraulis encrasicolus*	99.4	-
Gobiidae	*Caffrogobius natalensis*	99.7	-
Gobiidae*	nm	98 (3)	-
Haemulidae	*Pomadasys striatus*	100	-
Samaridae	*Samariscus triocellatus*	98.7 – 99.5	-
Soleidae	*Austroglossus pectoralis*	97.1 - 97.3	-
Sparidae*	*Diplodus sargus*	nm (2)	94.29 – 96.99
Sparidae	*Diplodus capensis*	99.6 - 100	-
Sparidae	*Pagellus natalensis*	98.6	-
Sparidae	*Porcostoma dentate*	100	-
Sparidae	*Pterogymnus laniarius*	100	-

*indicate sequences with no match and family level identification. Numerical values in brackets under BOLD similarity values column indicate number of sequences.

Of the ten sequences (as shown in Table 2.2) from different larval specimens where no match to the species level could be found, six were severely damaged and could not be identified using morphology. Four were identified as belonging to the families Soleidae, Gobiidae, Callionymidae and Blenniidae (Figure 2.4A, B, C, D, respectively), using traditional morphological techniques. Also, the GenBank close matches were below the probability percentage value adopted for this study.

Figure 2.4: Photographs of specimens with successful sequences that failed to find matches on the BOLD online platform and were identified through morphological characteristics. (A) Soleidae (FBIP0143); (B) Gobiidae (FBIP0087); (C) Callionymidae (FBIP0365) and (D) Blenniidae (FBIP0398).

The sequences that were obtained in this study were used in combination with the barcode data downloaded from BOLD and NCBI to construct a neighbor-joining (NJ) tree (Figure 2.5) for a graphic representation of the distance relationships among the sequences. The high bootstrap support, indicated on the nodes, supports the species identification for those larval specimens that could be identified through DNA barcoding.

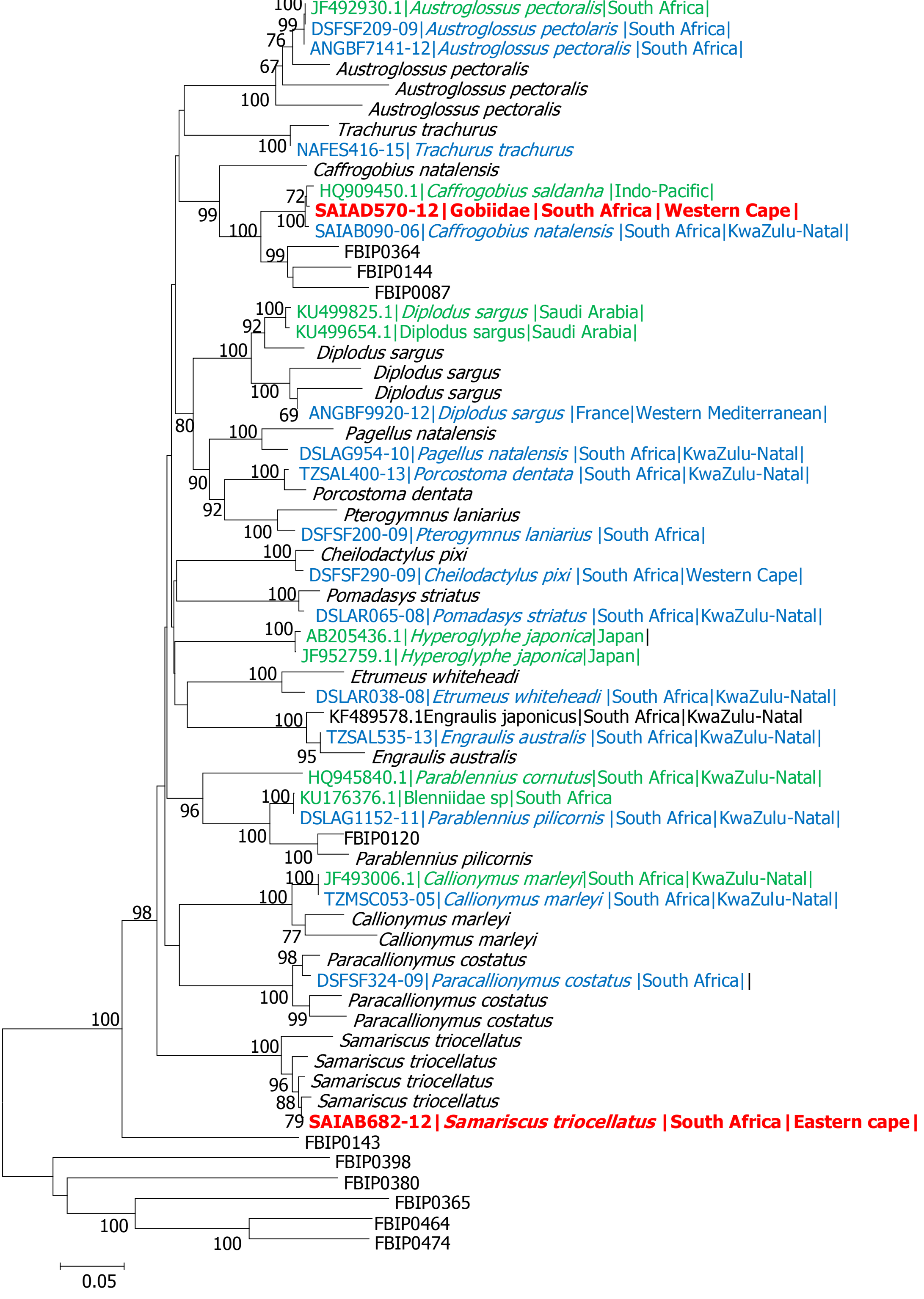

Figure 2.5: Neighbor-joining tree of the *COI* sequences showing the divergence patterns and distance relationships between sequences. Bootstrap values are represented by the number at the nodes. Black sample names represent the species names for larval samples of the current study. Red sample names represent unpublished sequence data from BOLD. Blue sample names are extracted from public data on BOLD. Green samples are the probability matches to from NCBI. Samples without a closest match on BOLD and NCBI are presented as FBIP number.

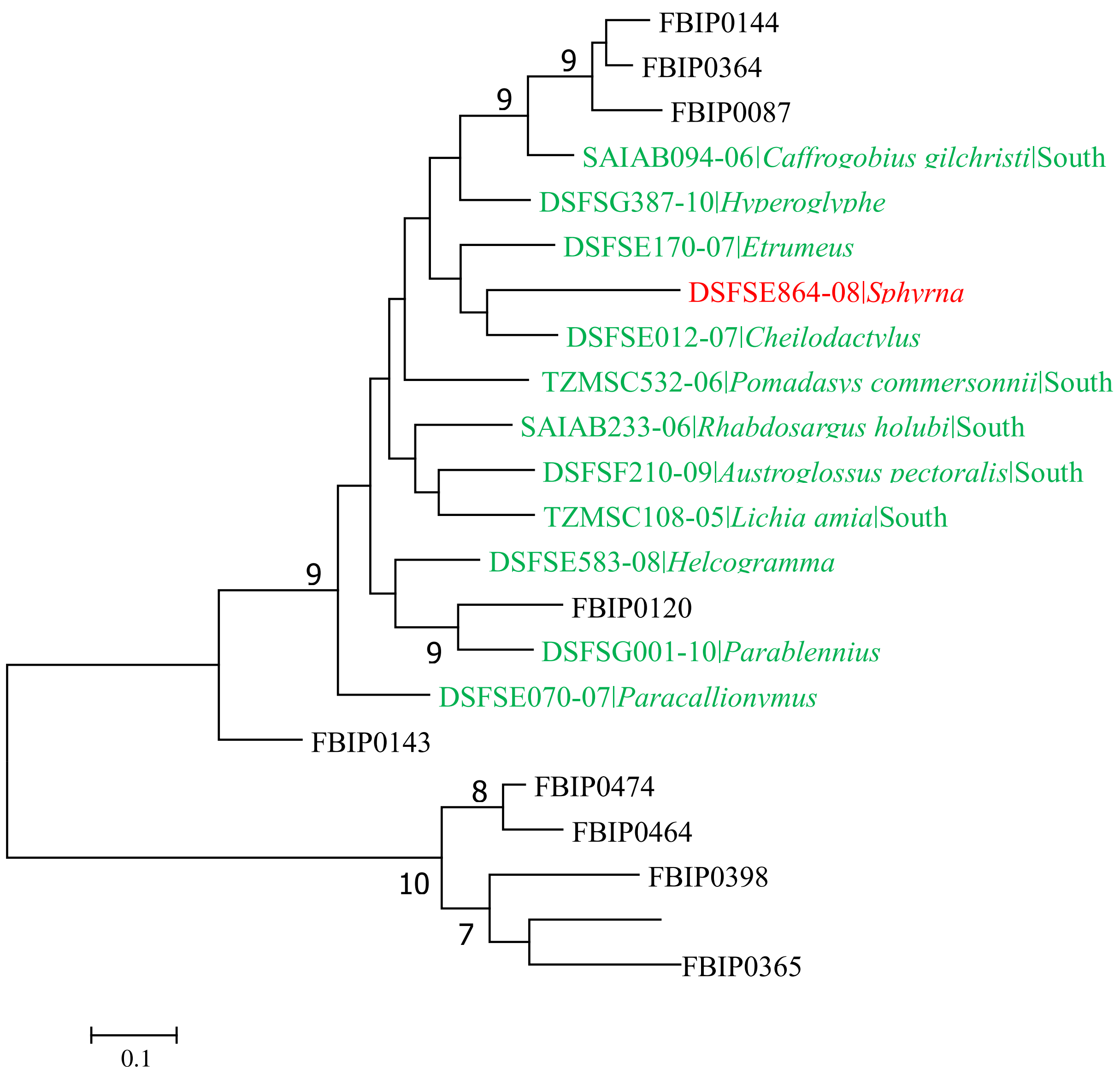

Figure 2.6: Maximum likelihood phylogenetic tree of the *COI* sequences showing the divergence patterns and distance relationships between sequences. Bootstrap values are represented by the number at the nodes. Black sample names represent the unidentified specimens of the current study. Red sample names represent a bony fish sequence data from BOLD. Green sample names represent the sequences of fish species found in South Africa.

All of the sequences from the unmatched larval specimens remained unidentified to the species level, probably because of the incomplete reference database. Therefore, the GenBank database was used to download sequences that were similar to the unknowns (Table 2.3).

Table 2.3: Sequences without closest match, together with those identified to family level on BOLD. Column two indicates the similarity values from GenBank.

Unknown Sequences	GenBank (%)	Species	Accession numbers	Authors
FBIP0087	92.82	*Caffrogobius saldanha*	HQ909450.1	Thacker *et al.* 2011
FBIP0120	96.89	Blennidae *sp.*	KU176376.1	Steinke *et al.* 2016
FBIP0144	92.20	*Caffrogobius saldanha*	HQ909450.1	Thacker *et al.* 2011
FBIP0143	82.26	*Hypereglyphe japonica*	AB205436.1	Yanagimoto & Kobayashi, 2012
FBIP0464	94.29	*Diplodus sargus*	KU499825.1	Rabaoui *et al.*2016
FBIP0364	93.07	*Caffrogobius saldanha*	HQ909450.1	Thacker *et al.* 2011
FBIP0380	80.50	*Hypereglyphe japonica*	AB205436.1	Yanagimoto & Kobayashi, 2012
FBIP0398	93.99	*Parablennius cornutus*	HQ945840.1	Steinke *et al.* 2016
FBIP0474	96.99	*Diplodus sargus*	KU499654.1	Rabaoui *et al.*2016
FBIP0365	94.29	*Callionymus marleyi*	JF493006.1	Steinke *et al.* 2016

Genetic distances were calculated among the unknown larval sequences and the closest matches from NCBI. A K2P distance range of 0.05 to 0.63 was revealed between the unknown sequences and all NCBI close matches (Table 2.4). The lowest genetic distance of 0.05 was between one of the unidentified larval specimens (FBIP0120) and the combtooth blenny *Parablennius pilicornis* (Blenniidae). A 0.24 K2P distance was revealed between FBIP0143 and a *Hypereglyphe japonica* (Centrolophidae) sequence.

Table 2.4: K2P genetic distances among the close matches and the sequences categorised as unknowns. The lowest genetic distance between FBIP0120 and Blenniidae *sp.*is highlighted in bold.

	Sequences	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	*Hypereglyphe japonica* (AB205436.1)												
2	*Parablennius cornutus* (HQ945840.1)	0.22											
3	*Callionymus marleyi* (JF493006.1)	0.22	0.22										
4	Blenniidae *sp*. (KU176376.1)	0.17	0.15	0.21									
5	*Diplodus sargus* (KU499654.1)	0.18	0.18	0.22	0.19								
6	*Diplodus sargus* (KU499825.1)	0.18	0.18	0.22	0.19	0.01							
7	FBIP0120	0.22	0.19	0.25	**0.05**	0.22	0.22						
8	FBIP0143	0.24	0.28	0.27	0.26	0.27	0.27	0.25					
9	FBIP0464	0.66	0.64	0.64	0.63	0.63	0.64	0.62	0.63				
10	FBIP0380	0.65	0.65	0.63	0.64	0.64	0.64	0.63	0.62	0.25			
11	FBIP0398	0.64	0.64	0.65	0.63	0.64	0.64	0.63	0.62	0.24	0.27		
12	FBIP0474	0.65	0.64	0.63	0.63	0.64	0.64	0.62	0.62	0.06	0.24	0.22	
13	FBIP0365	0.64	0.64	0.61	0.63	0.63	0.63	0.62	0.61	0.25	0.24	0.25	0.23

Between FBIP0364 and the Gobiidae sequences, there was a 0.09 distance; also FBIP0144 and FBIP0087 showed genetic distances of 0.12 and 0.13, respectively, to *Caffrogobius saldanha* (Gobiidae) (Table 2.5). The remaining five sequences showed a genetic distance range of 0.63 to 0.64 compared to all other sequences.

DNA barcoding failed to resolve the taxonomy of three specimens of the Gobiidae and these were only identified to the family level (Table 2.2). The K2P genetic distances among sequence data generated from this study and from both BOLD (*Caffrogobius natalensis*) and NCBI (*Caffrogobius saldanha*) ranged from 0.08 to 0.23 (Table 2.5). The genetic distance values between the three reference sequences and the larval fish samples from this study ranged from 0.09 and 0.17.

Table 2.5: K2P values between Gobiidae species sourced from BOLD and NCBI, together with the larval sequences derived from this study. The lowest and highest genetic distances among sequences are highlighted in bold.

	Sequences/Species	**1**	**2**	**3**	**4**	**5**	**6**
1	*Caffrogobius saldanha* (HQ909450.1)						
2	FBIP0364	0.10					
3	FBIP0140	**0.17**	0.17				
4	FBIP0144	0.12	0.08	0.18			
5	FBIP0087	0.14	0.12	0.23	0.11		
6	Gobiidae (SAIAB570-12)	0.01	**0.09**	0.15	0.11	0.13	
7	*Caffrogobius natalensis* (SAIAB090-06)	0.01	0.10	0.15	0.11	0.13	0.01

Table 2.6: Photographs and identifications of individual specimens, based on traditional morphology and DNA barcoding techniques.

Sample ID	Specimen photo ID	Morphological ID	DNA barcoding ID
FBIP0375		Blenniidae	Failed PCR
FBIP0118		Blenniidae	Engraulidae (*Engraulis australis*)
FBIP0398		Blenniidae	Successful sequence, No match on databases
FBIP0417		Blenniidae	Sparidae (*Diplodus sargus*)
FBIP0431		Callionymidae	Callionymidae (*Paracallionymus costatus*)

FBIP0432		Callionymidae	Callionymidae (*Paracallionymus costatus*)
FBIP0449		Callionymidae	Callionymidae (*Callionymus marleyi*)
FBIP0516		Callionymidae	Callionymidae (*Paracallionymus costatus*)
FBIP0518		Callionymidae	Callionymidae (*Callionymus marleyi*)
FBIP0365		Callionymidae	Successful sequence, No match on databases
FBIP0431		Callionymidae	Callionymidae (*Paracallionymus costatus*)
FBIP0280		Carangidae	Carangidae (*Trachurus trachurus*)

FBIP0433		Centrolophidae	Failed PCR
FBIP0109		Cynoglossidae	Failed PCR
FBIP0139		Cynoglossidae	Soleidae (*Austroglossus pectoralis*)
FBIP0234		Cynoglossidae	Failed PCR
FBIP0115		Engraulidae	Blenniidae (*Parablennius pilicornis*)
FBIP0458		Gobiesocidae	Cheilodactylidae (*Cheilodactylus pixi*)

FBIP0227		Gobiesocidae	Failed PCR
FBIP0087		Gobiidae	Gobiidae
FBIP0082		Not identified	Dussumieriidae (*Etrumeus whiteheadi*)
FBIP0405		Not identified	Haemulidae (*Pomadasys striatus*)
FBIP0115		Not identified (suspected to be Engraulidae spp.)	Blenniidae (*Parablennius pilicornis*)
FBIP0380		Not identified	Successful sequence, No match on databases

FBIP0480		Not identified	Sparidae (*Porcostoma dentata*)
FBIP0464		Not identified	Successful sequence, No match on databases
FBIP0474		Not identified	Successful sequence, No match on databases
FBIP0144		Not identified (suspected to be a Gobiidae spp.)	Gobiidae
FBIP0273		Phycidae	Failed PCR
FBIP0283		Scorpaenidae	Failed PCR

FBIP0078		Soleidae	Samaridae (*Samariscus triocellatus*)
FBIP0098		Soleidae (*Austroglosssus*)	Soleidae (*Austroglossus pectoralis*)
FBIP0103		Soleidae	Samaridae (*Samariscus triocellatus*)
FBIP0111		Soleidae	Soleidae (*Austroglossus pectoralis*)
FBIP0143		Soleidae	Successful sequence, No match on databases
FBIP0214		Soleidae (*Solea turbynei*)	Failed PCR
FBIP0413		Sparidae	Sparidae (*Pterogymnus laniarius*)

FBIP0267		Sparidae	Failed at PCR
FBIP0222		Sparidae	Failed at PCR
FBIP0338		Syngnathidae	Failed PCR
FBIP0491		Tetraodontiformes	Failed PCR
FBIP0278		Triglidae	Failed PCR
FBIP0120		Tripterygiidae	Successful sequence, No match on databases

2.3 Discussion

The results from this study revealed that DNA barcoding can be used as a tool to identify larval fish specimens to the species level. Almost 80% of the sequences generated successfully in this study were identified to species level through *COI* barcodes. This falls within the same region as several studies that have been conducted using the DNA barcoding approach to investigate ichthyoplankton assemblages in various marine habitats worldwide (Pegg *et al.,* 2006; Valdez-Moreno *et al.,* 2010; Ko *et al.,* 2013; Friedheim, 2016; Ardura *et al.,* 2016). Recently, Steinke *et al.* (2016) used this approach to link adults and immature individuals (both larvae and juveniles) of South African marine fishes along the KwaZulu-Natal coast. The present study, however, provides the first report of the taxonomic composition of ichthyoplankton assemblages from the shallow nearshore waters along the Eastern Cape coast, South Africa, using DNA barcoding.

DNA-established identification cannot be done without the contribution and expertise of taxonomists who identify type specimens from which the reference sequences can be attained (Hebert *et al.,* 2003; Hebert & Gregory 2005; Coyne & Orr, 2004). Identifying a larval specimen using morphological characters can be inconsistent among taxonomists (Ko *et al.,* 2013; Neira *et al.*, 2015; Azmir *et al.,* 2017). This inconsistency can result from different levels of expertise in larval fish identification (Ko *et al.,* 2013). A prime example of misidentification includes the commercially important Atlantic Bluefin tuna (Scomberidae: *Thunnus thynnus*), where larvae collected from the Mediterranean Sea were misidentified by the crew of these larval surveys (Puncher *et al.,* 2015). Although phenotypic characteristics can be misleading at times, experience is imperative in taxonomy, especially when dealing with specimens that are morphologically similar (Zhang & Hanner, 2011; Friedheim, 2016). Morphological misidentification can also be caused by limited available references for identification of larvae (Azmir *et al.,* 2017), and this is particularly true in South Africa. The preflexion larval developmental stage dominated (98%) the larval fish assemblage in the present study. Misidentification of preflexion larvae to the species level is highly likely because preflexion larvae present ontogenetic features in body shape that may lead to a single specimen being assigned to different groups (Leis & Carson-Ewart, 2000). Secondly, preflexion larvae often lack obvious diagnostic characteristics (Ko *et al.*, 2013), such as spines and pigmentation, used for morphological identification. As such, morphological misidentifications of these very early stage larvae have been documented in various fish families (Leis & Carson-Ewart, 2000).

The preflexion larval stages of benthic reef-associated species of the Blenniidae and Gobiidae families also dominated the catch in the present study. Similar catches of larval fishes in coastal

systems, comparable to the one targeted in the current study, were observed in the temperate nearshore of Portugal, where Blenniidae and Gobiidae encompassed 76% of larvae caught (Beladae *et al.*, 2006). The Blennioid families are generally difficult to identify using morphological characteristics (Ko *et al.*, 2013), due to the small size and the cryptic forms of the adults, with the early life history stages understudied (Leis, 2015).

Preflexion larvae from the Soleidae, Cynoglossidae and Engraulidae were also positively identified through morphology in the present study. The above-mentioned families also contribute to the nearshore larval collections in the Algoa Bay region and are found in abundance during the summer months (Beckley, 1986). Soles are pelagic spawners, some with potentially complex spawning migrations and early life history strategies, but the ecology and description of early life history stages is not fully elucidated for southern African species (Strydom *et al.*, 2015).

Morphological identification in the present study misidentified a preflexion Soleidae larva, which was then confirmed as a Samaridae larva through DNA barcoding. The Soleidae (true soles) and Samaridae (crested flounders) families belong to the same order, the Pleuronectiformes. The orders Pleuronectiformes and Blenniiformes each consist of species that are similar in appearance, to the point that the boundaries between families within each remain unresolved; however, these groups are said to be monophyletic (Berendzen & Dimmick, 2002; Lin & Hastings, 2013). Incomplete taxon sampling can also be the reason why blenniiform larvae could not be identified to the lowest taxonomic level through DNA barcoding (Leis & Carson-Ewart, 2000). Furthermore, difficulties in differentiating the larvae of the families through morphology alone were due to similarities such as the protruding gut (Leis & Carson-Ewart, 2000; Sasaki & Uyeda, 2002), moderate pigmentation, and similar body shape in their preflexion larval stages (Strydom *et al.,* 2015).

Samariscus triocellatus, identified by DNA barcoding in this study, is the only reef-dwelling species in the Samaridae family, but it is also found in shallow waters (Alfonso, 2001). The adults and juveniles of soleids and samarids are known to occur in coastal habitats of southern African waters (Smith & Heemstra, 2003). The ecology and description of the early life history stages are, however, not explained fully (Strydom *et al.,* 2015). The larval development has been described for only nine of the 56 species of the Pleuronectiformes from southern African waters (Wood, 2000; Thompson *et al.,* 2007), and none from the Samaridae family.

From the current study, larval fish specimens were assigned to 18 families through the traditional morphological technique. The number of families identified successfully was, however, reduced to 11 after comparing the DNA sequences on the databases, thereby confirming five families (Blenniidae, Callionymidae, Gobiidae, Soleidae and Sparidae) that were determined morphologically. Azmir *et al.* (2017) identified 177 larval fish samples from Malaysia, based on morphological characteristics as representatives of 19 different families; however, after DNA barcoding, only 11 families were identified.

Only one specimen from the Soleidae was positively identified to genus (*Austroglossus*), but not to species level, using morphology. DNA barcoding of this *Austroglossus* specimen led to the species level identification of individual FBIP0098 (Table 2.6). Low identification rates of early life history stage of fishes using morphological identification to the genus level were also reported by Shao *et al.* (2002) and Azmir *et al.* (2017). This common trend shows that specimens identified through morphology can potentially be matched to various species. DNA barcoding can, however, be effective enough to identify eggs or larvae to the species level. Ko *et al.* (2013) also revealed that only two out of 100 specimens that were identified to species level through the morphological approach were validated as correct by DNA barcoding. According to Ko *et al.* (2013), the accuracy of larval fish specimens identified to the species level using only morphological characters is far lower than those identified using DNA barcoding.

The Sparidae fish family is represented by 42 species in southern Africa, of which 50% are endemic (Smith, 1965). Sparid larvae mostly dominate the catch in bays on the south-eastern coast of South Africa (Strydom, 2008). The larval descriptions of most of these species are, however, not available, including those for *Porcostoma dentata* (Gilchrist & Thompson 1908) and *Pterogymnus laniarius* (Valenciennes 1830), which were positively identified by DNA barcoding in this study. Larvae of the sparid, *Diplodus sargus,* which were positively identified in this study, are commonly found as adults in habitats like the one under investigation (Beckley 1986). *Diplodus sargus* is a species that aggregates in groups (about eight to fifteen individuals) of both males and females in the near-surface shallow water to simultaneously release sperm and eggs in the water column (Richardson *et al.,* 2011). Although spawning in sparids occurs inshore (Connell, 2012), preflexion larvae are commonly found in the shallow nearshore (Pattrick & Strydom, 2008). Juveniles of this species maintain a very nearshore distribution as they are commonly found in tidal pools and also in the surf zone adjacent to sandy beaches (Lasiak, 1981).

The inability to positively identify seven unknown specimens using DNA barcoding was a result of sequences from those species being absent on the BOLD online database. Ko *et al.* (2013) also noted that the *COI* database on BOLD is incomplete, especially for the families that are

difficult to identify and of those that are not of economic importance. Overall, positive identifications obtained from DNA barcoding showed a high probability percentage (>97%) on the BOLD database, suggesting that the specimens identified belong to that exact species or genus on the database.

The incomplete taxonomic coverage in the DNA reference databases (Dettai *et al.,* 2011) remains a major impediment for accurately identifying fish larvae to species level (Ardura *et al.,* 2016). Three of the ten specimens with no sequence match were identified morphologically and included species from the Soleidae, Callionymidae and Blenniidae families. Those three specimens could not be validated through the *COI* sequences; however, these individuals are morphologically similar to the ones positively identified to species level as *Austroglossus pectoralis, Callionymidae marleyi* and *Parablennius pilicornis* by DNA barcoding. One of the major problems for DNA barcoding with *COI* is that it is unable to differentiate species if there is little variation between them (Mier *et al.,* 2006). Basically, in order for DNA barcoding to be able to identify the species, the interspecific sequence variation has to be higher than the intraspecific variation in the gene region (Ward *et al.,* 2005; Meyer & Paulay, 2005). According to Avise (2000) divergences between similar species are normally rarely greater 2% and most are less than 1%.

Investigations based on the K2P values revealed that FBIP0120 is a representative of the Blennidae, with 0.05 genetic distance, which confirms the morphological identification. Also, FBIP0364 had a genetic divergence of 0.09 to the Gobiidae sequence, which falls within Avise's (2000) 2% threshold reported above, indicating that these are members of the same species. Additionally, the other two goby sequences (FBIP0144 and FBIP0087) had a genetic distance range of 0.12 to 0.13 from the reference data (Table 2.5), which is not different from other studies. The above-mentioned genetic distances between these sequences is not far from the patterns detected in fish barcoding studies. For example, the K2P genetic values among individuals of the same species of Australian fishes ranged from 0.39 – 14.08% (Ward *et al.*, 2005) and the within-species range for Argentine marine fishes was 0 – 12.83% (Mabragana *et al.*, 2011). The unknown larval specimens from southern Africa with genetic divergence that falls within the 2% threshold were characterised as the conspecifics (Ockhuis *et al.*, 2017)

Unknown sample number FBIP0143 with a genetic distance of 0.24 from the downloaded reference data shows that the specimen is more likely to be a representative of the Centrolophidae family. Nonetheless, larger genetic divergence values were revealed among sequences of the unknown specimens and the reference data. For the sequences to show divergence, it means that the individual specimens belong to different categories in higher taxonomic levels, for example,

family. Hebert *et al.* (2003) reported that divergences increase with taxonomic level, meaning the generic level will be lower than familial or ordinal level. Also, sequence divergences are considerable larger among species than within species (Hebert *et al.* 2004).

Genetically, each individual fish specimen or species is unique, meaning no genomic data are identical to one another (Hebert *et al.,* 2003; Friedheim, 2016) which is why species level identification is made easy by DNA barcoding. In this study, morphological identification of larval specimens FBIP0115; FBIP0417 and FBIP0118 (Figure 2.5) contradicts the DNA barcoding results. There are many reasons for the above-mentioned results: most of DNA barcodes with which the sequence data from the larval specimens were compared were generated from field identified specimens (Swartz *et al.,* 2008). Also, some of the specimens were not removed from their natural habitats, but a photograph and a fin clip, or small tissue sample was taken from the individuals for DNA barcoding (Jaafar *et al.,* 2012, Weigt *et al.,* 2012). Additionally, a very small amount of DNA is used in barcoding; therefore, there are high chances of contamination (Karaiskou *et al.,* 2007; Friedheim, 2016; Bingpeng *et al.,* 2018). Even though, ideally, a contaminated sample would not pass PCR stage and the sequence will not be accurate enough to get positive identification, when working with large number of samples, mixing of labelling could cause a mismatch in species identification. Mismatch between the morphological and DNA barcoding has been documented in larval fishes, where the species level identification between NCBI and BOLD would show contradictory information about a particular species name, but similar geographic region at which it is found, or the other way around (Friedheim, 2016). This highlights the need for a stronger rulebook for sequences that are deposited on the online databases to decrease inconsistencies associated with human error.

The mismatch between the morphological and DNA barcoding approach once more indicates the gaps that still need be filled in terms of additional sequences to be added and robustly documented on the online databases. Another possible drawback in the DNA barcoding procedure, especially with the *COI* gene, is the amplification of the imperfect copies of the functional gene, referred to as pseudogenes (Brownell *et al.*, 1979; Smith *et al.*, 2007), which obstructs the positive identification of specimens (Song *et al.,* 2008; Buhay, 2009). Rare pseudogenes are mostly identified in fishes because they contain the stop codons (Venkatesh *et al.,* 2006). BOLD has a function responsible for identifying and flagging sequences that contain such pseudogenes when uploaded unintentionally (Becker *et al.,* 2010). Nevertheless, species identification cannot always be resolved by only one technique. Hence, it is once again recommended that, especially for ichthyoplankton, both DNA barcoding and morphological techniques should be used for species identification to avoid biases and errors in the process (Ko *et al.,* 2013; Puncher *et al.,* 2015). A broad database of *COI* sequences (Meyer & Paulay, 2005),

supported with identified voucher specimens, reference material such as photographs, good drawings and corresponding adult vouchers has the potential to further change the circumstances (Steinke *et al.,* 2016) and resolve larval fish identification.

DNA barcoding of fish eggs and larvae has provided positive identification and accurately confirmed the presence of the early life stages of particular species within a geographical coastal area. The adults that are known to occur in this geographical area were well represented by the identified larval species, for example the sparids, gobiids and blenniids (Beckley, 1986). Information from this study could be used by biologists to further define spawning areas and to study fish population dynamics (Burrows *et al.,* 2018).

In conclusion, the present study revealed that positive and successful identification of fish larvae using the traditional morphological approach was largely possible only to family level. Nevertheless, DNA barcoding proved to be a fast and reliable method for species level identification, but this approach also has methodological limitations that still need to be addressed for larval barcoding (Ko *et al.,* 2013; Harada *et al.,* 2015). For example, small-bodied benthic species, such as goby species (Whitfield, 1998), which are an important component of the fish biomass (Jennings *et al.,* 2002) and the food web, and play an important role in estuarine ecology, were identified only to family level. A comprehensive reference library is especially warranted for cryptic and ecologically important species which may not be important from a commercial or recreational fishery perspective. For example, some of these species, although not iconic, are essential to maintain the functioning and connectivity of coastal ecosystems (Strydom, 2003; 2008).

CHAPTER THREE

The use of DNA barcoding for the identification of the early life history stages of a cryptic benthic fish species

3.1 Introduction

Taxonomy is the study of discovering, describing, identifying and categorising organisms (Winston, 1999). Taxonomic identification of marine fish eggs and larvae is often complex due to the paucity of reference material available (Shao *et al.,* 2002; Ko *et al.,* 2013; Puncher *et al,* 2015), but also because of the biology and mechanisms of larval growth. Larval fishes often bear little resemblance to the adults (Leis, 1991), and both (fish eggs and larvae) are not just minute in size, they also have limited morphological development (Richardson *et al.,* 2006; Webb *et al.,* 2006) all making it hard to link the early developmental stages to the adults. The obvious morphological differences between adults and early life stages include the lack of morphological features that are found in adults, such as scales, pigment and even fins (Wilbur, 1980; Leis & Carson-Ewart, 2000; Leis, 2015). From a taxonomic point of view, positive identification of the early life history stages of fishes is challenging, due to overlapping meristic measurements among species, shared morphological similarities among different taxa, change in morphological characteristics during early development, and a lack of experienced identifiers (Richardson *et al.,* 2007; Ko *et al.,* 2013; Azmir *et al.,* 2017; John *et al.,* 2018; Sachithanandam & Mohan, 2018).

In South Africa, it is estimated that about 25% of all endemic fish species are undescribed, mainly because of the size of the fish (von der Heyden, 2011). Small-bodied (<10 cm) marine species are cryptic and inhabit regions that are difficult to sample (Zapata & Robertsons, 2007). Families such as the Blennidae, Gobiidae and Tripterygiidae (with species that are endemic to South Africa) are among the most abundant and speciose families found in temperate coral reefs (Paulin & Roberts, 1992; Randall *et al.,* 1997). Longenecker & Langston (2005) acknowledge that the early life history developmental stages of small, benthic fishes, especially the Tripterygiidae species, are mostly unidentified and undescribed. Also, it is likely that these fishes have remained undescribed because they are not significant food or ornamental species (Mace, 2004; Victor, 2013) or because of limited research interest in small-bodied species (Brook *et al.,* 2006). From an ecological perspective, small-bodied fishes from the nearshore environment (shore to 50 m depth) are most likely to be undescribed (von der Heyden, 2011) because small fishes mature at younger ages, have a shorter lifespan (Marrett *et al.*, 1991; Payne & Punt, 1995; Moranta *et al.*, 2007) and have benthic spawning as a reproductive strategy (a characteristic that involves parental care of eggs) so as to enhance the

survival of the offspring (Munday & Jones, 1998; Hendry *et al.,* 2001). Evidence from local guide books of South Africa's marine fauna (Smith & Heemstra, 2003; Heemstra & Heemstra, 2004; Zsilavecz, 2005) suggests that there is a high species richness and endemism. A hindrance in marine and coastal taxonomic progress is, however, blatantly obvious in South Africa, due to limited staff and project funding (Klopper *et al.*, 2002). Additionally, small-sized fishes that occupy shallow areas (especially those in the subtropical and tropical waters on the east coast) are subjected to under-sampling (von der Heyden, 2011).

Notwithstanding the pioneering work of Allan Connell in raising and identifying fish eggs and larvae from South African waters (fisheggsandlarvae.com), larval descriptions for almost all coastal fish species in South Africa are still lacking (Swartz *et al.*, 2008). Larval fish taxonomy has however, increased exponentially worldwide in recent years, fuelled by the increasing movement towards molecular techniques, including DNA barcoding, to identify species (Friedheim, 2016; Azmir *et al.,* 2017; Bingpeng *et al.,* 2018; Hulley *et al.,* 2018; Overdyk *et al.,* 2016). This research remains, however, in its infancy in South Africa.

Despite the obvious gaps in knowledge of cryptic, small benthic fishes, current information suggests that these communities of fishes have the ability to make a significant and substantial contribution to reef ecosystems (Ackerman & Bellwood, 2002). The contribution of these communities to trophic pathways is largely defined by their consumption of micro-crustaceans and their status as prey items for larger reef fishes (Depczynski & Bellwood, 2003). In addition, these small fishes have a role in utilising detritus as a major dietary component in trophic pathways (Wilson, 2000).

The Blennioid fishes are a suborder of the perciform teleosts, which, as adults, are small, show cryptic behaviour with sedentary habits (Thompson, 1983), and are difficult to identify (Syms, 1995). Blennioids comprise six families, 151 genera and 883 species of tropical and warm temperate fishes worldwide (Hastings & Springer, 2009). Presently, the six families that are recognised include the Clinidae (clinid kelpfish), Labrisomidae (labrisomid kelpfish), Chaenopsidae (tube blennies), Tripterygiidae (triplefin blennies), Blenniidae (combtooth blennies) and Dactyloscopidae (sand stargazers) (Springer, 1998). The tripterygiids and blenniids are distributed worldwide, while the Clinidae are restricted to the temperate regions and the remaining three families are neotropical (Stepien *et al.,* 1993; Hastings, 2009). The taxonomic relationships of the tripterygiid group and all other Blennioid fishes at the highest levels remain poorly resolved (Lin & Hastings, 2013).

The Tripterygiidae comprise bottom-dwelling Blennioid fishes, commonly found along cold, temperate, tropical and subtropical shores (Fricke, 2009; Nelson *et al.* 2016) throughout the Atlantic, Indian and Pacific Oceans (Stepien *et al.*, 1997). Fishes from this group are understudied,

especially when it comes to their taxonomic classification (Fricke, 1994; Holleman, 2005; Victor, 2015). Most triplefin fishes establish a very small territory of about 1–2 m² in which all feeding and breeding occur all year round, and recruits occupy the same habitat as conspecific adults (Thompson, 1983). Species of the Tripterygiidae family prey on various invertebrates, especially crustaceans (Feary, 2001; Langlois *et al.*, 2005). Triplefins mostly occupy and adapt to specialised conditions such as turbid waters (Harris *et al.,* 1999), feed on small invertebrates and therefore may positively or negatively affect algal growth (Wheeler, 1985; Allen & Robertson, 1994). Most of the tripterygiids appear to be habitat responders (Jones & Andrew, 1993), with their abundance and distribution typically associated with physical and biological characteristics of the reef substratum (Greenfield & Johnson, 1990; Hofrichter & Patzner, 2000). Above all, the Tripterygiidae fishes are the most numerically dominant and speciose group among the blennioids that occupy intertidal and shallow subtidal temperate coral reefs in New Zealand (Paulin & Roberts, 1992; Chesson, 2000; Munday, 2004).

Investigations into the ecology of triplefins have mainly focused on single species within the intertidal zone (Willis & Roberts, 1996; Beckley, 2000). Studies on the larval fishes in South African nearshore waters reveal that preflexion larval fish specimens from the Tripterygiidae are commonly found (Harris *et al.,* 1999; Harris *et al.,* 2001; Pattrick & Strydom, 2008). Harris *et al.* (1999) reported a 16% contribution of an unidentified Tripterygiidae larva in the nearshore coastal waters off St Lucia Estuary, KwaZulu-Natal. Furthermore, this fish family contributed 2% towards the total larval fish catch in the shallow nearshore of Algoa Bay, Eastern Cape (Pattrick & Strydom, 2008).

During the current research, a common preflexion larva from shallow nearshore larval fish surveys within the Eastern Cape, South Africa was identified through traditional morphological techniques as a species from the Tripterygiidae family (Leis & Carson-Ewart, 2000). DNA barcoding was used to validate the morphological identification. A fragment of the cytochrome *c* oxidase subunit 1 (*COI*) gene, a well-studied gene region for larval fish species identification (Ko *et al.,* 2013; Friedheim, 2016), was sequenced. Another mitochondrial marker, 16S rRNA, was also used (Aoyama *et al.,* 2001; Karaiskou *et al.,* 2007; Lelievre *et al.,* 2012, Ockhuis *et al.,* 2017). The main reason for using only mitochondrial DNA during the present study was because the samples contained a very low concentration of DNA template, and the mitochondrial DNA has a low resistance to deterioration compared to nuclear DNA (Karaiskou *et al.,* 2007; Scwarz *et al.,* 2009). The DNA molecule is small in size, circular in structure, protected by a cellular compartment of the mitochondrion and has a high nucleotide substitution rate, which makes it more suitable to use in species identification (Aoyama *et al.,* 2001; Karaiskou *et al.,* 2007; Lelievre *et al.,* 2012).

Mitochondrial genes are thus ideal as DNA barcodes (Saccone *et al.,* 1999). On account of the high diversity of the larvae of cryptic species observed in larval fish catches in the shallow nearshore of the Eastern Cape, therefore, the main aim of this study was to use DNA barcoding was to clarify the morphological identification of the egg and larval specimens of the Tripterygiidae family to species level.

3.2 Material and Methods

The specimens forming the basis of this chapter were collected from the same study area, using the same collection, sorting and morphological analytical techniques as described in Chapter 2.

3.2.1 DNA extraction, PCR amplification and sequencing

In the laboratory, DNA extraction from one fish egg and 30 larval specimens followed the same protocol as Chapter 2. Polymerase chain reaction (PCR) for *COI* was carried out using the primer pair HCO2198 and LCO1490 (Folmer *et al.,* 1994), both at 10 µM concentration. The PCR reaction contained 10 µl of Ready Mix Taq, 4 µl of molecular grade water; 0.5 µl of each forward and reverse primers, and 5 µl of DNA template. The PCR thermal cycling regime comprised an initial denaturing at 95°C for 2 minutes, followed by 35 cycles of denaturing at 95°C for 1 minute, annealing at 50°C for 1 minute and extension at 72°C for 1 minute, followed by a final extension at 72°C for 7 minutes.

In addition, a fragment of the 16S rRNA mtDNA marker was amplified, using 10 µM 16Sar and 16Sbr universal primers (Palumbi, 1996). The PCR thermal cycling profile for 16S rRNA was as follows: an initial denaturation at 94°C for 5 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 48°C for 30 seconds and extension at 72°C for 45 seconds; then final extension was at 72°C for 10 minutes. After electrophoresis, PCR products were visualised under ultraviolet light on 1% agarose gel, stained with ethidium bromide. Indication of a single band showed a successful amplification. The amplified samples were purified and sequenced in the same way as described in Chapter 2.

3.2.2 Sequence alignment and analysis

Both *COI* and 16S sequences were observed separately in Chromas Pro v.2.1.0 (Technylesium Pty Ltd) to verify quality after sequencing. SeqMan Pro (DNASTAR® Lasergene® 7.2) was used to edit trace files and to correct possible base ambiguities prior to assembling contigs for alignment. ClustalX version 2 (Larkin *et al.,* 2007) was used for final alignment of the sequences. For the comparison of sequences to identify the species, the Barcode of Life Data System (BOLD) (Ratnasingham & Hebert, 2007; https://www.boldsystems.org/) was used for the *COI* fragment only. *COI* and 16S sequences were both compared on the National Center for Biotechnology Information (NCBI) database using the standard nucleotide Basic Local Alignment Search Tool (BLAST) (https://www.ncbi.nlm.nih.gov/) (Altschul *et al.,* 1990).

3.2.3 Phylogenetic analysis

Molecular Evolutionary Genetics Analysis (MEGA) software version 6.0 (Kumar *et al.,* 2012) was used to analyse the sequence data. A Kimura two-parameter (K2P) model (Kimura, 1980), implemented in MEGA version 6.0 (Tamura *et al.,* 2013) was used to calculate the genetic distance between sequences. The K2P model is widely used in barcoding studies and is regarded as being reliable in estimating evolutionary distances between sequences (Ward, 2009; Randall & Victor, 2013; Tamura *et al.,* 2013; Puncher *et al.,* 2015). Its use here allowed better comparison of the results and sequence divergences with previous studies.

The genetic relationships between the sequences of 28 Tripterygiidae larval specimens generated in this study, three Tripterygiidae larvae (FLIL013-14; FLIL063-14; FLIL089-14) downloaded from BOLD, and one adult Clinidae (HQ168634.1) from the GenBank databases were assessed. The assessment was conducted by constructing a neighbour-joining (NJ) tree (Saitou & Nei, 1987), using K2P divergences for the *COI* sequences. The bootstrap analysis (Felsenstein, 1985), with 1000 replicates, was performed to test the statistical support at the nodes. The values for the bootstrap analysis are indicated as percentages, where a value equal to or larger than 70% equates to a probability of 0.95 or significance at $p < 0.05$ support for the clade (Hillis & Bull, 1993).

3.3 Results

3.3.1 *COI* sequences

A total of 30 larval specimens was identified using a traditional morphological approach as belonging to the Tripterygiidae family. Of the 31 specimens (including one egg), 28 DNA extracts were successfully amplified for the *COI* to confirm the morphological identification. The total lengths of the 28 sequences for the *COI* fragment were between 550 and 680 nucleotides. As a result of the differences in sequence lengths, all 28 of the DNA sequences of the *COI* gene region were compared with data on BOLD. For 26 sequences, a probability match of 99 – 100% was found to sequences from Tripterygiidae larvae on BOLD (Table 3.1). The other two, FBIP0396 and FBIP0152, had 91% and 95% matches, respectively, to the tripterygiid sequences in BOLD.

GenBank was used as an alternative online database for both *COI* and 16S sequences to search for a possible match using a BLAST search (as in Chapter 2). All 28 *COI* sequences were identified to species level, with 99% similarity to a sequence indicated as being derived from *Pavoclinus profundus* (Family Clinidae) (Table 3.1). An unexpected challenge was presented when this *Pavoclinus profundus* sequence from GenBank (accession number HQ168634.1), published by Lin and Hastings (2013), was downloaded and compared to BOLD. Surprisingly, this *P. profundus* sequence was assigned the same Barcode Index Number [BIN, which is the BOLD's unpublished interim taxonomic and identification system (Ratnasingham & Hebert, 2013)], BIN AAX1770, as species from the order Blenniiformes, family Tripterygiidae. The sequences of the three larval specimens, FLIL013-14; FLIL063-14; FLIL089-14 (Table 3.1), from the above-mentioned BIN were collected in 2013 at Port Elizabeth, South Africa, photographed and identified by Dr Paula Pattrick.

Table 3.1: Top matching sequences for the *COI* gene fragment from BOLD and NCBI databases.

BOLD Process ID	**Sequence match (%)**
FLIL089-14 – Tripterygiidae	99 -100
FLIL063-14 – Tripterygiidae	99 -100
FLIL013-14 – Tripterygiidae	99 -100
NCBI Accession no.	
HQ168634.1 - *Pavoclinus profundus*	99

K2P genetic distances were calculated among the *COI* sequences generated for this chapter and those from the online databases (Table 3.2). A neighbor-joining tree was constructed using K2P distance values to provide a graphic presentation of distances between sequences (Figure 3.1).

Sequences generated in this study grouped well with data from BOLD (FLIL-sequences), and the GenBank sequence (*Pavoclinus profundus*). FBIP0152, one of the 28 sequences generated for this study, had genetic distances that ranged between 0.04 and 0.09 to all the sequences comparisons; as a result it was slightly differentiated on the phylogenetic tree.

3.3.2 *16S* sequences

Twenty-six of the 31 DNA extracts were also successfully PCR amplified for the 16S gene fragment. The length for the 26 obtained 16S rRNA gene region sequences was 586 nucleotides. Sequences obtained from the 16S rRNA gene region matched with a 93.49 – 97.63% similarity to *Blennophis anguillaris* (Family Clinidae) on GenBank (Table 3.2). The similarity value in most of these comparisons was less than the threshold value of 97% adopted for species identification based on *COI* in this study (Meyer & Paulay, 2005; Meier *et al.,* 2006). Although the 16S gene is highly conserved (Stackebrandt & Rainey, 1995; Gurtler & Stanisich, 1996), using a fixed threshold cannot reliably identify species (Schloss, 2010) and these comparisons are not likely to represent conspecific comparisons and accurate species identification. Thus, 16S sequences were not used for further analysis, such as the construction of a neighbor-joining tree.

Table 3.2: K2P genetic distances between sequences from BOLD (Tripterygiidae), GenBank (*P. profundus*) and the unidentified specimens calculated in MEGA based on *COI* sequences.

	Sequences	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	FLIL063-14																															
2	FBIP0451	0.00																														
3	FBIP0151	0.00	0.01																													
4	FLIL013-14	0.00	0.00	0.00																												
5	FLIL089-14	0.00	0.00	0.00	0.00																											
6	FBIP0460	0.00	0.00	0.02	0.00	0.00																										
7	FBIP0429	0.00	0.00	0.02	0.00	0.00	0.00																									
8	FBIP0464	0.00	0.00	0.01	0.00	0.00	0.00	0.00																								
9	FBIP0433	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00																							
10	FBIP0081	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00																						
11	FBIP0396	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02																					
12	FBIP0445	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02																				
13	FBIP0468	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00																			
14	FBIP0385	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01																		
15	FBIP0102	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01																	
16	FBIP0122	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00																
17	FBIP0446	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00															
18	FBIPUNK2	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00														
19	FBIP0458	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.05	0.05	0.04	0.04	0.04	0.04													
20	FBIPUNK5	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.05												
21	FBIP0357	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00											
22	FBIP0126	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00										
23	FBIP0428	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00									
24	FBIPUNK1	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00								
25	FBIP0448	0.01	0.01	0.02	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.01	0.01	0.02	0.00	0.00	0.00	0.01	0.05	0.01	0.00	0.00	0.00	0.00							
26	FBIP0088	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.01						
27	FBIPUNK3	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.01	0.00					
28	FBIP0422	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00				
29	FBIP0332	0.00	0.00	0.02	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.02	0.00	0.01	0.02	0.01	0.01	0.01	0.00	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01			
30	FBIP0453	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01		
31	*P. profundus*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	
32	FBIP0152	0.06	0.07	0.08	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.09	0.07	0.07	0.04	0.07	0.07	0.07	0.07	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06

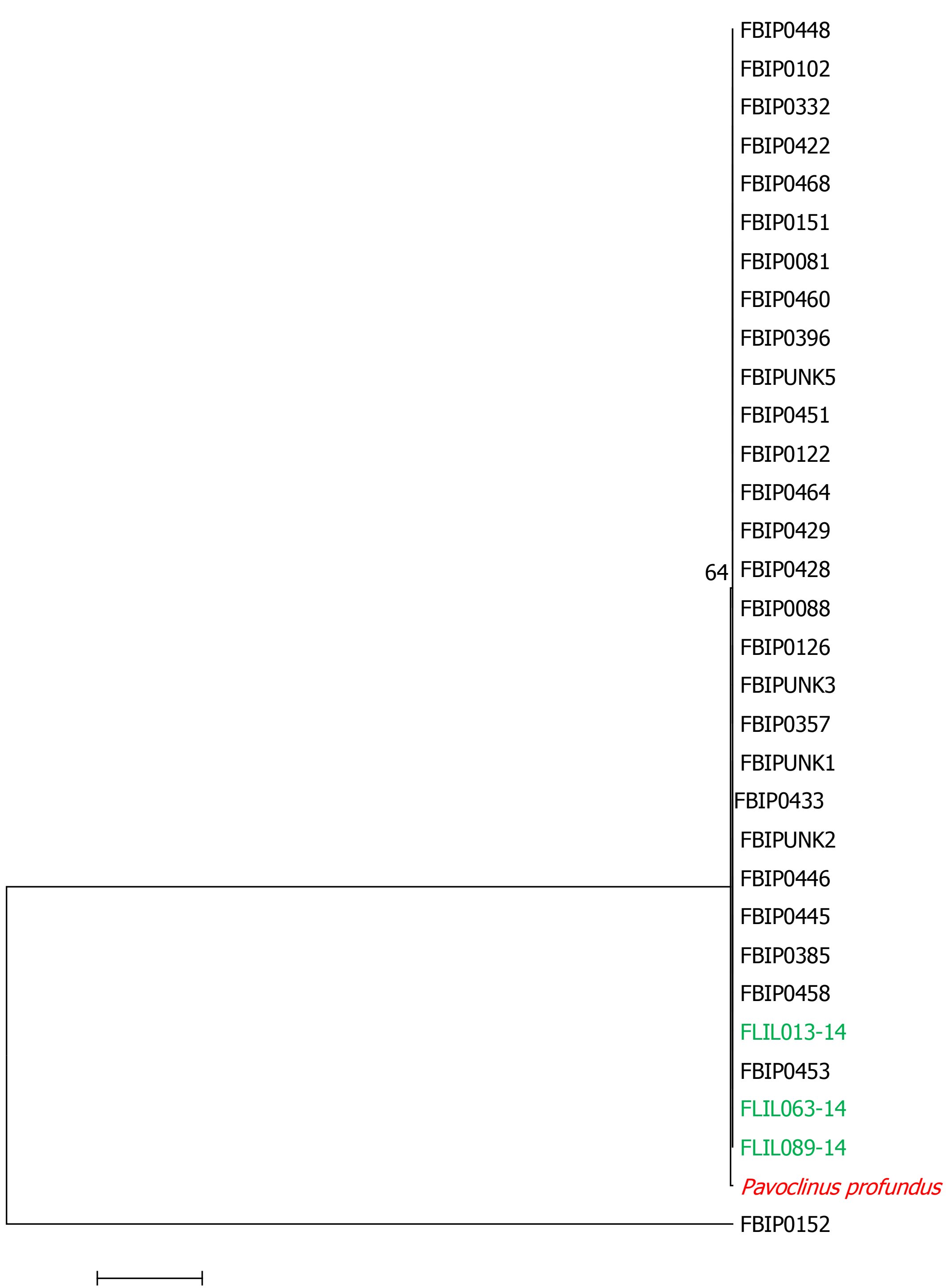

Figure 3.1: Neighbor-joining tree based on K2P genetic distances of the *COI* sequences for an egg and larval fishes (in black) as well as accessioned sequences of Tripterygiidae larvae from BOLD (in green) and a Clinidae adult from GenBank (in red). The scale bar represents 0.05 sequence divergence.

3.4 Discussion

The taxonomy of the early life history stages of the triplefin blennies is among the least-studied of the blennioides (Robertson, 1994; Nelson, 1994; Springer, 1998). Using morphological techniques, larval specimens from the present study were identified as those from the Tripterygiidae family. Morphological characters were however, not sufficient to resolve the identification beyond the family level. Blenniiformes generally contribute substantially towards larval fish catches both locally and internationally (Harris & Cyrus, 1999; Sabatés *et al.,* 2003; Beldade *et al.,* 2006). Tripterygiid larvae have been identified in ichthyoplankton studies in temperate nearshore regions in the Mediterranean with *Tripterygion tripternotus* dominating spring and summer catches (Sabatés *et al.,* 2003). Likewise, in nearshore Australian waters, Tripterygiidae dominated the larval catches (Gray, 1993). Tripterygiidae larvae have also been found in abundance in the nearshore coastal waters off the St Lucia Estuary, South Africa, but specimens were also not identified beyond the family level (Harris, *et al.,* 1999). In the same coastal nearshore region as the current study, Tripterygiidae larvae contributed 2% towards the total catch (Pattrick & Strydom, 2008).

Eggs and larvae of different species in the Blenniiformes are morphologically identical, which makes it difficult to classify them to genus, let alone the species level (Shao *et al.,* 2002; Duke *et al.,* 2018). Also, not all species that are genetically different exhibit distinguishing morphological characters, but they can, however, be identified through molecular analysis (von der Heyden *et al.,* 2011). For example, the preflexion larval stages of Sparidae without recognisable diagnostic characters are easily misidentified as Scorpaenidae, Scombridae, Haemulidae and Terapontidae using morphological techniques (Ko *et al.,* 2013). Also, a specialised knowledge of early life history stages is the minimum requirement for traditional morphological identification (Ko *et al.,* 2013; Overdyk *et al.*, 2016).

In the present study, the morphological identification of unknown preflexion larval specimens as species from the Tripterygiidae were confirmed through DNA barcoding. The specimens could, however, not be identified to the lowest taxonomic level using both the morphological identification and the DNA barcodes. Although the taxonomic representation of sequences in the BOLD reference database is not easy to evaluate for large geographic regions because of inaccuracy in identification, cryptic species diversity and incomplete species lists, small-bodied blennioides still have a lower taxonomic coverage than large-bodied families (Victor *et al.*, 2015). There has been an effort to barcode small-bodied reef-associated fishes found in the Indo-Pacific, including regions like French Polynesia (Hubert *et al.*, 2012), the South China Sea

(Zhang *et al.*, 2012), Queensland (Australia) and Bali (Indonesia) (Ward *et al.*, 2005) as well as southern Africa (Swartz *et al.*, 2008). Molecular investigations using the mitochondrial genes (*COI*, 16S and 12S) have been conducted on the phylogeny of adult triplefins; however, these studies were constrained by regional and taxon sampling (Geertjes *et al.,* 2001; Carreras-Carbonell *et al.*, 2005; Hickey & Clements, 2005). The popularity of DNA barcodes as a primary identification tool and as a highly efficient marker has resulted in the development of a large database for the *COI* mitochondrial gene fragment (Ward *et al.,* 2009). Nonetheless, not all existing species are documented on the database.

The NJ phylogenetic tree, derived from a fragment of the *COI* mitochondrial gene, strongly suggests that one of the sequences generated from this study could be of the same genus, but of a different species. The highest K2P genetic distance of 0.09 was between sequences FBIP0152 and FBIP0396, which suggests an interspecific comparison. An interspecific genetic distance of 0.073 has been reported among deep sea species of the genus *Chelidoperca* (Family: Serranidae) from Indian Ocean waters (Bineesh *et al.,* 2015). The latter value is similar to, or not too far from, the K2P interspecific genetic distances reported for other marine fishes using *COI.* An interspecific genetic distance of less than 5% was described within specimens of the same genus by Zhang and Hanner (2012). Also, a 6.50% K2P distances within genera was reported in fish species of the Taiwan Strait (Bingpeng *et al.,* 2018).

The barcode index number system (BIN) is an online structure in the BOLD database where barcode sequences are grouped together based on geographic information, descriptive metadata and images for specimens that probably belong to the same species (Ratnasingham & Hebert, 2013). Upon further investigation in this study, it was found that two fish families, Clinidae and Tripterygiidae, have been assigned to the same BOLD database BIN AAX1770. A similar shortcoming has been highlighted by Steinke *et al.* (2016) for the southern African species of *Engraulis encrasicolus/Engraulis japonicus* (Family Engraulidae), and also for *Trachurus delagoa; Trachurus murphyi* and *Trachurus trachurus* (Family Carangidae) that shared the same BIN on BOLD respectively. The sharing of the same BIN for *Engraulis encrasicolus/Engraulis japonicus* at least, was due to the change in name for that species (Steinke *et al.,* 2016). BIN sharing has been documented more frequently for the commercially important tuna, genus *Thunnus.* B BIN sharing of unresolved taxonomies could be resolved by sequencing additional mitochondrial or nuclear genes to authenticate identified species (Abdullah & Rehbein, 2014).

In the case of hybridisation or genetically closely-related fish species, the use of nuclear genes together with DNA barcoding could be useful for species assignment (Kochzius, 2009;

Teletchea, 2009). For this study, haploid and conserved genes were used to identify the specimens, and ambiguities were noticed at the species level with results yielding two different species identifications. Larvae, especially of small-bodied fish species, are difficult to resolve taxonomically due to the lack of prominent morphological differences among the closely-related families that could lead to misidentifications of specimens. As has been shown in this study, Clinidae and Tripterygiidae families shared a BIN on the BOLD online database. Mismatch at the family level shows that incorrect identification of small and cryptic fishes can lead to increased misidentification rates.

There are currently no safeguards against a BOLD contributor misidentifying a specimen, and when a name has been added into a record list, it may be challenging for a third party to demonstrate that it could be different (Ratnasingham & Hebert, 2007). A validated BIN, however, permits an evaluator to query the species identifications that do not match the validated identifications, whether to correct the specimen identity or to point out ambiguities in the BIN (Ratnasingham & Hebert, 2013). This study shows that there is a need for thorough investigation of the reef-associated, cryptic, small-bodied fish species found along the south-east cost of South Africa. There is evidence that many of the small reef-associated species represent cryptic, genetically different lineages (Bickford *et al.,* 2007) that need to be assessed taxonomically (von der Heyden, 2011). The sequence data generated can be an additional important asset in the identification of small reef fishes once data is submitted to the BOLD online database with supplementary information such as photographs (Ratnasingham & Hebert, 2007).

It is well recognised that the taxonomy of the Blenniiformes is uncertain (Lin & Hastings, 2013), due to the absence of a well-resolved phylogeny (Hasting & Springer, 2009). In southern Africa in particular, approximately 25% of all endemic marine fishes are undescribed and this includes the small (<10 cm) benthic individuals residing in shallow marine environments (von der Heyden, 2011). A closer taxonomic investigation of the early life history stages of the Tripterygiidae and Clinidae species will help to prevent BIN sharing of sequences in the online BOLD database.

Limited taxonomic information on the Blenniiformes poses a risk to the understanding of their biodiversity, ecology, and the overall functioning of the systems these fish inhabit (Hooper *et al.,* 2005). Failure to recognise species that are known to be common in the marine environment affects our understanding of the evolution and ecology of marine communities and ultimately how environmental factors affect them (Knowlton, 1993). The major constraint in this study could be that there were insufficient barcode data to compare the generated sequences to. It has

been shown that, if DNA barcode libraries were complete, the barcodes could generally perform very well for identification (Kress & Erickson, 2008; Bingpeng *et al.,* 2018). Problems occur when the reference databases are incomplete, as in the case of this study. One of the most serious limitations to DNA barcoding as an applied resource for molecular diagnostics is not necessarily associated with the quality of the DNA, but rather with the human inaccuracy in producing and curating a DNA barcoding reference library (Overdyk *et al.,* 2016; Hulley *et al.,* 2018).

In this study, DNA barcoding failed to identify to species level an egg and larvae of a seemingly common specimen in the shallow coastal waters of South Africa. This result shows that the effectiveness of both approaches is limited by the lack of accurate larval descriptions for the morphology and an incomplete reference database for the barcoding. The major limitation to this study was that there are no keys to identify the larval specimens morphologically to the lowest taxonomic level in South Africa. Additional sampling of adults and larvae therefore should be conducted.

CHAPTER FOUR

4.1 General Discussion

The groundwork for scientific studies such as conservation biology, behaviour, systematics, ecology, phylogeny and evolution are all based on the correct identification of species (Wilson, 2004; Ohl, 2014). In the field of taxonomy, where the primary task is to identify and delineate species, there is a crisis of limited funding and very few specialists for major and economically important groups (Godfray, 2002; Wilson, 2004; Swartz *et al.*, 2008; Griffiths *et al.*, 2010). After three centuries of inventorying marine living organisms, the earth's biodiversity, which consists of not less than 10 million described species to date, is still poorly known and includes many species that carry potential economic and societal applications that are still to be discovered (Vernooy *et al.,* 2010; Griffiths *et al.,* 2010). The records from three Canadian ocean provinces (Arctic, cold temperate Northwest Atlantic, and cold temperate Northeast Pacific) show that the inventory of marine fish taxa is constrained by limited knowledge of the diversity in the deep sea (Coad & Reist, 2004; Mora *et al.,* 2008; Archambault *et al.,* 2010). Africa at large has a great diversity when it comes to fish species, but documentation and assessments of biodiversity are hindered by slow progress in the taxonomy of adults (Swartz *et al.,* 2008). In South Africa, currently more than 12,914 species have been recorded, although individuals that are small in size remain undescribed (Griffiths *et al.,* 2010). It has been estimated that, with the current rate of fish descriptions, it will take about 50 years to document and describe the total fish fauna in South Africa (von der Heyden, 2011).

Regardless of the importance of identifying species for either academic or societal purposes, the interest in taxonomy and in pursuing the inventory of earth's living organisms has declined since its earlier development during the 18th century by Carl Linnaeus (Mallet & Willmott, 2003). The decline is mostly triggered by the retirement of specialists in taxonomy and systematics, and the difficulty in training new ones (Herbert, 2001; Griffths *et al.,* 2010; Archambault *et al.,* 2010; von der Heyden, 2011). As the commonly used and long-practised procedure of morphological identification has its limitations, more suitable approaches for species identification are needed (Hebert *et al.*, 2003; Swartz *et al.*, 2008; Ko *et al.,* 2013). Fishes undergo different life stages, making species description based on morphological characteristics alone inconsistent (Leis & Carson-Ewart, 2000; Friedheim, 2016). Dayrat (2005) proposed an integrative taxonomic approach, where a combination of molecular and morphological approaches can be used for effective species delineation. Several marine fish species have been described successfully using

the above-mentioned approach (e.g., Shao *et al.,* 2002; Victor *et al.,* 2009; Lelievre *et al.,* 2012; Randall & Victor, 2013).

This study tackled this black box of species identification by targeting the taxonomy of the early life history stages of fishes occurring on the south-east coast of South Africa, using an integrated morphological and molecular approach. Individuals from the family Gobiidae and Tripterygiidae were identified morphologically only to family level. Also, only one Soleidae specimen was positively assigned to species level based on morphological characteristics. *Diplodus sargus* and *Austroglossus pectoralis* specimens positively identified in this study are endemic to southern Africa and are a major component of the line and trawl fishery respectively (Heemstra & Gon, 1986; Wood, 2000); an important finding, considering that the South African fisheries sector contributes approximately 1% to the national gross domestic product (FAO, 2009; SAC-EU EPA, 2017).

Inconsistencies in the DNA barcode data analysis of the larval specimens, as seen for the tripterygiid larva, revealed a need for more intense ichthyoplankton surveys and the coupling of integrative approaches such as the one used in my research. Intensified ichthyoplankton surveys will not only serve ecological and conservation purposes (Termvidchakom & Hortle, 2013; Rodriguez *et al.,* 2017; van der Lingen & Huggett, 2003), but will also improve taxonomic identifications (Fox *et al.,* 2005). When the DNA barcode database has been further developed, the scientific and practical benefits of small-bodied fish barcoding will be diversified. DNA barcoding will provide more information for species descriptions.

There was, indeed, limited barcoding data when it came to the small-bodied fishes found in the study region. Limited information on these fishes shows that some of the world's biodiversity is not fully explored, especially for those species that are not economically important (von der Heyden, 2011). Also, species descriptions are often based on few specimens with inadequate knowledge about the distribution range and the ecology of the species (Cowen *et al.,* 2000). This poor taxonomic knowledge of early life stages results in limited knowledge of population connectivity, and can jeopardise management and conservation efforts of marine fish species (von der Heyden, 2011; Ardura *et al.,* 2016; Bingpeng *et al.,* 2018;). Despite these challenges, DNA barcoding is a powerful tool for rapid and accurate species identification (Azmir *et al.,* 2017; Bingpeng *et al.,* 2018; John *et al.,* 2018). DNA barcodes are very useful for describing biodiversity in ecosystems that are species-rich, difficult to evaluate, and poorly catalogued

(Swartz *et al.,* 2008; April *et al.,* 2012; Ardura *et al.,* 2016); through DNA barcoding, commercially important, small-bodied, and endemic fishes were identified in this study.

I have found that the ability of morphological techniques to resolve larval fish specimens to the species level is not significantly different from the ability of DNA barcoding, especially when considering the least-studied individuals such as Tripterygiidae and Gobiidae species. There are however, a few differences that were noticed between the two techniques caused by the inability of both DNA barcodes and morphological characteristics to identify Tripterygiidae and Gobiidae specimens beyond the family level. Nonetheless, I established that DNA barcoding is more robust than morphological characteristics in species identification for other taxonomic groups. This is also evident from other studies conducted to identify different species of fish (Ward *et al.,* 2009; Hubert *et al.,* 2010; Zhang *et al.,* 2011; Ko *et al.,* 2013, Ockhuis *et al.,* 2017).

DNA barcodes were also successful in identifying eggs and damaged larval specimens to the species level, when morphological identification was impossible. DNA barcoding has the ability to function, regardless of the physical conditions of the specimens. In the case of fish eggs that are difficult to identify based only on morphology, I would suggest the use of another gene region, in addition to the *COI.* Multiple gene sequencing would allow for the verification of identities (Zhang & Hanner, 2011). In an event where divergence is not so recent, nuclear markers can be used for phylogenetic analysis (Lin & Danforth, 2004). Further, mitochondrial protein-coding genes can now be sequenced using a Roche 454 platform at relatively low cost, providing markers for systematics applications, and additional data to potentially improve DNA barcode resolution (Timmermans *et al.,* 2010).

The solution to the mismatch in species identification when sequences generated from this study are compared on the online database could be solved by obtaining a DNA barcode or sequence from the type specimens. This would not be possible in most cases when types are unavailable or cannot be used to extract DNA. In such instances, DNA identification would be based on newly collected individuals which are examined by an experienced taxonomist to confirm their identity (Tauz *et al.,* 2003). Secondly, a database with a centralised, universal system for regional species (before submission of the data to the international dataset such as BOLD) which could be established from current efforts, would help to minimise mistakes and also improve the current species identification output, complementing the traditional morphological technique. Undoubtedly, the technical difficulties related PCR failure, mismatch, and misidentifications

should be overcome as to improve successful sequencing. With the incorporation of a single-copy nuclear marker (Prada *et al.,* 2014), the success rate of DNA barcoding is increasing.

In conclusion, what I can infer from this study is that the Phylogenetic Species Concept which gives information about the evolutionary history of an organism (Mayden, 1997; Coyne & Orr, 2004) based on morphological characteristics and molecular barcode (Cracraft, 1989; Coyne & Orr, 2004) is more suitable for identifying the early life history stages of fish. The above-mentioned species concept complements the integrative approach very well when it comes to species level identification of larval fishes. Use of the integrated species concept, which involves both DNA barcoding and morphological characteristics, could increase the number of positive identifications of fish species. The DNA barcoding method produces a short genetic marker to enable identification of a particular specimen; this can be done even by non-specialists in a very short period of time (Ardura *et al.,* 2016). However, if the barcodes are unable to identify the species, another approach is the nearest-neighbour BLAST which can be used to delimit anonymous reads (Zhang *et al.* 2013). Although the unknown sequence query depends on the completeness of the taxon sampling in the reference database, the use of a sequence similarity gap can reveal previously unsampled species as distinct (Meyer & Pulay, 2005). With the decline in the number of taxonomists, the use of DNA barcoding in identifying damaged or physically intact specimens could reveal cryptic and least-studied individuals. Even though DNA is regarded as accurate and rapid when it comes to species level identification, there is a pressing need for development of conventional taxonomic keys for ichthyoplankton samples that can be used in cross-questioning genetic data.

4.2 Conclusion and Recommendations

Molecular techniques have been used extensively and have recently gained popularity for the identification of different fish species (Zhang & Hanner, 2011; Ardura *et al.*, 2016). Prior to the use of DNA barcoding in species identification, taxonomists relied only on morphology. A morphological approach however, requires specimens to be intact and relies heavily on the expertise of the taxonomist in charge of that particular species under study (Ko *et al.*, 2013). Molecular techniques were hence designed to produce a more reliable method that can be used in species identification and to accommodate specimens that are no longer morphologically intact. Even though there is an increased application of DNA barcoding and morphological techniques in species identification, there are no studies that have previously used both techniques to identify the larval fishes from the south-east coast of South Africa.

The morphological approach in this study was, however, constrained by the limited reference data for species identification beyond the family level, while DNA barcodes enabled identification to the lowest taxonomic level. Similarities were observed between the two techniques, these include the inability of DNA barcoding to delineate Tripterygiidae specimens to the species level; and the inability of present morphological approaches to identify specimens beyond the family level. However, use of DNA barcoding led to species level identification of fish eggs. These two techniques could not be compared in the identification of fish eggs and larvae in this study because of limited tools, such dichotomous keys and a reference library for species found on the south-east coast. A comprehensive reference library is especially necessary for cryptic and ecologically important species, which may not be important from a commercial or recreational fishery perspective. I therefore conclude that despite the inability of DNA barcoding to resolve benthic cryptic species, it is still a more robust technique for the identification of fish eggs and larvae than to a morphological approach. While this demonstrates that morphological identifications by an expert ichthyologist are highly accurate, it shows that DNA barcoding is a slightly more accurate technique, when DNA barcoding is able to identify the specimen to the species level with an accuracy of 99-100%.

However, for future research on various life history phases of fishes that are distributed across different spatial and temporal scales (Pattrick & Strydom, 2008), I would recommend another gene to be used in addition to the COI gene for DNA barcoding. This would be suitable for understanding the spawning habitats which strongly affect recruitment success (Borja *et al.,* 2008) and, consequently, the demographic stability of a population (Bellier *et al.,* 2007). Hence, the location, quality and size of spawning areas are essential factors influencing the spatial and

temporal dynamics of fish populations (Russ & Alcala, 2004). The only obstacle would be the development of a comprehensive database for alternative genes. This study was directly impacted by the limited reference data, time available for second gene sequencing and lack of experience in the morphological identification of larval fishes beyond the family level. Therefore, there is a need to train junior researchers who are interested in the taxonomy of early life history stages of fish; alternatively, the development of a book with photographs, and specimen descriptions of the early life history stages of different species occurring in this region.

References

ABDULLAH A., REHBEIN H. 2014. Authentication of raw and processed tuna from Indonesian markets using DNA barcoding, nuclear gene and character-based approach. *European food research and technology*, **239**(4): 695–706.

ACKERMAN L.J., BELLWOOD D.R. 2002. Comparative efficiency of clove oil vs rotenone for sampling tropical reef fish assemblages. *Journal of fish biology,* **60**: 893–901.

ALDHEBIANI A.Y. 2018. Species concept and speciation. *Saudi journal of biological sciences,* **25**: 437-440.

ALLEN G., ROBERTSON D. 1994. *Fishes of the tropical eastern Pacific*. Honolulu, HI: University of Hawaii Press.

ALMANY G.R., BERUMEN M.L., THORROLD S.R., PLANES S., JONES G.F. 2007. Local replenishment of coral reef fish populations in a Marine Reserve. *Science*, **316**:742–744.

ANDERSON D.R. 2008. Model based inference in the life sciences: a primer on evidence. Springer, New York.

ANDERSSON L. 1990. The driving force: species concepts and ecology. *Taxon,* **39**(3): 375–382.

AOYAMA J., ISHIKAWA S., OTAKE T., MOCHIOKA N., SUZUKI U., WATANABE S., SHINODA A., INOUE J., LOKMAN P., INAGAKI T., OYA M., HASUMOTO H., KUBOKAWA K., LEE T., FRICKE H., TSUKAMOTO K. 2001. Molecular approach to species identification of eggs with respect to determination of the spawning site of the Japanese eel Anguilla japonica. *Fisheries science,* **67**:761–763.

APRIL J., MAYDEN R.L., HANNER R.H., BERNATCHEZ L, 2012. Genetic calibration of species diversity among North America's freshwater fishes. *PNAS,* **108**(26): 10602–10607.

ARCHAMBAULT P., SNELGROVE P.V.R., FISHER J.A.D., GAGNON J.M., GARBARY D.J., HARVEY M., KENCHINGTON E.L., LESAGE V., LEVESQUE M., LOVEJOY C., MACKAS D.L., MCKINDSEY C.W., NELSON J.R., PEPIN P., PICHE L., POULIN MICHEL. 2010. From Sea to Sea: Canada's Three Oceans of Biodiversity. *PLoS ONE*, **5**(8): e12182.

ARDURA A., MOROTE E., KOCHZIUS M., GRARCIA-VAZQUEZ E. 2016. Diversity of planktonic fish larvae along a latitudinal gradient in the Eastern Atlantic Ocean estimated through DNA barcodes. *Peer reviewed journal,* **4**:e2438.

ARMSTRONG M.J., CONNOLLY P., NASH R.D.M., PAWSON M.G., ALESWORTH E., COULAHAN P.J., DICKEY-COLLAS M., MILLIGAN S.P., O'NEILL M.F., WITTHAMES P.R., WOOLNER L. 2001. An application of the annual egg production method to estimate the spawning biomass of cod (Gadus morhua L.), plaice (Pleuronectes platessa L.) and sole (Solea solea L.) in the Irish Sea. ICES *Journal of Marine Science*, **58**(1): 183– 203.

AVISE J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press.

AWAD A.A., GRIFFITHS C.L., TURPIE J.K. 2002. Distribution of South African marine benthic invertebrates applied to the selection of priority conservation areas, *Diversity and distributions,* **8**: 129-145.

AZMIR I.A., ESA Y., AMIN S.M.N., Md YASIN I. S., Md YUSOF, F. Z. 2017. Identification of larval fish in mangrove areas of Peninsular Malaysia using morphology and DNA barcoding methods. *Journal of applied ichthyology*, **33**(5): 998–1006.

BALDWIN C.C. 2013. The phylogenetic significance of colour patterns in marine teleost larvae. Zoological Journal of the Linnean Society, **168**(3): 496-563.

BALON E.K. 1979. The theory of saltation and its application in the ontogeny of fishes: steps and thresholds. *Environmental biology of fishes,* **4***:* 97–101.

BALON E.K. 1981. Saltatory processes and altricial to precocial forms in the ontogeny of fishes. *American zoologist,* **21**: 567–590.

BALON E.K. 1986. Saltatory ontogeny and evolution. *RivIsta Di biologia – biology forum,* **79**(2): 151–190.

BALON E.K. 1990. Epigenesis of an epigeneticist: the development of some alternative concepts on the early ontogeny and evolution of fishes. *Guelph ichthyology reviews,* **1**: 1–42.

BALON E.K. 1999. Alternative ways to become a juvenile or a definitive phenotype (and some persisting linguistic offenses). *Environmental biology of fishes,* **56**: 17–38.

BARRINGTON E.J.W. 1961. Metamorphic processes in fishes and lampreys. *American zoologist,* **1**: 97–106.

BARRINGTON E.J.W. 1968. Metamorphosis in lower chordates. In: Etkin, W. & Gilbert, L.I. (eds), Metamorphosis a Problem in Developmental Biology. *North-Holland publishing company, Amsterdam*, 223–270.

BECKER R.A., SALES N.G. SANTOS G.M., SANTOS G.B. CARVALHO D.C. 2015. DNA barcoding and morphological identification of neotropical ichthyoplankton from the Upper Parana and Sao Francisco. *Journal of fish biology,* **87**: 159-168.

BECKER S., HANNER R., STEINKE D. 2010. Five years of FISH-BOL: brief status report. *The Journal of DNA mapping, sequencing, and analysis,* **22**(S1): 3-9.

BELDADE R., BORGES R., GONCALVES E.J. 2006. Depth distribution of nearshore temperate fish larval assemblages near rocky substrates. *Journal of plankton research,* **28**: 1003-1013.

BERNARDI G. 2013. Speciation in fishes. *Molecular ecology*, **22**: 5487-5502.

BICKFORD D., LOHMAN D.J., SODHI N.S., NG P.K.L., MEIER R., WINKER K., INGRAM K.K., DAS I. 2007. Cryptic species as a window on diversity and conservation. *Trends in ecology & evolution,* **22**: 148–155.

BINEESH K.K., MOHITHA C., VINEESH N., BASHEER V.S., JOSELET M., PILLAI N.K.G., JENA J.K., GOPALAKRISHNAN A. 2015. Molecular identification of three deep-sea fish species of the genus *Chelidoperca* (Perciformes: Serranidae) from Indian waters. *Indian journal of fish* **62**: 104-108.

BINGPENG X., HESHAN L., ZHILAN Z., CHUNGUANG W., YANGUO W., JIANJUN W. 2018. DNA barcoding for identification of fish species in the Taiwan Strait. *Plos ONE*, **13**(6).

BLAXTER H. S. 1986. Development of sense organs and behaviour of teleost larvae with special reference to feeding and predator avoidance. *Transactions of the American fisheries society*, **115**: 98-1 14.

BOEHLERT G.W., MUNDY B.C. 1993. Ichthyoplankton assemblages at seamounts and oceanic islands. *Bulletin of Marine Science*, **53** (2): 336–361.

BOWEN B.W., MUSS A., ROCHA L.A., GRANT W.S. 2006. Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion of the Indian Ocean. *Journal of heredity,* **97**: 1-12.

BROOK B.W., BRADSHAW C.J.A., KOH L.P, SODHI N.S. 2006. Momentum drives the crash: mass extinction in the tropics. *Biotropica*, **38**: 302–305.

BROTHERS E.B., WILLIAMS D.M., SALE P.F. 1983. Length of larval life in twelve families of fishes at "one tree lagoon", Great Barrier Reef, Australia. *Marine biology*, **76**: 319–324.

BROWN S.D., COLLINS R.A., BOYER S., LEFORT M.C., MALUMBRES-OLARTE J., VINK C.J., CRUICKSHANK R.H. 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Molecular ecology resources*, **12**(3): 562–565.

BROWNELL C.L. 1979. Stages in the early development of 40 marine fish species with pelagic eggs from the Cape of Good Hope. Ichthyol. Bull. J.L.B. Smith Institute of Ichthyology. **40:** 1-84.

BUHAY J.E. 2009. "*COI*-Like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal crustacean biology*, **29**: 96–110.

BURROWS M., BROWNING J.S., BREITBART M., MURAWSKI S.A., PEEBLES E.B. 2018. DNA barcoding reveals clear delineation between spawning sites for neritic versus oceanic fishes in the Gulf of Mexico. *Fisheries oceanogr*aphy, **00**:1–12.

CALEY M.J., CARR M.H., HIXON M.A., HUGHES T.P., JONES G.P., MENGE B.A. 1996. Recruitment and the local dynamics of open marine populations. *Annual review of ecology, evolution and systematics*, **27**:477–500.

CARRERAS-CARBONELL J., MACPHERSON E., PASCUAL M. 2005. Rapid radiation and cryptic speciation in Mediterranean triplefin blennies (Pisces: Tripterygiidae) combining multiple genes. *Molecular phylogenetics and evolution*, **37**:751–761.

CLARIDGE M. F., DAWAH H. A., WILSON M. R. 1997. Species: The Units of Biodiversity. Chapman & Hall, London.

COAD B.W., REIST J.D. 2004. Annotated list of the Arctic marine fishes of Canada. *Canadian data report of fisheries and aquatic sciences*, 2674: 112.

COLLINS R.A., CRUICKSHANK R.H. 2013. The seven deadly sins of DNA barcoding. *Molecular ecology resources*, **13**(6): 969–75.

COSTA F.O., CARVALHO G.R. 2007. The Barcode of Life Initiative: synopsis and prospective societal impacts of DNA barcoding of fish. *Genomics society and policy*, **3**: 29–40.

COWEN R.K., LWIZA K.M., SPONAUGLE S., PARIS C.B., OLSON D.B. 2000. Connectivity of marine populations: open or closed? *Science*, **287**: 857–859.

COYNE J. A. ORR H. A. 2004. Speciation. Sinauer Associates, Sunderland, MA.

CRACRAFT J. 1983. Species concepts and speciation analysis. *Current Ornithology* **1**: 159–187.

CRACRAFT J. 1989. Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In Speciation and its Consequences (D. Otte and J. A. Endler, Eds.), pp. 28–59. Sinauer Associates, Sunderland, MA.

CRAWFORD S.S., BALON E.K. 1996. Cause and effect of parental care in fishes, an epigenetic perspective. *Advances in the study of behavior,* **25**: 53–107.

CRAWFORD, R.J.M., SHANNON, L.V., POLLOCK, D.E., 1987. The Benguela ecosystem. Part IV. The major fish and invertebrate resources. *Oceanography and marine biology annual review,* **25**: 353-505.

DAVID O., LARÉDO C., LEBLOIS R., SCHAEFFER B., VERGNE N. 2012. Coalescent-based DNA barcoding: multilocus analysis and robustness. *Journal of computational biology*, **19**(3): 271–278.

DAYRAT B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean society,* **85**: 407-415.

DEPCZYNSKI M., BELLWOOD D.R. 2003. The role of cryptobenthic reef fishes in coral reef tropho-dynamics. *Marine ecology progress series*, **256**: 183-191.

De-QUEIROZ K. 2007. Species concepts and species delimitation. *Systematic biology,* **56**(6): 879–886.

DETTAI A., LAUTREDOU A.C., BONILLO C., GOIMBAULT E., BUSSON F., CAUSSE R., COULOUX A., CRUAUD, C., DUHAMEL G., DENYS G., HAUTECOEUR M., IGLESIAS S., KOUBBI P., LECOINTRE G., MOTEKI M., PRUVOST P., TERCERIE S., OZOUF C. 2011. The actinopterygian diversity of the CEAMARC cruises: barcoding and molecular taxonomy as a multi-level tool for new findings. *Deep-sea research,* **58:** (250-263).

DOBZHANSKY, T.H. 1950 Mendelian populations and their evolution. *The American naturalist,* **84**: 401–418.

DONOGHUE M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist*, **88**(3): 172–181.

DUARTE C.M., ALCARAZ M. 1989. To produce many small or few large eggs: a size-independent reproductive tactic of fish. *Oecologia*, **80**: 401-404.

DUKE E.M., HARADA A.E., BURTON R.S. 2018. Large inter-annual variation in spawning in San Diego marine protected areas captured by molecular identification of fish eggs. *Marine ecology progress series,* **604**: 199-210.

FAITH D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biology and conservation*, **61**: 1-10.

FAITH D.P., BAKER A.M. 2007. Phylogenetic diversity and biodiversity conservation: some bioinformatics challenges. *Evolutional bioinformatics online* **2**: 121- 128.

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

FOLMER O., BLACK M., HOEH W., LUTZ R., VRIJENHOEK R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology,* **3**: 294–299.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). 2009. Fishery and aquaculture country profiles: South Africa. Available: http://www.fao.org/fishery/countrysector/FI-CP_ZA/en.

FOX C., TAYLOR I., PEREYRA R., VILLASANA M.I., RICO C. 2005. TaqMan DNA technology confirms likely overestimation of cod (*Gadus morhua* L.) egg abundance in the Irish Sea: implication for the assessment of the cod stock and mapping of spawning areas using egg-based methods. *Molecular ecology*, **14**: 897-884.

FOX C.J., TAYLOR M., DICKEY-COLLAS M., FOSSUM P., KRAUS G., ROHLF N., MUNK P., VAN DAMME C.J.G., BOLLE L.J., MAXWELL D.L., WRIGHT P.J. 2008. Mapping the spawning grounds of North Sea cod (*Gadus morhua*) by direct and indirect means. *Proceedings of royal society biology,* **275**:1543-1548.

FRICKE R. 1994. Tripterygiid fishes of Australia, New Zealand and the Southwest Pacific Ocean (Teleostei). Koenigstein: Koeltz Scientific Books.

FRICKE R. 2009. Systematics of the Tripterygiidae (triplefins). In The Biology of Blennies. Edited by Patzner RA, Gonçalves EJ, Hastings PA, Kapoor BG. Enfield: Science Publishers; 31–67.

FRIEDHEIM S. 2016. Comparison of Species Identification Methods, DNA barcoding versus Morphological Taxonomy. *Manoa horizons*, **1**: 74-86.

GEERTJES G.J., KAMPING A., VAN DELDEN W., VIDELER J.J. 2001. Genetic relationships among one non-endemic and two endemic Mediterranean triplefin blennies (Pisces, Blennioidei). *Marine ecology progress series*, **22**: 255–265.

GLEASON L.U., BURTON R.S. 2012. High-throughput molecular identification of fish eggs using multiplex suspension bead arrays. *Molecular ecology resources*, **12**: 57-66.

GODFRAY, H.C.J. 2007. Linnaeus in the information age. *Nature,* **446**: 259-260.

GOVONI J.J., LABAN E.H., HARE J.A. 2003. The early life history of sword fish (*Xiphias gladius*) in the western North Atlantic. *Fishery bulletin,* **101**: 778-789**.**

GRAY, C. A. 1993 Horizontal and vertical trends in the distribution of larval fishes in coastal waters off central New South Wales, Australia. *Marine biology,* **116:** 649–666.

GREENFIELD D. W., JOHNSON R. K. 1990. Community structure of western Caribbean blennioid fishes. *Copeia*, **1990**: 433–448.

GREER A.T., WOODSON C.B., GUIGAND C.M., COWEN R.K. 2016. Larval fishes utilize Batesian mimicry as survival strategy in the plankton. *Marine ecology progress series*, **551**: 1-12.

GRIFFITHS C.L., ROBINSON T.B., LANGE L., MEAD A. 2010. Marine Biodiversity in South Africa: An Evaluation of Current States of Knowledge. *PLoS ONE,* **5**(8): e12008.

GRIFFITHS M.H. 2000. Long term trends in catch and effort of commercial line fish off South Africa's Cape Province: Snapshots of the 20th Century. *South African journal of marine science,* **22**: 81–110.

GROSBERG R. K., LEVITAN D. R. 1992. For adults only? Supply-side ecology and the history of larval biology. *Trends in ecology and evolution*, **7**: 130–133.

GURTLER V., STANISICH V.A., 1996. New approaches to typing and identification of bacteria using the 16S–23S rDNA spacer. *Microbiology,* **142**: 3–16.

HAJIBABAEI M., SINGER G.A.C., HEBERT P.D.N, HICKEY D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetic and population genetics. *Trends genetics,* **23**(4): 167–172.

HARE J.A., QUINLAN J.A., WERNER F.E., BLANTON B.O., GOVONI J.J., FORWARD R.B., SETTLE L.R., HOSS D.E. 1999. Larval transport during winter in the Sabre study area: Results of a coupled vertical larval behavior-three-dimensional circulation model. *Fisheries Oceanography,* **8**(1): 57–76.

HARRIS S.A., CYRUS D.P. 1999. Composition, abundance and seasonality of fish larvae in the mouth of Durban Harbour, KwaZulu-Natal, South Africa, *South African journal of marine science*, **21**(1): 19-39.

HARRIS S.A., CYRUS D.P., BECKLEY L.E. 1999. The larval fish assemblage in nearshore coastal waters off the St Lucia estuary, South Africa. *Estuarine, coastal and shelf science,* **49**: 789-811.

HARRIS S.A., CYRUS D.P., BECKLEY L.E. 2001. Horizontal trends in larval fish diversity and abundance along an ocean-estuarine gradient on the Northern KwaZulu-Natal Coast, South Africa. *Estuarine, coastal and shelf science,* **53**: 221-235.

HASTINGS P.A. 2009. Biogeography of New World Blennies. In The Biology of Blennies. Edited by Patzner RA, Gonçalves EJ, Hastings PA, Kapoor BG. Enfield: Science Publishers, 95–118.

HASTINGS P.A., SPRINGER V.G. 2009. Systematics of the Blennioidei and the included families Dactyloscopidae, Chaenopsidae, Clinidae and Labrisomidae. In The Biology of Blennies. Edited by Patzner RA, Gonçalves EJ, Hastings PA, Kapoor BG. Enfield: Science Publishers, 3–30.

HEBERT P.D., GREGORY T.R. 2005. The promise of DNA barcoding for taxonomy. *Systematic biol*ogy, **54**: 852–859.

HEBERT P.D.N., CYWINSKA A., BALL S.L., DEWAARD J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the royal society*, **270**:313–321.

HEBERT P.D.N., PENTON E.H., BURNS J.D.H., HALLWACHS W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgreator*. *Proceedings of the national academy of sciences of the United States America*, **101**(41):14812–14817.

HEEMSTRA P.C., GON O., 1986. Family no.262: Soleidae. In: M.M. Smith and P.C. Heemstra (eds), *Smith's sea fishes,* 868-874. Macmillan, South Africa.

HEEMSTRA P.C., HEEMSTRA E. 2004. Coastal Fishes of Southern Africa. South African Institute for Aquatic Biodiversity and National Inquiry Service Centre, Grahamstown, 486.

HENDRY A.P., DAY T., COOPER A.B. 2001. Optimal size and number of propagules: allowance for discrete stages and effects of maternal size on reproductive output and offspring fitness. *The American naturalist,* **157**: 387-407.

HERBERT D.G. 2001. Museum natural science and the NRF: crisis times for practitioners of fundamental biodiversity science. *South African journal of science,* **97**: 168–172.

HICKEY A.J.R., CLEMENTS K.D. 2005. Genome size evolution in New Zealand triplefin fishes. *Journal of heredity,* **7**:356–362.

HILLIS D.M., BULL J.J. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic biology,* **42**: 182-192.

HOFRICHTER R., PATZNER R. A. 2000. Habitat and microhabitat of Mediterranean clingfishes (Teleostei: Gobiesociformes: Gobiesocidae). *PSZN Marine Ecology* **1**, 41–53.

HOLLEMAN W. 2005. A review of the triplefin fish genus *Enneapterygius* (Blennioidei: Tripterygidae) in the western Indian Ocean, with descriptions of four new species. *Smithiana, Publications in aquatic biodiversity, bulletin*, **5**: 1–25.

HOOPER D. U., CHAPIN F. S., EWEL J. J., HECTOR A., INCHAUSTI P., LAVOREL S., LAWTON J. H., LODGE D. M., LOREAU M., NAEEM S., SCHMID B., SETALA H., SYMSTAD A. J., VANDERMEER J., WARDLE D. A.. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*, **75**: 3–35.

HOPKINS G.W., FRECKLETON R.P. 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal conservation*, **5**: 245–249.

HORNE J.B., van HERWERDEN L., CHOAT J.H., ROBERTSON D.R. 2008. High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in the three reef fish congeners. *Molecular phylogenetics and evolution,* **49**: 629-638.

HUBERT N., DELRIEU-TROTTIN E., IRISSON J.-O., MEYE C., PLANES S. 2010. Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. Molecular phylogenetics and evolution, **55**: 1195–1203.

HUBERT N., DELRIEU-TROTTIN E., IRISSON J.O., MEYER C., PLANES S. 2010. Identifying early stages of coral reef fishes through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. *Molecular phylogenetics and evolution*, **55**: 1195–1203.

HUBERT N., HANNER R., HOLM E., MANDRAK N.E., TAYLOR E., BURRIDGE M., WATKINSON D.,DUMONT P., CURRY A., BENTZEN P., ZHANG J., APRIL J., BERNATCHEZ L. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One* **3** (6): 1–8.

HUBERT N., MEYER C.P., BRUGGEMANN H.J., GUÉRIN F., KOMENO R.J.L., ESPIAU B., CAUSSE R., WILLIAMS J.T., PLANES S. 2012. Cryptic Diversity in Indo-Pacific

Coral-Reef Fishes Revealed by DNA-Barcoding Provides New Support to the Centre-of-Overlap Hypothesis, *PLoS ONE*, **7** (3).

HULL F.M. 1970. Some new genera and species of bee flies from South America: Diptera Bombyliidae. *Journal of the Georgia Entomological Society,* **5**: 163–166.

HULLEY E.N., TAYLOR N.D.J., ZARNKE A.M., SOMERS C.M., MANZON J.Y.W., BOREHAM D.R. 2018. DNA barcoding vs. morphological identification of larval fish and embryos in Lake Huron: Advantages to a molecular approach*, Journal of great lakes research.*

HUTCHINGS L., VAN DER LINGEN C.D., GRIFFITHS M., ROBERTS M.R., BECKLEY L.E., SUNDBY S. 2002. Spawning on the edge: spawning grounds and nursery areas around the South African coast. *Marine and freshwater research*. **53**: 307–318

HYDE J. R., LYNN E., HUMPHREYS J.R., MUSYL M., WEST A. P., VETTER R. 2005. Shipboard identification of fish eggs and larvae by multiplex PCR, and description of fertilized eggs of blue marlin, short bill spearfish, and wahoo. *Marine ecology progress series*, **286**:269–277.

JAAFAR T.N.A.M., TAYLOR M.I., NOR S.A.M, BRUYN M., CARVALHO G.R. 2012. DNA barcoding reveals cryptic diversity within commercially exploited Indo-Malay Carangidae (Teleosteii: Perciformes). *PLoS One*, **7**: e49623.

JERDE C. L., MAHON A. R., CHADDERTON W. L., LODGE, D. M. 2011. “Sightunseen” detection of rare aquatic species using environmental DNA. *Conservation letters*, **4**: 150–157.

JOHN B.A., SHEIKH H.I., JALAL K. C. A., KAMARUZZAMAN B. Y., SANOWER H., HANISAH M. N., RAHMAN M. H., ROZIHAN M. 2018. DNA barcoding of ichthyoplankton and juvenile fishes of a tropical river in Malaysia. *Springer international publishing.*

JOHNSON M. P. 2005. Is there confusion over what is meant by ‘open population’? *Hydrobiologia*, **544**: 333–338.

JONES G.P., ANDREW N.L. 1993. Temperate reefs and the scope of seascape ecology. In: Creese RG, Schiel DR, Battershill CN, Jones GP, MacDiarmid AB eds. *Proceedings of the Second International Temperate Reef Symposium.* Wellington, New Zealand, NIWA Publications. 6376.

JONES G.P., PLANES S., THORROLD S.R. 2005. Coral reef fish larvae settle close to home. *Current Biology*, **15**(1): 314–318.

JOOSTE C.M., OLIVER J., EMAMI-KHOYI A., TESKE P.R. 2018. Is the Wild Coast in the eastern South Africa a distinct marine bioregion? *Helgoland marine research,* **72**:6.

KARAISKOU N., TRIANTAFYLLIDIS A., ALVAREZ P., LOPES P., GARCIA-VAZQUEZ E., TRIANTAPHYLLIDIS C. 2007. Horse mackerel egg identification using DNA methodology. *Marine ecology*, **28**:429-434.

KENDALL A.W., AHLSTROM E.H., MOSER H.G. 1984. Early life history stages of fishes and their characters. In G. G. Moser, W. J. Richards, D. M. Cohen, M. P. Fahay, A. W. Kendall, & S. L. Richardson (Eds.), Ontogeny and systematics of fishes (pp.11- 22). Kansas: *American Society of Ichthyologists and Herpetologists.*

KENDALL A.W., AHLSTROM E.H., MOSER H.G. 1984. Early life history stages of fishes and their characters. In G. G. Moser, W. J. Richards, D. M. Cohen, M. P. Fahay, A. W. Kendall, & S. L. Richardson (Eds.), Ontogeny and systematics of fishes Kansas. *American society of ichthyologists and herpetologists,* 11-22.

KIMURA M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution,* **16**: 111-120.

KLOPPER R.R., SMITH G.F., CHIKUNI A.C. 2002. The Global Taxonomy Initiative in Africa. *Taxon,* **51**: 159–165

KNOWLTON N. 1993. Sibling species in the sea. *Annual review of ecology and systematics*, **24**: 189–216.

KO H.L., WANG Y.T., CHIU T.S., LEE M.A., LEU M.Y., CHANG K.Z., CHEN W.Y., SHAO K.T. 2013. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS ONE*, **8**(1): 53451.

KOCHZIUS M. (2009.) Trends in fishery genetics. In: Beamish RJ, Rothschild BJ,

KRESS W. J., ERICKSON D. L. 2008. DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the national academy of sciences of the United States of America*, **105**(8): 2761-2.

KRUGER M., STRYDOM N.A. 2010. Spatial and temporal variability in the larval fish assemblage of a warm temperate South African estuary, with notes on the effects of artificial channelling. *African zoology*, **45**(2): 195-212.

KUMAR S., STECHER G., PETERSON D., TAMURA K. 2012. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. *Bioinformatics,* **28**: 2685–2686.

KUMAR S., STECHER G., PETERSON D., TAMURA K. 2012. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. B*ioinformatics*, **28**:2685–2686.

LAMBERT T. C., WARE D. M. 1984. Reproductive Strategies of Demersal and Pelagic Spawning Fish. *Canadian journal of fisheries and aquatic sciences*, **41**(11): 1565–1569.

LANGLOIS T.J., ANDERSON M.J., BABCOCK R.C. 2005. Reef associated predators influence adjacent soft sediment communities. *Ecology*, **86**: 1508-1519.

LARKIN M.A., BLACKSHIELDS G., BROWN N.P., CHENNA R., MCGETTIGAN P.A., MCWILLIAM H., HIGGINS D.G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* **23**: 2947-2948.

LEIS J. M. 1991. The pelagic stage of reef fishes: the larval biology of coral reef fishes. Academic Press Inc, San Diego, California, 183– 230.

LEIS J. M., CARSON-EWART B. M. 2000. The Larvae of Indo-Pacific Coastal Fishes: An Identification Guide to Marine Fish Larvae. *Leiden, the Netherlands: Brill publisher*.

LEIS J.M. 2006. Are larvae of demersal fishes plankton or nekton? Ichthyology, Australian, Sydney, Australia.

LEIS J.M. 2015. Taxonomy and systematics of larval Indo-Pacific fishes: a review of progress since 1981. *Ichthyology research,* **62**: 9-28.

LEIS J.M., SIEBECK U.E., HAY A.C., PARIS C.B., CHATEAU O., WANTIEZ L. 2015. *In situ* orientation of fish larvae can vary among regions. *Marine ecology progress series,* **537**: 191-203.

LEIS J.M., TRNSKI T. 1989. The larvae of Indo-Pacific shore fishes. New South Wales University Press, Sydney, **371**.

LELIEVRE S., JEROME M., MAES G.E., VAZ S., CALAIVANY S., VARREZ-BAGNIS V. 2012. Integrating molecular identification of pelagic eggs with geostatistical mapping to improve the delineation of North Sea fish spawning grounds. *Marine ecology progress series*, **445**: 161-172**.**

LIN C.P., DANFORTH B.N. 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined data sets. *Molecular phylogenetics and evolution,* **30**(3): 686–702.

LIN H.C., HASTINGS PA. 2013. Phylogeny and biogeography of a shallow water fish clade (Teleostei: Blenniiformes). *BMC evolutionary biology,* **3**(13): 210.

LONGENECKER K., LANGSTON R. 2005. Life history of the Hawaiian blackhead triplefin, *Enneapterygius artriceps*. *Environmental biology of fishes,* **73**: 243-251.

MABRAGAÑA E, DÍAZ DE ASTARLOA JM, HANNER R, ZHANG J, GONZÁLEZ CASTRO M. 2011. DNA Barcoding Identifies Argentine Fishes from Marine and Brackish Waters. *PLoS ONE*, **6**(12): e28655.

MACE G.M. 2004. The role of taxonomy in species conservation. *Philosophical transactions of the royal society,* **359**: 711–719.

MALLET J., WILLMOTT K. 2003. Taxonomy: renaissance or tower of babel? *Trends in ecology and evolution,* **18**(2): 57-59.

MARDIS E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends in genetics*, **24**: 133–141.

MATARESE A. C., SPIES I. B., BUSBY M. S., ORR J. W. 2011. Early larvae of *Zesticelus profundorum* (family Cottidae) identified using DNA barcoding. *Ichthyological research*, **58**(2): 170–174.

MAYDEN, R. L. 1997. A hierarchy of species concepts: the denouement of the species problem *in* The Units of Biodiversity - Species in Practice Special Volume 54 (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Systematics Association.

MAYR E. 1942. Systematics and the origin of species from the viewpoint of a zoologist.

MAYR E. 1982. The Growth of Biological Thought. Diversity, Evolution, and Inheritance. Belknap, Cambridge, MA.

MAYR E., 2000. In: Wheeler, Q.D., Meier, R. (Eds.), The biological species concept in species concepts and phylogenetic theory: a debate. Columbia University Press.

MEIER R., SHIYANG K., VAIDYA G., PETER K. L. 2006. DNA Barcoding and Taxonomy in Diptera: A Tale of High Intraspecific Variability and Low Identification Success, *Systematic biology*, **55**(5): 715–728

MEYER C., PAULAY G. 2005. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, **3**(12).

MEYER C.P. 2003. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society,* **79**: 401–459.

MILLER B.S., KENDALL A.W. 2009. Early Life History of Marine Fishes. University of California Press.

MINAMOTO T., YAMANAKA H., TAKAHARA T., HONJO M. N., KAWABATA Z. 2012. Surveillance of fish species composition using environmental DNA. *Limnology*, **13**: 193–197.

MORA C., TITTENSOR D.P., MAYERS R.A. 2008. The completeness of taxonomic inventories for describing the global diversity and distribution of fishes. *Proceedings of the royal society B: biological sciences*, **275**: 149–155.

MORANTA J., MASSUTI E., PALMER M., GORDON J.D.M. 2007. Geographic and bathymetric trends in abundance, biomass and body size of four grenadier fishes along the Iberian coast in the western Mediterranean. *Progress in oceanography*, **72**: 63–83.

MOSER H.G., AHLSTROM E.H. 1970. Development of lantern fishes (family Myctophidae) in the California current. Part I. species with a narrow-eyed larvae. *Bulletin of the Los Angeles county museum,* **7**: 145.

MOSER H.G., RICHARDS W.J., COHEN D.M., FAHAY M.P., KENDALL A.W., RICHARDSON S.L. 1984. Ontogeny and systematics of fishes, special publication No.1. *American society of ichthyologists and herpetologists*, Lawrence, Kansas.

MUNCH K., BOOMSMA W., HUELSENBECK J.P., WILLERSLEV E., NIELSEN R. 2008. Statistical assignment of DNA sequences using Bayesian phylogenetics. *Systematic biology*, **57**(5):750–757.

MUNDAY P.L., JONES G.P. 1998. The ecological implications of small body size among coral-reef fishes. *Oceanography and marine biology an annual review,* **36**: 373-411.

MWALUMA, J.M., KAUNDA-ARARA, B., STRYDOM, N. 2014. A guide to commonly occurring larval stages of fishes in Kenyan coastal waters. WIOMSA Book Series No. **15**:73.

NEIRA F. J., PERRY R. A., BURRIDGE C. P., LYLE J. M., KEANE J. P. 2015. Molecular discrimination of shelf-spawned eggs of two co-occurring *Trachurus* spp. (Carangidae) in south eastern Australia: a key step to future egg-based biomass estimates. *ICES Journal of marine science*, **72:** 614–624.

NEIRA F.J., MISKIEWICZ A.G., TRNSKI T. 1998. Larvae of temperate Australian fishes: laboratory guide for larval fish identification, UWA publishing.

NELSON J.S. 1994. Fishes of the world – third edition. Upper Saddle River: prentice-hall.

NELSON J.S. 2006. Fishes of the world, 4th ed. *John Wiley and sons*, New Jersey.

NELSON J.S., GRANDE T.L., WILSON M.V.H. 2016. Fishes of the world. Fifth Edition. John Wiley & Sons, Hoboken, NJ, USA, xli + 707.

NIXON K. C., WHEELER Q. D. 1990. An amplification of the phylogenetic species concept. *Cladistics*, **6**(3): 211–223.

OCKHUIS S., HUGGETT J.A., GOUWS G., SPARKS C. 2017. The 'suitcase hypothesis': Can entrainment of meroplankton by eddies provide a pathway for gene flow between Madagascar and KwaZulu-Natal, South Africa? *African journal of marine science*, **39**(4): 435-451.

OHL M. 2014. Principles of taxonomy and classification: current procedures for naming and classifying organisms. In: Henke W, Hardt T, Tattersall I (eds), *Handbook of Palaeoanthropology* Springer-Verlag, Heidelberg, 141-166.

OVERDYK L.M., HOLM E., CRAWFORD S.S., HANNER R.H. 2016. Increased taxonomic resolution of Laurentian Great Lakes ichthyoplankton through DNA barcoding: A case study comparison against visual identification of larval fishes from Stokes Bay, Lake Huron. *Journal of great lakes research*.

PACKER L., GIBBS J., SHEFFIELD C., HANNER R. 2009. DNA barcoding and the mediocrity of morphology. *Molecular ecology research,* **9**: 42–50.

PALUMBI S.R. 1996. Nucleic acids II: the polymerase chain reaction. *Molecular systematics,* **2**: 205–247.

PATTRICK P., STRYDOM N.A. 2008. Composition, abundance, distribution and seasonality of larval fishes in the shallow nearshore of the proposed Greater Addo Marine Reserve, Algoa Bay, South Africa. *Estuarine, coastal and shelf science,* **79**: 251-262.

PATTRICK P., STRYDOM N.A. 2014. The effects of exposure in sandy beach surf zones on the early life history stages of fishes within Algoa Bay, South Africa. *Journal of Fish Biology*, **84**: 1354-1376.

PATTRICK P., STRYDOM, N. A. 2016. Variability in spatial and temporal occurrence of presettlement and settlement-stage fishes associated with shallow reefs. *Journal of Fish Biology*, **90**(3): 847–866.

PAYNE A.I.L., PUNT A.E. 1995. Biology and fisheries of South African Cape hakes (M. capensis and M. paradoxus).

PEGG G.G., SINCLAIR B., BRISKEY L., ASPDEN W.J. 2006. MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia. *Scientia marina*, 7–12.

PENAZ M. 2001: A general framework of fish ontogeny: a review of the ongoing debate. *Folia zool*ogica, **50**(4): 241–256.

PINCEEL J., JORDAENS K., BACKELJAU T. 2005. Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *Journal of evolutionary biology*, **18**: 1264 -1280.

PINEDA J., HARE J.A., SPONAUGLE S.U. 2007. Larval transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography,* **20**(3): 22-39.

PONS J., BARRACLOUGH T.G., GOMEZ-ZURITA J., CARDOSO A., DURAN D.P., HAZELL S., VOGLER A.P. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic biology*, **55**(4):595–609.

POWLES H., MARKLE D.F. 1984. Identification of larvae. In: H.G. Moser, W.J. Richards, D.M. Cohen, M.P. Fahay, A.W. Kendall & S.L. Richardson, eds. Ontogeny and Systematics of Fishes. *American society of ichthyologists and herpetologists*, 31–33.

PUILLANDRE N, LAMBERT A, BROUILLET S, ACHAZ G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular biology,* **21**(8):1864–1877.

PUNCHER G.N., ARRIZABALAGA H., ALEMANY F., CARIANI A., ORAY I.K., KARAKULAK F.S., BASOLONE G., CUTTITTA A., MAZZOLA S., TINTI F. 2015. Molecular identification of Atlantic bluefin tuna (*Thunnus thynnus*, Scombridae) larvae and development of a DNA character-based identification key for Mediterranean scombrids. *PLoS ONE,* **10**(7): e0130407.

RALSTON S., BENCE J.R., ELDRIDGE M.B., LENARZ W.H. 2003. *An approach to estimating rockfish biomass based on larval production, with application to Sebastes jordani. Fishery bulletin*, **101**(1): 129-146.

RANDALL J.E., ALLEN G.R., STEEN R.C. 1997. Fishes of the Great Barrier Reef and Coral Sea. *Crawford house publishing Pty Ltd,* Bathurst.

RANDALL J.E., VICTOR B.C. 2013. *Bodias atrolumbus* (Valenciennes 1839), a valid species of labrid fish from the southwest Indian Ocean. *Journal of the ocean science foundation,* **8**: 44-61.

RANDALL J.E., VICTOR B.C. 2013. *Bodias atrolumbus* (Valenciennes 1839), a valid species of labrid fish from the southwest Indian Ocean. *Journal of the ocean science foundation,* **8**: 44-61.

RATNASINGHAM S., HEBERT P.D.N. 2007. BOLD: the barcode of life data System. *Molecular ecology notes*, **7**: 355–364.
RATNASINGHAM S., HEBERT P.D.N. 2007. BOLD: the barcode of life data System. *Molecular ecology notes*, **7**: 355–364.
RATNASINGHAM S., HEBERT P.D.N. 2013. A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS ONE,* **8**: e66213.

RICHARDSON D.E., COWEN R.K. 2004. Diversity of *leptocephalus* larvae around the island of Barbados (West Indies): relevance to regional distributions. *Marine ecology progress series,* **282**: 271–284.

RICHARDSON D.E., VANWYE J.D., EXUM A.M., COWEN R.K., CRAWFORD D.L. 2006. High through put species identification: from DNA isolation to bioinformatics. *Molecular ecology notes,* **7:** 199–207.

RICHARDSON D.E., VANWYE J.D., EXUM A.M., COWEN R.K., CRAWFORD D.L. 2007. High through put species identification: from DNA isolation to bioinformatics. *Molecular ecology notes*, **7**: 199–207.

RICHARDSON T. J., POTTS W. M., SANTOS C. V., SAUER, W.H.H. 2011. Comparison of the population structure and life-history parameters of *Diplodus capensis* (Sparidae) in exploited and unexploited areas of southern Angola. *African journal of marine science* **33**: (2) 191–201.

RIDLEY M. 2004. Evolution, 3rd edition. Blackwell publishing, Ltd.

RIGINOS C. AND VICTOR B. C. 2001. Larval spatial distribution sand other early life-history characteristics predict genetic differentiation in eastern Pacific blennioid fishes. *Proceedings of the royal society,* **268**: 1931-1936.

RIGINOS C., DOUGLAS K. E., JIN Y., SHANAHAN D. F., TREML E. A. 2011. Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, **34**: 566–575.
RIGINOS C., DOUGLAS K. E., JIN Y., SHANAHAN D. F., TREML E. A. 2011. Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, **34**: 566–575.
ROBERTSON A.G.D. 1994. Fishes of the tropical eastern pacific. Honolulu, HI: *University of Hawaii Press.*

RODRIGUEZ J.M., ALEMANY F., GARCIA A. 2017. A guide to the eggs and larvae of 100 common Western Mediterranean Sea bony fish species. FAO, Rome, Italy, 256.

ROSEN D. E. 1979. Fishes from the uplands and inter-montane basin of Guatemala: Revisionary studies and comparative geography, *Bulletin of the American museum of natural history*, **162**: 267–376.

SABATE´S A., ZABALA M., GARCI´A-RUBIES A. 2003. Larval fish communities in the Medes Islands Marine Reserve (North-west Mediterranean). *Journal of plankton research,* **24**: 1035–1046.

SACCONE C. GIORGI C.D. GISSI C. PESOLE G. REYES A. 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene,* **238**: 195–209.

SACHITHANANDAM V., MOHAN P. 2018. A review on DNA barcoding on Taxonomy in India. *DNA barcoding and molecular phylogeny.*

SAITOU N., NEI M. 1987. The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular biology and evolution,* **4**: 406-425.

SALE P. F. 1970. Distribution of larval Acanthuridae off Hawaii. *Copeia,* 765–766.

SALE P.F. 1991. Reef fish communities: open non-equilibrium systems. 564-598.

SAUNDERS G. 2005. Applying DNA barcoding to red macro-algae, a preliminary appraisal holds promise for future applications. *Philosophical transactions the royal society series biology*, **360**: 1879–1888.

SCHELTEMA R. S. 1986. Long-distance dispersal by planktonic larvae of shoal-water benthic invertebrates among Central Pacific islands. *Bulletin of marine science*, **39**: 241–256.

SCHLOSS P.D. 2010. The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS computational biology,* **6**: e1000844.

SCHWARZ G. 1974. Estimating the dimension of a model. *Annals of statistics*, **6**:461–464.

SEBERG O., HUMPHRIES C.J, KNAPP S., STEVENSON D.W. PETERSEN G., SCHARFF N., ANDERSEN N.M. 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. *Trends in ecology & evolution*. **18**: 63–65.

SERAFY J. E., FAUNCE C. H., LORENZ J. L. 2003. Mangrove shoreline fishes of Biscayne Bay, Florida. *Bulletin of marine science,* **72:** 161–180.

SHAO K.T., CHEN K.C., WU J.H. 2002. Identification of marine fish eggs in Taiwan using light microscopy, scanning electric microscopy and mtDNA sequencing. *CSIRO*, **53**: 355–65.

SIMON T.P., LYONS J. 1995. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. In: Davis, W.S., Simon, T.P. (Eds.),

Biological Assessment and Criteria: tools for water resource planning and decision-making. *Lewis Publishers*, Boca Raton, 245–262.

SINCLAIR M. 1988. Marine Populations: An essay on population regulation and speciation. Washington Sea Grant Program, Seattle, WA, **252**.

SMITH M.M. HEEMSTRA P.C., SMITH M.M. 2003. Smith's sea fishes. 2nd edition. Cape Town Publishers.

SMITH M.M., HEEMSTRA P.C. 2003. Smiths' sea fishes, 1st edition. Struik Publishers, Cape Town.

SONG H., BUHAY J.E., WHITING M.F., CRANDALL K.A. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are co-amplified. *Proceedings of the National Academy of sciences of the United States of America,* **105**: 13486–13491.

SPRINGER V. 1998. Blennies. In: W Eschmeyer, J Paxton, eds. Encyclopedia of Fishes – second edition. San Diego, CA: *Academic press*, 216.

STACKEBRANDT E., EBERS J. 2006. Taxonomic parameters revisited: tarnished gold standards. *Microbiology today*, **33**:152–155.

STACKEBRANDT, E., RAINEY, F.A., 1995. Partial and complete 16S Wang, G., VAN Dam, A.P., Le Fleche, A., Postic, D., Peter, O., rDNA sequences, their use in generation of 16S phylogenetic trees and their implications in molecular ecological studies. In: Molecular microbial ecology manual, Kluwer Dordrecht, 1–17.

STEINKE D., CONNELL A.D., HEBERT P.D.N. 2016. Linking adults and immatures of South African marine fishes. *NRC research press*, **59**: 959-967.

STEINKE D., ZEMLAK T.S., BOUTILLIER J.A. 2009. DNA barcoding Pacific Canada's fishes. *Marine biology,* **156**: 2641–2647.

STEPIEN C.A., DILLON A.K., BROOKS M.J., CHASE K.L., HUBERS A.N. 1997. The evolution of Blenniod fishes based on an analysis of mitochondrial 12S rDNA. *Molecular systematic of fishes,* 245.

STEPIEN C.A., DIXON M.T., HILLIS D.M. 1993. Evolutionary relationships of the blennioid fish families Clinidae, Labrisomidae and Chaenopsidae: congruence between DNA sequence and allozyme data. *Bulletin of marine science*, **52**(1): 496-515.

STRATHMANN R.R., HUGHES T.P., KURIS M.A., LINDEMAN K.C., MORGAN G.S., PANDOLFI J.M. WARNER R.R. 2002. Evolution of local recruitment and its consequences for marine populations. *Bulletin of marine science*, **70**(1): 377-396.

STRYDOM N.A. 2003. Occurrence of larval and early juvenile fishes in the surf zone adjacent to two intermittently open estuaries, South Africa. *Environmental biology of fishes*, **66**: 349-359.

STRYDOM N.A., COETZER C.J., PATTRICK P. 2015. The complex early life history of a marine estuarine opportunist fish species, *Solea turbynei* (Soleidae) from temperate South Africa. *Scientia marina*, **79**(2): 169178.

STUESSY T.F. 1972. A revision of the genus Melampodium (Compositae: Heliantheae). *Rhodora,* **74**:1–219.

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

SWEARER S.E., SHIMA J.S., HELLBERG M.E., THORROLD S.R., JONES G.P., ROBERTSON D.R., MORGAN S.G., SELKOE K.A., RUIZ G.M., WARNER R.R. 2002. Evidence of self-recruitment in demersal marine populations. *Bulletin of marine science*, **70**: 251–271.

SYMS C. 1995. Multi-scale analysis of habitat association in a guild of blennioid fishes. *Marine ecology progress series*, **125**: 31-43.

TAMURA K., STECHER G., PETERSON D., FILIPSKI A., KUMAR S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution*, **30**: 2725–2729.

TAUTZ D., ARCTANDER P., MINELLI A., THOMAS R.H. VOGLER A.P. 2003. A plea for DNA taxonomy. *Trends in ecology & evol*ution, **18**: 70–74.

TAYLOR H. R., HARRIS W. E. 2012. An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. *Molecular ecology resources*, **12**: 377– 388.

TERMVIDCHAKORN A., K.G. HORTLE. 2013. A guide to larvae and juveniles of some common fish species from the Mekong River Basin. MRC technical paper no.38. Mekong River Commission, Phnom Penh.

THE EUROPEAN UNION AND THE SOUTHERN AFRICAN DEVELOPMENT COMMUNITY. 2017. South African fisheries and the SACD-EU economic partnership agreement. Available at https://sadc-epa-outreach.com/images/files/sadc-eu-epa-fisheries-july-2017.pdf.

THOMPSON E.F., STRYDOM N.A., HECHT T. 2007. Larval development of Dagetichthys marginatus (Soleidae) obtained from hormone induced spawning under artificial conditions. *Scientia marina*, **71**(3): 421–428.

THOMPSON S. 1983. Homing In a territorial reef fish. *Copeia*, 832–83.

THOMSEN P. F., KIELGAST J., IVERSEN L. L., WIUF C., RASMUSSEN M., GILBERT M. T. P., ORLANDO L., WILLERSLEV E. 2012. Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular ecology*, **21**: 2565–2573.

TIMMERMANS M. J. T. N., DODSWORTH S., CULVERWELL C. L., BOCAK L., AHRENS D., LITTLEWOOD D. T. J., PONS J., VOGLER A. P. 2010. Why barcode? High throughput multiplex sequencing of mitochondrial genomes for molecular systematics. *Nucleic acids research*, **38**:e197.

TURPIE J.K., BECKLEY L.E., KATUA S.M. 2000. Biogeography and the selection of priority areas for the conservation of South African coastal fishes. *Biological conservation*, **92**: 59–72.

URHO L. 2000. Characters of larvae – what are they? *Folia zoology,* **51**(3): 161–186.

VALDEZ-MORENO M., VÁSQUEZ-YEOMANS L., ELÍAS-GUTIÉRREZ M., IVANOVA N.V., HEBERT P.D.N. 2010. Using DNA barcodes to connect adults and early life stages of marine fishes from the Yucatan Peninsula, Mexico: potential in fisheries management. *Marine and freshwater research*, **61**(6): 655–71.

VAN DER LINGEN C.D., HUGGET J.A. 2003. The role of ichthyoplankton surveys in recruitment research and management of South African anchovy and sardine. *The Big Fish Bang, Proceedings of the 26th Annual larval fish conference.*

VAN VALEN L. 1976. Ecological species, multispecies and oaks. *Taxon*, **25:** 233–239.

VERNOOY R., HARIBABU E., MULLER M.R., VOGEL J.H., HEBERT P.D.N., SCHINDEL D.E., SHIMURA J., SINGER G.A.C. 2010. Barcoding Life to Conserve Biological Diversity: Beyond the Taxonomic Imperative. *PLoS ONE,* **8**(7): e1000417.

VICTOR B.C. 1986. Duration of the planktonic larval stage of one hundred species of Pacific and Atlantic wrasses (family Labridae). *Marine biology*, **90**: 317–326.

VICTOR B.C. 2009. Identification of the larval and juvenile stages of the Cubera Snapper (Lutjanus cyanopterus) using DNA barcoding. *Zootaxa*, **36**(2215): 24–36.

VICTOR B.C. 2013. The Caribbean roughhead triplefin (*Enneanectes boehlkei*): DNA barcoding reveals a complex of four West Indian sympatric cryptic species (Teleostei: Blenniodei: Tripterygiidae). *Journal of the ocean science foundation,* **7**: 44–73.

VICTOR B.C. 2015. How many coral reef fish species are there? Cryptic diversity and the new molecular taxonomy, In: Mora C. (Ed.), Ecology of Fishes on Coral Reefs, Cambridge University Press, Cambridge.

VICTOR B.C., HANNER R., SHIVJI M., HYDE J., CALDOW C. 2009. Identification of the larval and juvenile stages of the Cubera snapper, *Lutjanus cyanopterus,* using DNA barcoding. *Zootaxa*, **2215**: 24–36.

VON DER HEYDEN S., BOWIE R. C. K., PROCHAZKA K., BLOOMER P., CRANE N. L., BERNARDI G. 2011. Phylogeography patterns and cryptic speciation across oceanographic barriers in South African intertidal fishes. *Evolutionary biology,* **24:** 2505–2519.

WARD R.D. 2009. DNA Barcoding: DNA barcode divergence among species and genera of birds and fishes. *Molecular ecology resources*, **9**: 1077–1085.

WARD R.D., ZEMLAK T.S., INNES B.H., LAST P.R., HEBERT P.D.N. 2005. DNA barcoding Australia's fish species, *Philosophical transactions of the royal society B*, **360**: 1847–1857.

WARD, R.D., HANNER, R., & HEBERT, P.D. 2009. The campaign to DNA barcode all fishes. *Journal of Fish Biology,* **74**: 329–356.

WEBB J.F. 1999. Larvae in fish development and evolution. In B. K. Hall, &M. H. Wake (Eds.), the origin and evolution of larval forms. San Diego, CA: *Academic Press*. 109–158.

WEBB K.E., BARNES D.K.A.A., CLARK M.S., BOWDEN D.A. 2006. DNA barcoding: a molecular tool to identify Antarctic marine larvae. *Deep Sea Research Part II: Topical studies in oceanography,* **53**: 1053–1060.

WEIGT L.A., DRISKELL A.C., BALDWIN C.C., ORMOS A. 2012. DNA barcoding fishes. *Methods in molecular biology*, **858**: 109–126.

WHEELER A. 1985. *The World Encyclopedia of Fishes - second edition*. London: Macdonald.

WILBUR H.M. 1980. Complex life cycles. *Annual review of ecology and systematics*, **8**: 67-93.

WILEY E. O. 1978. "The Evolutionary Species Concept Reconsidered". *Systematic zoology,* **27**.1: 17.

WILSON E.O. 2004. Taxonomy as a fundamental discipline. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 739.

WILSON S.K. 2000. Trophic status and feeding selectivity of blennies (Blenniidae: Salariini). *Marine biology*, **136**: 431–437.

WINSTON J.E. 1999. Describing species: practical taxonomic procedure for biologists. *Columbia University Press,* New York.

WOOD A.D. 2000. The description of *Austroglossus pectoralis* (Teleostei: Soleidae) larvae from the south-east coast of South Africa. *Scientia Marina*, **64**(4): 387-392.

WORLD CONSERVATION MONITORING CENTRE. 1992. Development of a national biodiversity index: A discussion paper. Cambridge: World Conservation Press.

ZAPATA F.A., ROBERTSON D.R. 2007. How many species of shore fishes are there in the Tropical Eastern Pacific? *Journal of biogeography*, **34**: 38–51.

ZHANG J. B., HANNER R. 2011. DNA barcoding is a useful tool for the identification of marine fishes from Japan. *Biochemical systematics and ecology*, **39**(1): 31–42.

ZHANG J., HANNER R. 2012. Molecular Approach to the Identification of Fish in the South China Sea, *PLoS ONE*, **7**(2): e30621.

ZSILAVECZ G. 2005. Coastal fishes of the Cape Peninsula and False Bay: a divers' identification guide. Southern Underwater Research Group, Cape Town.