An investigation of the population connectivity of sardines (*Sardinops sagax*) of the KZN sardine run using meristic, morphological and genetic data

BRENT CHIAZZARI

Submitted in fulfilment of the academic requirements
for the degree of
MASTER OF SCIENCE

In the School of Life Sciences
Faculty of Science and Agriculture
University of KwaZulu-Natal
Durban

April 2014

Supervisors:
Dr Angus H.H. MacDonald
Dr Sean O'Donoghue

Abstract

The Sardine run occurs annually when large schools of sardine (*Sardinops sagax*) move from the Agulhas Bank towards KwaZulu-Natal, and has significant ecological and anthropogenic importance. Recent investigation has highlighted the nature and mechanisms resulting in the sardine run, however, critical questions about why the sardine run occurs remain unanswered. Therefore, the aim of this project was to elucidate the population diversity, connectivity and structure of sardines undertaking the sardine run. Sardines were sampled at four sites along the South African coast, and their morphology assessed using meristic data, multivariate, and geometric morphometrics. Nine exon-primed, intron-crossing (EPIC) DNA markers and the mitochondrially encoded cytochrome oxidase I (mtCOI) region of DNA were used for population and phylogeographic genetic analyses. Morphological analyses revealed significant differences between head size and shape of sardine run stock compared with other regions, and supports the delineation of a western, southern and eastern South African stock. Phylogeographic analysis using cytochrome oxidase I data, supported the idea that the *Sardinops* genus is monotypic. Genetic analyses using EPIC data confirmed low levels of segregation between sardines from the sardine run and the Western Cape stock. However, larvae spawned in KwaZulu-Natal demonstrated moderate levels of isolation from the Western Cape stock. The results reveal that there is successful recruitment of KwaZulu-Natal juveniles to the adult stock undertaking the sardine run. KwaZulu-Natal juveniles also recruit to the Western Cape population, although, to a lower degree. Results suggest sardines from the West Coast and Agulhas Bank partake in the sardine run. However genetic evidence suggests a certain subpopulation of the Agulhas Bank and a sub-stock of the Western Cape stock spawn successfully in KwaZulu-Natal. These results support the hypothesis that the sardine run represents a subpopulation spawning migration of *Sardinops sagax* in South Africa.

Key words: *Sardinops sagax*, KZN sardine run, morphology, genetic connectivity, stock structure

Preface

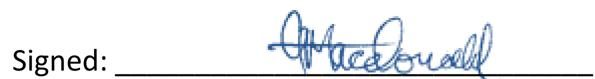
The work described in this dissertation was carried out at the School of Life Sciences at the University of KwaZulu-Natal, Westville. Field work was conducted at various sites along the KwaZulu-Natal, Eastern Cape and Western Cape coastline, under the supervision of Drs Angus H.H. MacDonald and Sean O'Donoghue. This dissertation represents original work by the author and has not otherwise been submitted in any form for any degree or diploma to any tertiary institute. Where use has been made of the work of others, it is duly acknowledged in the text.

I certify that the above statement is correct:

Signed: ________________________________

Brent Chiazzari; April 2014

As the candidate's supervisor, I agree with the above, and have approved this dissertation for submission:

Signed: ________________________________


Dr Angus H.H. MacDonald, Supervisor; April 2014

Signed: ________________________________

Dr Sean O'Donoghue, Supervisor; April 2014

Declaration

Plagiarism

I, Brent Chiazzari, declare that:

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

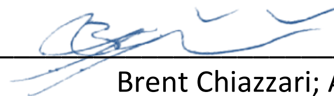
2. The thesis has not been submitted for any degree or examination at any other university.

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

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. Their words have been re-written but the general information attributed to them has been referenced.

    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

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

Signed:  ______________________

Brent Chiazzari; April 2014

Publications

Detail of contribution to publications that form part and/or include research presented in this thesis.

Publication 1 (in preparation):

B. Chiazzari, S. O'Donoghue and A.H.H. MacDonald. The population connectivity of the KZN Sardines (*Sardinops sagax*) using meristic, morphological and genetic data.

Author contributions:

Brent Chiazzari, Angus H.H. MacDonald and Sean O'Donoghue conceived the paper. Brent Chiazzari collected and analysed the data, Angus H.H. MacDonald provided financial support for the study, and both Angus H.H. MacDonald and Sean O'Donoghue contributed comments on the manuscript.

Publication 2 (in preparation):

B. Chiazzari, S. O'Donoghue and A.H.H. MacDonald. The optimisation and use of EPIC primers in stock identification of sardine (*Sardinops sagax*) in South Africa; a cheap and reliable alternative to sequence data analyses.

Author contributions:

Brent Chiazzari and Angus H.H. MacDonald conceived the paper. Brent Chiazzari collected and analysed the data, Angus H.H. MacDonald provided financial support for the study, and both Angus H.H. MacDonald and Sean O'Donoghue contributed comments on the manuscript.

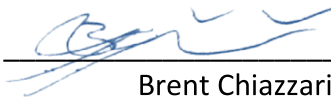
Other research outputs

Poster presentation:

B. Chiazzari, S. O'Donoghue and A.H.H. MacDonald. October 2013. The population connectivity of the KZN sardine, *Sardinops sagax*, using meristic, morphological and genetic data. Eighth WIOMSA scientific symposium, Maputo.

Oral presentation:

B. Chiazzari, S. O'Donoghue and A.H.H. MacDonald. July 2015. The population connectivity of the KZN sardine, *Sardinops sagax*, using meristic, morphological and genetic data. Southern African marine science symposium 2014, Stellenbosch.

Signed: 

Brent Chiazzari; April 2014

Table of contents

Abstract ii

Preface iii

Declaration iv

Plagiarism iv

Publications v

Other research outputs vi

Acknowledgements 1

Chapter one: 2

General introduction 2

1.1. The sardine run: an overview 2

1.2. Ecological phenomenon: The sardine run and it's fishery 4

1.3. Proximate and ultimate factors: current hypotheses about the sardine run 6

1.3.1. Ultimate factors 6

1.3.2. Proximate factors 11

1.4. Motivation and rationale for this study 12

1.4.1. Population genetic and morphological techniques in population delineation 12

1.4.2. Conclusions 15

1.4.3. Arrangement of dissertation 17

Chapter two: 18

Materials and methods 18

2.1 Study sites and sample collection 18

2.2 Meristic data collection and analysis 21

2.3. Morphological data collection and analysis 22

2.3.1. Material collection and preparation 22

2.3.2. Multivariate and Geometric analysis 23

2.4 Genetic material, primer and data collection, and analysis 26

2.4.1. Primer selection 26

2.4.2. Material collection and preparation 29

2.4.3. EPIC encoded data collection and analysis 30

2.4.4. mtCOI data collection and analysis 31

Chapter three: 35

Morphological characteristics of the sardine (*Sardinops sagax*) in South Africa 35

3.1. Introduction 35

3.2 Materials and methods 36

3.3. Results 37

3.3.1 Meristic results *37*

3.3.2. Morphometric results *42*

3.4. Discussion 51

Chapter four: 54

Genetic population structure of the KZN sardine (*Sardinops sagax*) using mitochondrial and exon-primed, intron-crossing (EPIC) markers 54

4.1. Introduction 54

4.2. Materials and methods 57

4.3. Results 57

4.3.1. Population genetic analyses, using EPIC primers *57*

4.3.2. Phylogenetic analyses, using mtCOI genetic data *63*

4.4. Discussion 69

4.4.1 Population structure of KZN Sardinops sagax using EPIC genetic data *69*

4.4.2 Analysis of Sardinops sagax in South African waters using mtCOI genetic data *72*

Chapter five: 77

Conclusions 77

References 81

Appendix 1 92

Appendix 2 92

Appendix 3 93

Appendix 4 94

Acknowledgements

First and foremost, I would like to thank my supervisors. Dr Angus H.H. MacDonald, who has gone above and beyond to mentor and guide me through the sea of pipettes and fish scales. Dr MacDonald has supplied me with the insights into learning the ins and outs of molecular ecology over the past few years. I am also greatly indebted to Dr Sean O'Donoghue, my co-supervisor, who throughout the last two years has supplied me with valuable personal insights, critiques, contacts and guidance specific to the sardine run. Their understanding and mutual interest in water sports has been valuable when the sea was good and my work ethic intermittent.

I would like to thank the members of the MAREVOL laboratory, who have offered guidance and support when needed. Thanks go to Rogan Harmer, for sharing project field trips with me, and aiding in the collection of samples in the Transkei. I am indebted to Roy Jackson, who offered the final shove when it came to Transkei free diving in limited visibility, with sardine and other marine mega fauna. Roy's ability to launch a 14ft Ski-Vee through the Umgazana river mouth in large swell during outgoing tides was also appreciated. He also aided in the collection of samples from the 2011 sardine run.

I would like to thank members of the Department of Agriculture, Forestry & Fisheries (DAFF), namely Dr Carl Van Der Lingen and Marc Russell Hendricks who aided in the procurement of samples from the Western and Eastern Cape, and to Dr Van Der Lingen who supplied me with his personal insights into Sardines in South Africa. I would like to thank Dr Allan Connell, for his insights into the KZN sardine run and their larvae, and for his supply of larvae that were used in this study. Thanks must also go to Greg Thomson at the KZN Sharks Board for regular sardine updates from the seat of the Sharks Board spotter plane. Thanks must go to the two anonymous reviewers of this dissertation. Their extensive criticisms and comments were appreciated and integral to the overall quality of this study. I would like to thank The Rufford Small Grant Foundation and the National research foundation (NRF) field work and personal funding, for which this work would have not been possible.

Lastly, to my family; Mike, Carol, and Dean Chiazzari, and Jessica Wiggill; for without their unconditional support, my pursuits would not be as successful. I would also like to thank them for putting up with the smell of over a 1000 defrosted sardines for weeks on end, at home, while processing was done.

Chapter one:

General introduction

1.1. The sardine run: an overview

The ecological phenomenon known as the 'Sardine run' occurs annually when large schools of sardines (*Sardinops sagax)* move from the Agulhas Bank towards KwaZulu-Natal (KZN) (van der Lingen, Coetzee*, et al.* 2010). This is a well-known event along the KZN and Eastern Cape coast, primarily because of its value as an ecologically important phenomenon, natural resource and more recently, in the ecotourism industry.

Freon*, et al.* (2010) defines the sardine run as: "the visible effects of the coastal, alongshore movement during early austral winter of a small and variable fraction of the South African population of sardines (*Sardinops sagax*) from the eastern Agulhas bank to the KwaZulu-Natal (KZN) coast, as far as Durban and the north coast of KZN". The sardine run typically includes the movement of sardines and their predators into the neritic waters along the South African east coast from the Agulhas bank, and congregates further to the east near Waterfall Bluff in late May (Natoli*, et al.* 2008, O'Donoghue, Drapeau, Dudley*, et al.* 2010, O'Donoghue, Whittington*, et al.* 2010). The large shoals of sardines continue up the coast, and if suitable conditions prevail, they will move into KZN, and have been known to move as far up the coast as the Tugela Bank (Natoli*, et al.* 2008, Freon*, et al.* 2010, O'Donoghue, Drapeau, Dudley*, et al.* 2010, O'Donoghue, Drapeau and Peddemors 2010, O'Donoghue, Whittington*, et al.* 2010).

The South African sardine, *Sardinops sagax* (Jenyns, 1842), is a member of the family Clupeidae. Females are oviparous and spawn pelagically on the continental shelf and in large open mouth bays; their larvae are pelagic. Individuals have a mean life span of about two to three years but can attain nine years and can attain a length of about 20 centimetres (van der Lingen*, et al.* 2000). Although this epipelagic shoaling species is considered a non-selective filter feeder, juveniles and adults mostly feed on zooplankton (Whitehead*, et al.* 1988, van der Lingen 2002). Worldwide distribution of *Sardinops sagax* is restricted to the Indo-Pacific and

Atlantic oceans and concentrated populations are found in areas of ocean with strong upwelling present, favouring cool water in the range of 10°C to 20°C (Whitehead*, et al.* 1988).

Congregation of predatory species and their winter migrations correspond with the sardine's migration up the east coast. Shoals of sardines usually overlap with the presence of other small shoaling clupeids (Coetzee 1996, Coetzee*, et al.* 2010). The influx of the sardine biomass causes the run to be an important perennial source of energy for the oligotrophic waters off the east coast of South Africa (Coetzee*, et al.* 2010, Hutchings*, et al.* 2010). The sardine run also supports a seasonal beach-seine net fishery in KZN, and is also a focus of international ecotourism along the Eastern Cape and KZN coasts. Myeza*, et al.* (2010) estimated the socioeconomic value of the sardine run and its related activity to R 34 - 54 million (US$ 3.23 – 5.14 million, R10 to US$ 1 accessed on 07/04/2014) annually, which benefits local coastal communities (Myeza*, et al.* 2010).

The poorly understood features of the sardine run prompted a collaborative effort by researchers which recently shed light on the ecology of this unpredictable event, and progress has been made into the mechanisms that trigger the sardine run. Investigations include the biology and ecology of sardines and their predators during the sardine run, physio-chemical aspects of the sardine run, and the socio-economic impacts of the sardine run. Nevertheless, the question of why the sardine run occurs remains unanswered. A recent review of sardine run related literature, conducted by Freon*, et al.* (2010), provided a number of possible hypotheses for further investigation into the reasons for and mechanisms that drive the sardine run.

The aim of the introductory chapter of this dissertation is to provide an overview of the ecology of sardines and a summary of the current understanding of the sardine run up the east coast of southern Africa and report on hypotheses concerning the nature of the KZN sardine run Finally, a genetic and morphological study on the sardine run is presented, and it will be argued that such research can elucidate the reasons for the sardine run, and will aid in the planning and management of the commercial fishing and ecotourism industry and most importantly the ecology of *Sardinops sagax* in South Africa.

1.2. Ecological phenomenon: The sardine run and it's fishery

The sardine run corresponds with the onset of austral winter, usually beginning in late May, when large shoals of sardines move up the east coast of South Africa from the Agulhas Bank, usually in association with water between 15 and 18°C (van der Lingen, Hendricks*, et al.* 2010). Sardines on the Agulhas Bank extend their range further east and up the eastern coast of South Africa (Figure 1). Sardines are most frequently observed between Port St Johns and Waterfall Bluff area, in the Eastern Cape, and are forced close to the coastline by the narrowing continental shelf, and bounded by the unfavourable warmer Agulhas current offshore. Inshore counter currents (Roberts*, et al.* 2010) and the cooling of inshore waters along the eastern coast (Heydorn*, et al.* 1978) have some positive effect on the magnitude of the event, but are not the key cause for it (Freon*, et al.* 2010). Thus the effective winter range of southern African sardines can extend from the Western Cape to KZN (Figure 1).

Connell (2010) interpreted the annual bi-modal peak in sardine presence along the KZN South Coast as two separate movements (north and then south) and that the southward second movement of fish is likely to occur at depth due to the warming of surface layers of the ocean during spring. Thus, in late winter sardines are thought to move offshore at depth and most are thought to return to the Agulhas bank by early summer (Connell 1996; 2010). Sardine eggs are observed in surface plankton trawls in southern KZN from the onset of the sardine run in late May until December, indicating the presence of some sardines into early summer (Connell 1996; 2010). Sardines have also occasionally been caught by KZN ski-boat fishermen on Yusuri Jig and in the gut contents of predatory teleosts into late spring. Nevertheless, the exact numbers of sardines that return to the Agulhas Bank are not known.

The migration of sardines encourages a host of ecological interactions between species. Predation on sardines is the most evident interaction. Cetacean species such as the long-beaked common dolphin (*Delphinus capensis*) and seasonally migrating Indo-Pacific bottlenose dolphin (*Tursiops aduncus*) congregate along the Eastern coast and target the shoaling sardines as they move into KZN (Peddemors 1993; 1999, O'Donoghue, Whittington*, et al.* 2010). The abundance of elasmobranch species, like the copper shark (*Carcharhinus brachyurus*), spinner shark (*C. brevipinna*), and dusky shark (*C. obscurus*) increases significantly in KZN between June and July (Dudley and Cliff 2010). Teleosts, like the king mackerel

(*Scomberomorus commerson*), garrick (*Lichia amia*), giant kob (*Argyrosomus japonicas*), geelbek (*Atractoscion aequidens*), kingfish (*Caranx* sp.), Elf (*Pomatomus saltatrix*) and yellowtail (*Seriola lalandi*) also prey upon the small fish (Fennessy*, et al.* 2010). Many of these species' spawning migrations also coincide with the sardine run, possibly to increase the amount of nutrition along their migration route along the KZN and Eastern Cape coasts. Broekhuysen*, et al.* (1961) and Crawford et al. (1983) described the Cape gannet (*Morus capensis*) a non-breeding winter visitor to KZN which exhibit a close association to the sardine run. More recently, spatio-temporal associations between these avian predators and the sardine run have been investigated (O'Donoghue 2009, O'Donoghue, Drapeau, Dudley*, et al.* 2010, O'Donoghue, Drapeau and Peddemors 2010, O'Donoghue, Whittington*, et al.* 2010). These examples of the relationship of different taxa to the sardine run highlight its ecological importance.

There are many ecologically important interactions during the sardine run. Nevertheless, the most important aspect of the sardine run, ecologically, is the delivery of atypical amounts of nutrients, in the form of lipid and protein rich sardine biomass, to the oligotrophic East Coast waters (Hutchings*, et al.* 2010). Whilst acknowledging the high level of inter-annual variability of sardine abundance during the sardine run, a mean estimation of 960 tons of organic nitrogen is supplied from the Agulhas bank to KZN waters by migrating sardines. The sardine run can be equivalent to between 24 % and 76 % of the total annual nitrogen input; depending on the strength of the run, and the total nitrogen input from other sources (Hutchings*, et al.* 2010).

The Eastern Cape and KZN sardine run beach-seine fishery is sporadic, seasonal and substantially smaller than the South African sardine fishery. In 2010, the catch of sardines by purse seine netters in South Africa was 126 386 tons. Conversely, the total catch by beach-seine netters in KZN and the Eastern Cape in 2010 was less than 50 tons. The largest ever annual catch by beach-seine netters was just less than 700 tons, in 1993 (van der Lingen, Coetzee*, et al.* 2010). Indeed, from a fishery perspective, the sardine run has considerably less economic value than the South African purse seine fishery. However, the benefit to local beach-seine fishermen and their communities, along with the ecotourism generated, is locally significant (Myeza*, et al.* 2010).

Most recently, a number of papers published in the *African Journal of Marine Science* have dealt with patterns of distribution of sardines and their predators in relation to environmental conditions (Dudley and Cliff 2010, Fennessy*, et al.* 2010, O'Donoghue, Drapeau, Dudley*, et al.* 2010, O'Donoghue, Drapeau and Peddemors 2010), biological characteristics of the KZN sardine (van der Lingen, Hendricks*, et al.* 2010), long term evidence of spawning of sardines in KZN (Connell 1996; 2010), biomass of sardine and it's ecological impact during the sardine run (Coetzee*, et al.* 2010, Hutchings*, et al.* 2010), socioeconomic impacts to local KZN and Eastern Cape communities (Dicken 2010, Myeza*, et al.* 2010), and environmental factors such as ocean currents and sea surface temperature (SST) (Roberts*, et al.* 2010). This and previous research has culminated in the development of a paper by Freon*, et al.* (2010) citing hypotheses about the proximate and ultimate factors relating to the cause of the sardine run.

In recent years, important insights about the nature and ecology of the sardine run have been revealed. The final publication in a special series of papers by Freon*, et al.* (2010) describes current hypotheses about the cause of the sardine run. Where sufficient data exists, some of these hypotheses were tested. However, for the sake of brevity, only those that were supported by current literature will be discussed.

1.3. Proximate and ultimate factors: current hypotheses about the sardine run

1.3.1. Ultimate factors

Freon*, et al.* (2010) listed a number of hypotheses that may contribute to the ultimate cause of the sardine run. In total, seven ultimate hypotheses were presented and tested, but only two were supported. The first of the supported hypotheses was proposed by Baird (1971), who proposed that the sardine run formed part of a subpopulation spawning migration. The second ultimate hypothesis supported was that sardines undertook their migration due to relic behavioural responses and possibly a bet-hedging strategy (Olofsson*, et al.* 2009, Rees*, et al.* 2010, Ripa*, et al.* 2010) during the last glacial maximum (Freon*, et al.* 2010). Both of these hypotheses were tested by Freon*, et al.* (2010) via the same tests, except one which was used to test the relic hypothesis. The tests used to support the two hypotheses are listed and discussed below.

Baird's (1971) and Coetzee, *et al. s '* (2010) hypotheses state that sardines of the KZN sardine run are in fact undergoing a migration. Freon*, et al.* (2010) list Dingle's (1996) conditions for a movement to be deemed a migration. Dingle's (1996) migration definition requires that sardines would need to display direct, predictable and continuous motion, swimming outside of their usual home range. Indeed, these fish undergo a journey of hundreds of kilometres clear of their home range (see Figure 1), in a direct manner and towards a predictable area i.e. KZN. Another requirement for the sardine run movement to be deemed migration is that the fish should ignore other resources along the migration route that would usually not be ignored (Dingle 1996). Sardines leave the plankton rich waters of the Agulhas Bank to swim into the nutrient poor, warm oligotrophic waters of KZN (Hutchings*, et al.* 2010). The poor condition of sardines partaking in the sardine run exemplifies the fact that these individuals do not move into KZN to feed (van der Lingen, Hendricks*, et al.* 2010).

Dingle (1996) stated that a migratory movement requires specific departure and arrival behaviour. Although these aspects of the sardine run remain unclear, Freon*, et al.* (2010) report that shoals of sardines on the eastern Agulhas bank leave for KZN in May. Sardine shoals usually arrive in the northern Eastern Cape near Port St. Johns by June, and if conditions are suitable, they may move into KZN periodically, in accordance with the Waterfall Bluff gateway hypothesis described by Roberts*, et al.* (2010), from June to September (Freon et al. 2010). The use of proxies such as occurrence of predators (such as Cape gannets and Common dolphin) has been used to determine the presence of sardine shoals, and thus the beginning of the sardine run (O'Donoghue, Drapeau and Peddemors 2010). In KZN, plankton trawls for sardine eggs, are used to detect spawning sardines. Indeed, sardine eggs are commonly found in plankton trawls until December (Connell 2010).

The last condition, stipulated by Dingle (1996), is that the migrant sardines should show physiological storage of energy in the form of fats and lipids. Long term research between 1953 to 2004 (van der Lingen and Durholtz 2004, van der Lingen*, et al.* 2006, van der Lingen, Hendricks*, et al.* 2010) showed that sardines from the southern cape and Agulhas bank (Figure 1) illustrate a higher mean year round conditioning factor than those found on the western

coast (Table 1). Individuals from the sardine run displayed the lowest conditioning factor (van der Lingen*, et al.* 2006) of all sampled areas.

Recent consideration has been given to the development of a multistock hypothesis of *Sardinops sagax* in South Africa, first conceived by van der Lingen*, et al.* (2005). Miller*, et al.* (2006) corroborated the multistock hypothesis of van der Lingen*, et al.* (2005) via modelling the transport and retention of sardine eggs and larvae in the southern Benguela. Based on these results, Miller*, et al.* (2006) confirmed two distinct spawning and nursery grounds of sardine between the Western Cape and on the Agulhas Bank (see Fig. 1). Further research by van der Lingen*, et al.* (2009) provided a morphological basis for the delineation of multiple sardine stocks, namely the western (west of Cape Agulhas) and southern (east of Cape Agulhas) sardine stocks. Further to this Freon*, et al.* (2010) deduced that sardines partaking in the sardine run likely come from sardines on the Agulhas Bank, based on previous morphological work completed by van der Lingen, Hendricks*, et al.* (2010) and Coetzee*, et al.* (2010). Freon*, et al.* (2010) concluded that KZN sardines lost their lipid reserves as they travelled northwards into KZN, and according to van der Lingen, Hendricks*, et al.* (2010) the sardine run sardines constituted a third "eastern" winter season stock. Thus, from the above information, the sardine run generally conformed to Dingle's (1996) definition of a migration.

Baird (1971) suggested that the sardine run formed part of a subpopulation spawning migration; Freon*, et al.* (2010) questioned whether sardines from the sardine run were phenotypically and genetically different (see Freon' test 3 and 4) from the rest of the SA population. Phenotypic data can be useful for indirect subpopulation separation and show post larval variation between varied environments (Freon*, et al.* 2010), as illustrated in Table 1 (van der Lingen, Coetzee*, et al.* 2010). Most recently, van der Lingen, Hendricks*, et al.* (2010) described the biological characteristics of sardines caught off KZN and compared them with the eastern and western stock, and concluded that the KZN stock represented a distinct stock or possibly a functionally discrete adult assemblage (FDAA). A FDAA is defined as a population of adults that are physically or reproductively isolated from another such group. These conclusions were drawn from significant differences in conditioning factor, vertebral count and

body shape of fish from the Agulhas Bank (eastern stock, Figure 1) and KwaZulu-Natal (van der Lingen, Coetzee*, et al.* 2010, van der Lingen, Hendricks*, et al.* 2010).

Definitive evidence of stock structure can be determined by the use of various modern genetic techniques. Nevertheless, this powerful tool has yet to be applied specifically to the sardine run. Hampton (pers. comm.) has undertaken a study to test the two stock hypothesis (Figure 1) by van van der Lingen*, et al.* (2009), de Moor*, et al.* (2008) and de Moor and Butterworth (2009) of the southern Benguela sardines. The lack of molecular data relating to the sardine run has left a key shortage in the collective knowledge of the sardine run. The use of these techniques is discussed further on.

Another method of identifying stock structure, using parasites as biological tags, has been proposed by Reed*, et al.* (2012). In their assessment of possible parasitic biological tags, it was illustrated that some members of the digenean tetracotyle metacercariae (a subclass of parasitic flatworms) have potential as biological tags. Studies employing biological tags (Reed, UCT; van der Lingen, DAFF; Froeschke, Stellenbosch University) and modern molecular techniques (within this study) are underway, and should produce a description of the stock structure of the sardine run complementary to this study.

Freon*, et al.* (2010) tested another requirement for the spawning migration hypothesis to be fully supported, and this was that migrating sardines would have to spawn in KZN. Connell's (2010, 2012) twenty-one year time-series of ichthyoplankton data confirms that sardines do spawn during the sardine run. Occurrences of sardine eggs in plankton trawls begin with the start of the run in June, and continue before slowly declining to almost no observations in December (although sardine eggs have been collected as early as March and late as January), when most sardines are thought to have returned to the Agulhas bank (Connell 1996; 2010; 2012). The sixth requirement highlighted by Freon*, et al.* (2010), is that a nursery area should exist in KZN waters. Indeed, larvae (Beckley and Hewitson 1994, Beckley and Naidoo 2003) and juveniles (PCN. 1978) have been documented in KZN. More recently, Freon*, et al.* (2010) described how Miller*, et al.* (2006) observed juvenile sardines during a survey of the waters between Port Elizabeth and Port St. Johns. Miller*, et al.* (2006) thus reasoned that these

sardines were possibly spawned in KZN and travelled south with the Agulhas current. Miller*, et al.* (2006) added that long term data by Connell (2010) suggested a causative link between the winter presence of sardines in KZN and the presence of eggs in the water column off KZN. Thus KZN is likely a nursery ground for the sardines spawned during the sardine run.

It is not yet known whether KZN spawned juveniles are successfully recruited to the adult stock (Freon*, et al.* 2010). Nevertheless, the number of juveniles recruited to the returning stock of sardines must be significant, especially if the KZN sardine run represents a genetically distinct subpopulation spawning migration or relic behaviour (Freon*, et al.* 2010). Modern molecular techniques can be used to compare relatedness of sardine larvae and juveniles sampled in KZN and along the Agulhas bank (Freon*, et al.* 2010). Other techniques such as otolith biochemistry can also aid in determining the origin of larvae and juveniles and biannual sampling of otoliths can allow one to determine the regions that sardines occupy during the year.

Freon, *et al.'s* (2010) final test of the subpopulation spawning migration hypothesis was to test whether all KZN spawned sardines return at sexual maturity to spawn only in KZN. Allopatric spawning in *Sardinops sagax* is possible; where sardine run individuals only spawn in KZN, and those of the western and eastern stock off the Agulhas bank only spawn on the Agulhas bank (Freon*, et al.* 2010). However, Freon*, et al.* (2010) established that due to the lengthy and bimodal spawning characteristics of the sardines, it is likely that larger fecund females spawning in KZN in June and again later in the year upon returning to the Agulhas Bank. The use of fine scale population genetic techniques, investigating whether spawning populations exist, remains the most feasible method of shedding light on the subject (Freon*, et al.* 2010).

Freon *et al.*'s (2010) final test, which only applies to the relic behaviour hypothesis, relates to past environmental conditions that may have influenced distribution and reproductive strategies, including bet hedging (Coetzee*, et al.* 2010). Conditions during the last glacial maximum approximately 18 000 years ago, included lower SST and sea levels (Freon*, et al.* 2010). These past environments are thought to have encouraged a larger suitable range for *Sardinops sagax*. It is likely that the west and east coast sardines were separated by the limits

of suitable habitat (see Figure 1) and that two large areas of the continental shelf off Mozambique were suitable for sardines. This area had increased fluvial input and may have been favourable to sardines as a nursery ground and thus could be a relic of a spawning migration of sardines during the last glacial maximum (Freon*, et al.* 2010). Nevertheless, without evidence of the presence of sardines off Mozambique in the fossil record, any assumptions remain conjecture (Freon*, et al.* 2010).

1.3.2. Proximate factors

Proximate factors are those that enable or trigger the sardine run. Freon*, et al.* (2010) tested and ultimately supported two of the five proximate factors involved in the sardine run. The first hypothesis to explain the triggering of the sardine run is the natal homing and imprinting hypothesis, defined by Carr (1967), where adults will return to their place of spawning and Stabell (1984) who added that natal homing was brought on by external signals which are imprinted at the egg or larval stage (Freon*, et al.* 2010). Certainly, this hypothesis could explain why such a small percentage of sardines migrate annually during the sardine run and that the amount of sardines partaking in the run shows no relationship to population size off the Agulhas bank and the Western Cape coast (Freon*, et al.* 2010). Vertebral counts may also yield important information about the above hypothesis. Sardines spawned in warmer waters, such as off the KZN coast during the sardine run have less vertebrae than those spawned off the Western Cape (Freon*, et al.* 2010). This may help to identify where fish may have been spawned. However, the use of vertebral counts as a proxy for spawning locality has a significant caveat in that sea surface temperature (SST) is variable along the southern African coastline throughout the year (Lutjeharms*, et al.* 2001).

The second proximate hypothesis is environmental forcing of sardines up the east coast via the Waterfall Bluff gateway, first observed by Armstrong*, et al.* (1991) and then confirmed by Roberts*, et al.* (2010). Here, sardines are hypothesised to be forced by currents of cooler water, bounded by unfavourable warm Agulhas current water, and the Natal pulse. Nevertheless, this hypothesis cannot be tested by means of the objectives set out below, thus for the sake of brevity the above mentioned and explanation will suffice.

The findings described in the sardine run special edition were constrained by available ocean current data. However, the need to establish objectives to fully test the hypotheses listed

above, is critical. Indeed, the need for targeted research has been reiterated by Freon*, et al.* (2010). One of the tools described as suitable to test all but the Waterfall Bluff gateway hypothesis, is the use of modern molecular genetics (Freon*, et al.* 2010). The work proposed, and ultimately conducted in this MSc dissertation is based primarily on modern molecular techniques briefly outlined in the following paragraphs, with a more in depth description given in chapter 2 (Methodology).

The need to establish population ranges and parameters such as identity, variance, relatedness, demographics, diversity, degrees of inbreeding, and migration is important to understanding the sardine run. Practically, this information could allow fisheries and conservation managers to determine populations under potential threat, and those that may be exploited to an optimum. Multi population fisheries may be implemented if necessary, thus generating a more realistic fisheries model. This minimises the effect of selection differentials on specific populations (Law 2007).

1.4. Motivation and rationale for this study

1.4.1. Population genetic and morphological techniques in population delineation

Taxonomy is important to fishery scientists for the delineation of fished resources, and aids in developing rational conservation strategies (Agüero and Rodríguez 2004). The taxonomy of the sardine globally, including other clupeid species has been contentious (Agüero and Rodríguez 2004), and this has been the case for *Sardinops sagax* in Southern Africa. Five global sub-species of *Sardinops sagax* were identified worldwide by Grant and Bowen (1998); including the sub species of *S. sagax* from southern Africa, designated the sub species name: *S. ocellatus*. Nevertheless, studies focused on stock structure and systematics have subsequently shown the genus to be monotypic (Beckley and van der Lingen 1999).

The sardine is referred to as *Sardinops sagax*, although it is considered a different sub-species to the South American *S. sagax sagax*. In southern Africa, *S. sagax* occurs year round on the Agulhas bank between Port Elizabeth in the east and into Namibia (sometimes as far as Angola) to the north west. Genetic and morphological analysis can be especially valuable in testing the above hypotheses, where broad scale structuring among sardines found on the

east and south or west coast during the sardine run may be determined. Finer scale connectivity and population structure amongst sardines within the sardine run may also be determined via genetic analysis. Genetic structuring of the sardines will allow one to test the current hypotheses outlined above, thought to be the cause of the sardine run (Freon, *et al.* 2010). Genetic techniques using microsatellites, exon-primed intron-crossed primers, single nucleotide polymorphisms and many other methods are common now to many fished stocks such as salmon (Ryynanen and Primmer 2006), Big eye tuna (Martínez, *et al.* 2006) and a host of other species, including Clupeidae (Touriya, *et al.* 2003, Pereyra, *et al.* 2004, Atarhouch, *et al.* 2006, Gonzalez and Zardoya 2007, Keski̇n and Atar 2012).

Genetic analyses can be used for testing the hypotheses proposed by contemporary literature, but are also crucial to determining the demography of the sardine stocks which are fished commercially on both the west and east coasts of Southern Africa. This may ultimately lead to the conservation of genetically delineated stocks. This is especially important in ensuring species and stock resilience to anthropogenic pressure and thus the sustainability of the sardine run as a fished stock and an economically important ecological phenomenon primarily through the preservation of genetic diversity.

Based on recent data (van der Lingen, *et al.* 2005, Miller, *et al.* 2006, Coetzee, Van der Lingen, *et al.* 2008), (van der Lingen, *et al.* (2009) the southern Benguela stock appears to be composed of two separate stocks, with a transition zone separating the stock west of Cape Agulhas and east of Mossel Bay (Figure 1). Miller, *et al.* (2006), de Moor and Butterworth (2009) and van der Lingen, *et al.* (2009) concluded that apart from some larval exchange within the transition zone, effectively two distinct stocks exist with limited larval and egg exchange. Hampton (*pers. comm.)* found some shallow genetic structuring, using various microsatellite markers.

Sardines found west of Cape Agulhas exhibited different phenotypic characteristics to those found east of the transition zone (see Figure 1). Generally, sardines found west of the transition zone had poorer conditioning, lower fat content, larger relative gonad size, smaller

relative body size, different otolith shape, body shape and larger head size (see table 4; van der Lingen*, et al.* (2009)). West coast sardines also fed on smaller copepods and were found in cooler waters, based on data from multiple years (see Table 4; Van der Lingen *et al.* (2009)). Most notable however, is the separation of the breeding stock and nursery grounds of the western and southern stocks (Miller*, et al.* 2006). Although Miller*, et al.* (2006) found some degree of egg and larval transport, mostly from the western stock to the east; the two stocks exhibit a measure of reproductive isolation.

Although all sardines along the southern African coast have overlapping ranges, new data describing shallow population structuring between the western and eastern Benguela sardine stock have been identified (Shannon and van der Lingen, Pers. Comm.). This may be explained by the fact that it is possible that southern Benguela sardines may form functionally discrete adult assemblages (FDAA's) (van der Lingen, Hendricks*, et al.* 2010). The delineation of these stocks may be caused by factors such as diet, conditioning (van der Lingen*, et al.* 2009), habit, breeding and dispersal ranges (Coetzee, Van der Lingen*, et al.* 2008, de Moor*, et al.* 2008, de Moor and Butterworth 2009). Limited larval and juvenile exchange in the area between Cape Agulhas and Mossel Bay may be a manifestation of two FDAA (Figure1; Miller*, et al.* (2006)).Thus, molecular techniques employed during this study must be able to identify fine scale genetic structuring as well as broader scale sub-species and species level phylogeographic relationships that may exist. This requires the use of molecular markers that evolve at different rates, to discern evolutionary relationships at different taxonomic resolutions.

Broad-scale markers, such as Cytochrome-b, D-loop and COI, have been used successfully on clupeids to elucidate phylogenetic relationships within and between species (Ivanova*, et al.* 2007). Recently, universal primer cocktails designed to amplify a 650 base pair region of the cytochrome oxidase I (COI) region have been developed at the Canadian centre for DNA barcoding and have been successfully used to barcode marine fishes. The barcoding of sardines has already been conducted on sardine larvae and adults in KZN and abroad (Ward*, et al.* 2005, Ward*, et al.* 2009, Zemlak*, et al.* 2009, Steinke*, et al.* 2011, Cawthorn*, et al.* 2012, Keskİn and Atar 2013).

The sardine run, as discussed above, may show some amount of genetic structuring. This necessitates the use of a spectrum of very fine, to moderate scale markers such as the microsatellite markers described by Pereyra*, et al.* (2004) and others such as the exon-primed, intron-crossing (EPIC) primers described by Touriya*, et al.* (2003) used on the clupeid *Sardina pilchardus*, and other non-model teleost fish (Hassan*, et al.* 2002, Li*, et al.* 2010). These primers can be utilised in many different ways to reveal genetic structure.

The mode of analysis depends on many factors, including the type of data required to meet the aims of the study, available laboratory equipment and the allocated budget required to process representative numbers of individuals per population. Recently, modern molecular biology has shifted towards the use of DNA sequencing as the means through which molecular analyses are carried out (Sunnucks*, et al.* 2000). Nevertheless, this method can be expensive and time consuming (Sunnucks*, et al.* 2000).The development of fast and relatively cheap yet effective protocols, capable of detecting accurate levels of sequence polymorphism are available and have been used successfully in population genetic studies (Sunnucks*, et al.* 2000). This includes length polymorphism gel electrophoresis, employed in this study, where many individuals may be analysed over multiple marker sets in a short period of time (details illustrated in chapter 2; methodology).

1.4.2. Conclusions

Considering the above, the aim of this project is to investigate the heterogeneity of the South African sardine stock (*Sardinops sagax*) using length polymorphism gel electrophoresis and sequencing of EPIC PCR amplicons. Population structure will also be investigated using meristic and morphological characters, using multivariate and geometric techniques. Samples collected off KZN sites, Port St. Johns, Port Elizabeth and sites in the Western Cape will be used (see methodology). Findings from the genetic analyses will be used to determine whether sardines caught in KZN return to their (possible) natal spawning grounds, and whether sardines

partaking in the KZN sardine run form part of a sub-stock of sardine in South Africa, thus testing Freon*, et al.* (2010) hypotheses.

If it is found that KZN sardines are a sub-stock, this would have important consequences for the management of this event, particularly from an ecosystem approach to managing the beach-seine and the purse-seine fishery in South Africa. The threat of potential near shore ocean warming, and effects on local upwelling cells, due to the strengthening of the Agulhas Current as a consequence of climate change, may also have negative effects on the sardine run (Rouault*, et al.* 2010). A number of potential causes including climate mediated changes in physical and ecological parameters have been observed in changes in the distribution and behaviour of a large number of species globally (see Brander (2010) for examples) These causes may also play a role in the south and eastward shift in sardine distribution in South Africa (Coetzee, Van der Lingen*, et al.* 2008). Given the importance of the annual sardine run movement, as outlined above, it is critical that further research is undertaken to provide both continuity in research effort and to create the knowledge necessary for the successful marine management of this event.

Information collected in this study will further elucidate the nature of this event and the ecology of the South African sardine, and practically; allow stock managers to plan conservation strategies based on morphological and genetic diversity, connectivity, and breeding patterns within the KZN sardine. Ultimately, understanding genetic relationships and breeding patterns of the migrating fish stocks will allow for the conservation of genetic diversity of the fish, and the preservation of the KZN sardine run. This will not only benefit the sardine, but also the multitude of taxa such as marine birds, teleosts, elasmobranchs, and cetacea that take advantage of the sardine run.

1.4.3. Arrangement of dissertation

This dissertation consists of five chapters; an introductory chapter (chapter 1), Materials and Methods (chapter 2), two chapters describing the morphometric (using meristic, multivariate and geometric morphology) and genetic (using mitochondrial and exon-primed, intron-crossing markers) population structure of *Sardinops sagax* from KZN and the Eastern Agulhas Bank and Western Cape stock (chapter 3 and chapter 4), and a concluding chapter (chapter 5). The materials and methods in chapter 3 and chapter 4 are presented together in chapter 2 to avoid repetition.

Chapter two:

Materials and methods

2.1 Study sites and sample collection

To assess the population connectivity of sardines along the South African coast and thus allow for comparison of the hypothesised KZN sardine stock, sampling was conducted west of Cape Agulhas, on the Agulhas bank, and in KZN between June 2011 and September 2013. The sites mentioned represented sampling sites within hypothesized populations (Figure 1). A total of 1039 sardines were collected (including 34 larvae samples collected in KZN) from five sites along the South African coast; namely Cape Town, Mossel Bay, Port Elizabeth, Port St. Johns, and the KZN South coast from Port Edward to Park Rynie (see Table 1).

Table 1: *Sardinops sagax* collected in South Africa from July 2011 to 2013, including sample numbers for morphometric, meristic and genetic analyses.

Sample	Sample type	Date collected	Location and Coordinates	Capture method	Morphometric analysis	Meristic analysis	mtDNA	EPIC
A	Adult	01/07/2011	KZN, Ifafa (30°27'S 30°42'E)	BSN	None	204	11	48
B	Adult	01/07/2012	KZN, Southbroom (30°56'S 30°18'E)	BSN	None	210	10	48
C	Adult	01/11/2012	KZN, Southbroom (30°18'S 30°44'E)	RR	None	7	None	None
D	Larvae	30/09/2012	KZN, Port Edward (30°18'S 30°44'E)	PN	N/A	N/A	5	22
E	Larvae	13/07/2012	KZN, Hibberdene (30°18'S 30°44'E)	PN	N/A	N/A	None	10
F	Larvae	22/11/2012	KZN, Hibberdene (30°18'S 30°44'E)	PN	N/A	N/A	8	4
G	Larvae	09/12/2012	KZN, Park rynie (30°18'S 30°44'E)	PN	N/A	N/A	None	None
H	Adult	04/02/2013	WP, Cape Town (33°04'S 17°50'E)	PSN	122	200	13	48
I	Adult	09/02/2013	EC, Port Elizabeth (33°44'S 25°58'E)	PSN	200	201	6	None
J	Adult	04/04/2013	WP, Mossel Bay (34°35'S 22°27'E)	PSN	200	200	None	None
K	Adult	08/07/2013	KZN, Port St. Johns (31°02'S 29°34'E)	RR	17	17	None	None
				Total:	522	1039	40	378

WP=Western Province, EC= Eastern Cape, KZN=KwaZulu-Natal, BSN=Beach-seine net, RR=Rod and reel, PN=Plankton net, PSN=Purse seine net. N/A=Not applicable.

Sardines were caught using purse seine nets with a 28mm mesh size aboard commercial purse seining vessels in Cape Town, Algoa Bay (Port Elizabeth) and Mossel Bay (Table 1) for research by the Department of Agriculture, Forestry and Fisheries (DAFF). Sardines were collected from beach-seine netters operating along the Eastern Cape's wild coast and KZN coast in winter. Sardines were netted in favourable surf conditions along sandy beaches where shoals of sardines came close enough to the shoreline (generally less than 300m from the shore) to be caught in beach-seine nets with a wing mesh size of at least 14mm and at least 13mm for the cod end, with a total net length of less than 100m. Sardines were also caught via Yusuri jig (Kingfisher Sabiki luminous live bait jig, No. 4 hook) on rod and reel in less than 50m depth of water, from a ski-boat in KZN and the wild coast (Table 1). Sardines were frozen at -20°C after capture for approximately two months before morphometric analysis.

Sardine eggs were collected via surface plankton trawls off Park Rynie (KZN) in the months after the sardine run, as part of an on-going ichthyoplankton survey started in 1987 and aimed at understanding the spawning patterns of pelagic fish off the KZN coastline (Connell 1996; 2010; 2012). A two-meter long cone shaped plankton net with an aperture of 1360cm^2 and a mesh size of 300µm was trawled along the water surface at approximately 2km/h for 10 minutes. The plankton trawls were conducted in 40 to 50 meters of water depth. Sardine eggs were sorted from the trawl samples based on external characteristics, hatched overnight, and incubated for five days, so as to aid identification, by Dr Allan Connell. Larvae were stored in 95 % ethanol at room temperature.

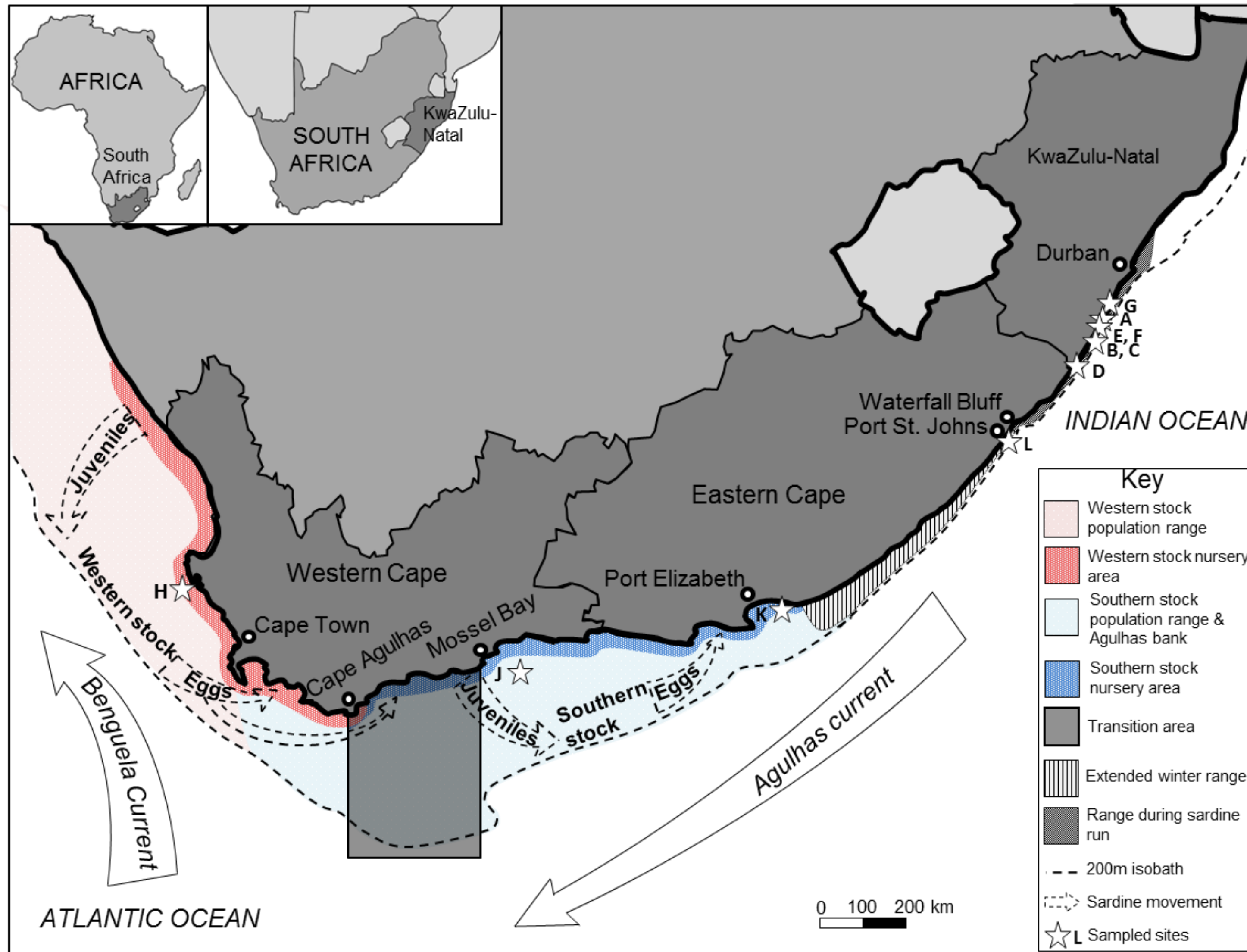


Figure 1: Map of *Sardinops sagax* occurrence and putative stock delineation in Southern Africa, including the winter and summer ranges and the spatial extent of the sardine run. Letters associated with sites (represented by stars) correspond to sites sampled in table 1.

2.2 Meristic data collection and analysis

In ichthyology, meristic data is that which can be counted and can include measurements such as wet body mass, caudal length, sex ratio, and number of vertebrae, and can be useful indicators of stock identity. These meristic traits are often sensitive to environmental factors such as stress, food type and availability, and water temperature; and can cause differences in these measurements between populations. Hence, these parameters have successfully been used on sardine stocks in South Africa for some time (van der Lingen and Durholtz 2004, van der Lingen*, et al.* 2005, Miller*, et al.* 2006, Coetzee, Van der Lingen*, et al.* 2008, Wessels*, et al.* 2010). Wet body mass and caudal lengths were recorded for all individuals; which were subsequently frozen at -20°C for two months; this was to standardize the amount of time all sardines were exposed to freezing and its effect on fish morphology (Wessels*, et al.* 2010). Vertebrae in the Clupeidae vary in number due to environmental temperature during development. Thus, vertebrae number between individuals of different populations is an important proxy for the origin stock identity of sardine populations in southern African waters. The difference in population variation in vertebral count is due to the temperature gradients along the coast. Sardines spawned in colder water generally have more vertebrae, and less in warmer water. After photographs were taken for geometric analysis (see: 2.3.2. Multivariate and Geometric analysis), sardines were dissected dorso-ventrally to count vertebrae and determine sex.

A correlation analysis between caudal length and wet body mass of fish was conducted. A non-parametric Spearman's rho correlation was computed after the assumptions for parametric analysis, via a one sample Kolmogorov-Smirnov test, were not met. R^2 values were also calculated for each population.

To test the possibility of sexual dimorphism (and thus the necessity to test population differences according to sex), a non-parametric Kruskal-Wallis ANOVA was conducted for wet body mass, caudal length, and vertebrae count within each population according to sex. This, after the assumptions of homeoscedacity (Levene’s test) and normality (1 sample Kolmogorov-Smirnov test) were not met. The parameters were split within their sites so intrapopulation

sexual dimorphism could be measured. No sexual dimorphism was observed for the measured parameters. This meant both sexes could be analysed together, between populations.

A non-parametric Kruskal-Wallis ANOVA was conducted for wet body mass, caudal length, and vertebrae count within each year class of the KZN population so any differences in sardines sampled between years during the sardine run could be measured. There were no significant differences observed between years. This, after the assumptions of homeoscedacity (Levene's test) and normality (1 sample Kolmogorov-Smirnov test) was not met. This meant that all year classes could be analysed together. Vertebrae and sex ratio frequency distributions were also calculated for each population. All statistical tests were undertaken with a 95 % confidence interval using IBM SPSS version 21.0 (IBM Corp. 2012).

2.3. Morphological data collection and analysis

2.3.1. Material collection and preparation

Sardines were stored frozen at -20°C for two months, before being photographed. Sardines were carefully placed in polystyrene boxes, in single layers, laid straight and belly-up to ensure minimal deformation of body shape (resulting from vertebral columns of fish freezing at varied angles) or tearing of the soft abdomen flesh. Sardines were slowly air thawed to room temperature before photographs were taken. Sardines were placed individually on a white plastic surface alongside a set square with millimetre marking. Individuals were placed upon a line drawn across the white plastic, with a marking to place the tip of the nose of the fish. This ensured consistent placement of the fish within the camera's focal range. Individuals were manipulated into a natural position with their mouths closed, vertebral column aligned; and flared dorsal, caudal, anal, pelvic and pectoral fins. The camera was placed on a stand 35 cm from the fish and so that the focal point was the centre of the fish. Photographs were taken without flash, and in good lighting conditions. Careful attention was given to keeping all components of the photographic apparatus in the identical position between photograph sessions, thus keeping a standardised photographed area. Digital photographs of the left side of each individual were taken with a Sony DSC-WX7 camera (resolution: 28.3 pixels.cm^{-1}).

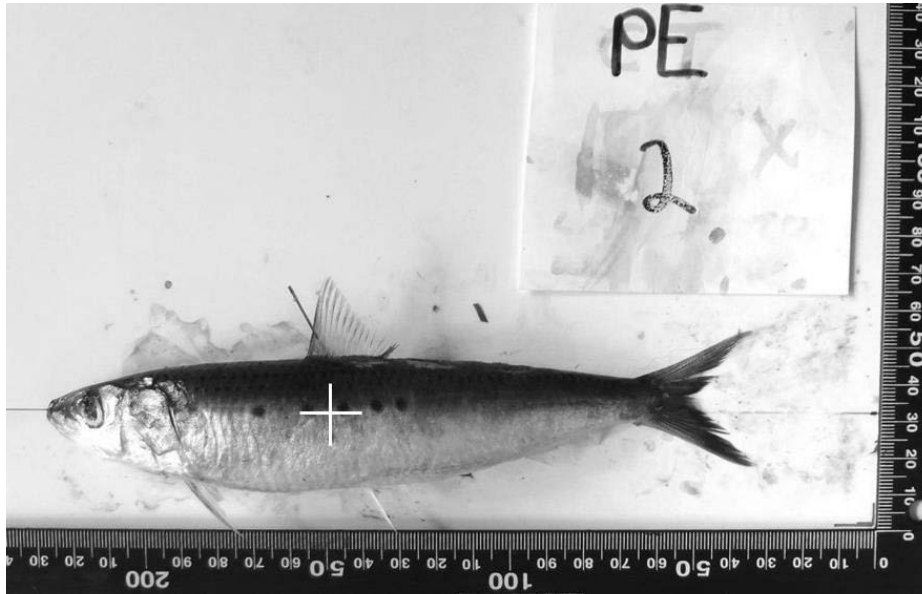


Plate 1: Typical photographed area (cropped to show only scale bar and sample) of *Sardinops sagax* for landmark placement before morphometric analysis. The white cross represents the consistent focal area.

2.3.2. ***Multivariate and Geometric analysis***

Morphometric analyses are a powerful set of tools in evaluating stock identity, discrimination, and delineation (Cadrin 2000). Although recently superseded by modern genetic techniques, this technique is important in delineating populations that can only be observed in an individual's phenotype. Thus, in combination with genetic techniques, geometric morphometrics offer a powerful tool to determine population boundaries according to phenotype and genotype (Cadrin 2000).

Eleven digitised landmarks were recorded upon the sardines photographed, with 21 truss measurements (Figure 2), using tpsDig2 software (Adams*, et al.* 2004). Easily recognisable and reproducible points on the sardine were chosen to minimise error associated with placing landmarks, and in accordance with previous studies for continuity and possible future

comparison (Silva 2003, Agüero and Rodríguez 2004, van der Lingen, Hendricks*, et al.* 2010, Wessels*, et al.* 2010).

Table 2: Description of landmark positions placed on *Sardinops sagax* samples for morphometric analyses.

Land mark coordinate	Landmark description
1	Tip of the snout
2	Top front corner of operculum, below occipital ridge
3	Front insertion point-dorsal fin
4	Back insertion point-dorsal fin
5	Upper insertion point-caudal fin
6	Lower insertion point-caudal fin
7	Front insertion point-anal fin
8	Front insertion point-pelvic fin
9	Front insertion point-pectoral fin
10	Jaw hinge, aligned with operculum
11	Point where last backbone meets caudal fin bones

Ten landmarks (1 - 10, see Figure 2) were used to produce 21 box trusses for subsequent multivariate analyses, by calculating relative distances between landmarks using PAST version 3.01 (Hammer*, et al.* 2012). A Principal Coordinates Analysis (PCA) was run and outlying samples from the 95 % confidence limits of the scatter plot were removed from further analysis. Seven females were removed due to abnormally distended stomachs. Truss variables were then corrected for by size, using the Burnaby method, after being log transformed (Burnaby 1966, Rohlf and Bookstein 1987). Truss variable size was corrected by projecting truss variables on the subspace orthogonal to the space covered by the size vector in PAST (Hammer*, et al.* 2012). Subsequently, confidence (95 % interval) limits indicated individual size vectors were not significantly different from those of other individuals, suggesting a common size transformation had standardised all samples.

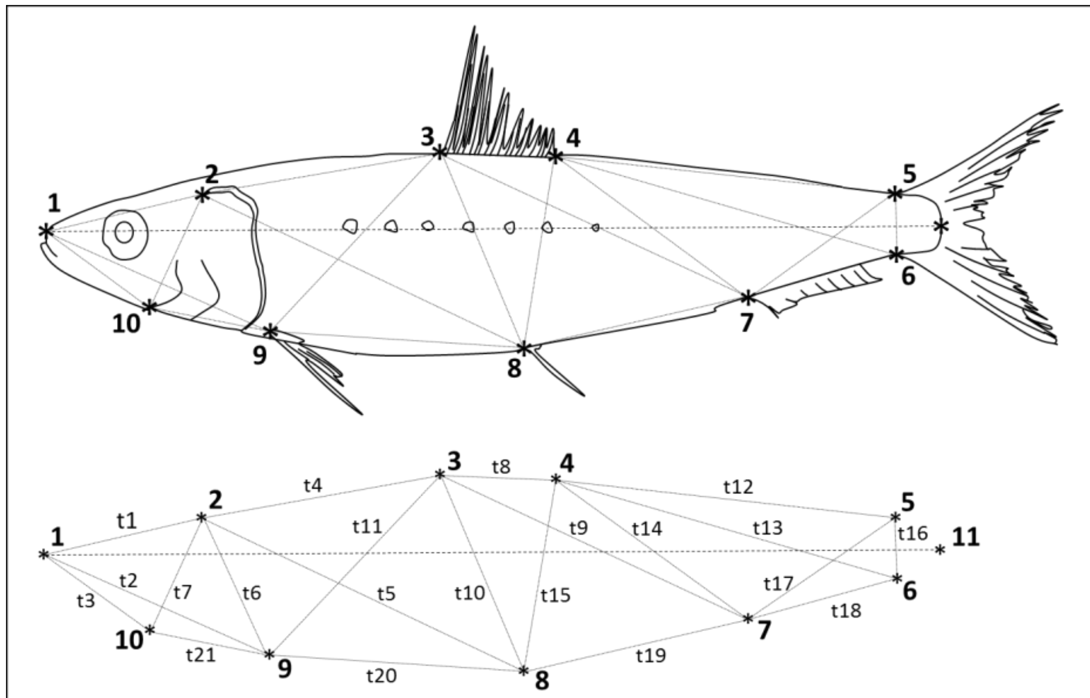


Figure 2: Placement of eleven morphometric landmarks and the truss network (t1 - t21) used in multivariate and geometric analysis, for *Sardinops sagax* caught at eight sites along the South African coastline. Dotted line (landmark 1 - 11) indicates the caudal length measurement.

A Principal Component Analysis (PCA) was conducted for the size-corrected truss variables, and principal coordinates calculated using SPSS (IBM Corp. 2012). Group centroids for individuals of each site, and 95 % asymptotic confidence limits of the scores on the first two principal components, were computed for each sample. Discrimination functions between truss measurements according to site were calculated using a Discriminant Function Analysis (DFA). DFA was used to calculate the contribution of each truss measurement to site separation and individual difference to hypothetical mean measurements (McGarigal 2000). Relative discriminatory power of each truss measurement was also calculated for each truss measurement.

Geometric analyses were conducted using ten landmarks (1 - 10), in PAST (Hammer*, et al.* 2012). A generalised orthogonal Procrustes 2D analysis was completed for each sardine, through superimposition, to determine size corrections for all individuals using PAST (Hammer*,*

et al. 2012). Sample configurations were centred, scaled, and rotated to minimise the sum of squares distance between homologous landmarks for all individuals, allowing for the removal of the effect of distance from the shapes. Mean shapes (consensus configurations) were calculated for sites including relative warps. The Procrustes distances were then analysed by hierarchical clustering with 100000 bootstraps (Stevens 2012). Two outgroups namely *Etrumeus whiteheadi* and *Trachurus capensis* were chosen to root the cluster analysis of mean Procrustes distances. Differences among group means were tested using the Mahalanobis distance using single linkage, and the consistency of the groups was evaluated by computing the misclassification rates of new individuals (Stevens 2012).

2.4 Genetic material, primer and data collection, and analysis

2.4.1. Primer selection

Exon-primed, intron-crossing (EPIC) primers are an invaluable tool to molecular ecologists and taxonomists alike, due to their ability to deliver both inter- and intra-specific genetic information relatively cheaply (Li*, et al.* 2010). The homology of EPIC-amplified sequences can also be determined by comparison of their exon or intron (Li*, et al.* 2010). Nevertheless, there have been few markers developed for non-model teleost fishes, thus little work has been conducted using them (Hassan*, et al.* 2002, Touriya*, et al.* 2003, Ryynanen and Primmer 2006, Chenuil*, et al.* 2010, Li*, et al.* 2010). Nevertheless, ten EPIC primers previously shown to work on teleost fishes were chosen for this study (Table 2). The first aim was to identify EPIC markers that were amplifiable and showed polymorphisms within the sardine DNA. The markers that worked and showed polymorphisms within or between populations were then used for a population level assessment of genetic diversity amongst the Southern African Sardine.

Primers that amplify the mitochondrially encoded cytochrome oxidase I gene (mtCOI) have recently been designed at the Canadian Centre for DNA Barcoding (CCDB), and are the primary markers used in the global effort to genetically barcode the world's teleost ichthyofauna (Ivanova*, et al.* 2007). This marker has been used successfully on Clupeiformes, including

Sardinops sagax, locally (Steinke*, et al.* 2011, Cawthorn*, et al.* 2012) and abroad (Ward*, et al.* 2005, Ivanova*, et al.* 2007, Steinke*, et al.* 2009, Zemlak*, et al.* 2009, Keskİn and Atar 2013).

Table 3: Mitochondrial DNA and (mtDNA) exon-primed intron-crossing (EPIC), primers used on *Sardinops sagax*.

	Gene	Marker abbreviation	Primer name	Sequence set (5'-3')	Ta	Size of focal intron	Reference
mtDNA primer cocktail	Cytochrome Oxidase sub unit 1	COI	FishF2-t1	TGTAAAACGACGGCCAGTCGACTAATCATAAAGATATCGGCAC	46°C	652	(Ivanova, Zemlak et al. 2007)
			FishR2-t1	CAGGAAACAGCTATGACACTTCAGGGTGACCGAAGAATCAGAA			
M13 sequencing primers	-	-	M13F (-21)	TGTAAAACGACGGCCAGT	N/A	N/A	(Messing 1983)
			M13R (-27)	CAGGAAACAGCTATGAC			
EPIC primers	Opsin	Ops	OPS-1-F	GCTCATGGGCCTGCAGACCACAA	52°C	1022	(Touriya, Rami et al. 2003)
			OPS-1-R	CCTGCTCAACCTGGCCATGGC			
	Actin	Act	Act-2-F	GCATAACCCTCGTAGATGGGCAC	58°C		(Touriya, Rami et al. 2003)
			Act-2-R	ATCTGGCACCACACCTTCTACAA			
	Chymo-trypsin B	ChymB	ChymB-6-F	GCATGAGGGCTGTGACTCGGG	54°C	382	(Touriya, Rami et al. 2003)
			ChymB-6-R	ATCGTGTCCGAGGCTGACTGCAA			
	Myosin light chain 3	MLc	MLc-3-F	AGTAATGACGTCGCAGATGTTCT	54°C	674	(Touriya, Rami et al. 2003)
			MLc-3-R	CGACAGGTTCACTCTCGAGGAG			
	Calmodulin	Cam-3	Cam-3-F	TGACGGAGCTCTGCAGCACTGAC	54°C	491	(Touriya, Rami et al. 2003)
			Cam-3-R	GTGAGGAGGAGCTCCGTGAGGC			
	Glyceraldehyde-3-phosphate dehydrogenase	GPd	GPd-2-F	GCCATCAATGACCCCTTCATCG	46°C	303	(Hassan, Lemaire et al. 2002)
			GPd-3-R	TTGACCTCACCCTTGAAGCGGCCG			
	Aldolase B, intron 4	AldoB-4	Aldo-5-F	GCCAGATATGCCAGCATCTGCC	54°C	134	(Hassan, Lemaire et al. 2002)
			Aldo-3.1-R	GGGTTCCATCAGGCAGGATCTCTGGC			
	Gonadotropin-releasing hormone 3	GnRH	GnRH-2-F	AGAAGTGTGGGAGAGCTAGAGGC	Failed	-	(Hassan, Lemaire et al. 2002)
			GnRH-2-R	AGAGACACCACTTCTCCTGTACCC			
	Alpha tropomyosin	TR-1	Tr-1-F	AGGGAACAGAGGATGAGCTGGAC	52°C	841	(Hassan, Lemaire et al. 2002)
			Tr-1-R	TCTCAGCTTCCTCCAGCTTGGTG			
	Aldolase B	AldoB	AldoB1-1-F	GCTCCAGGAAAGGGAATCCTGGC	54°C	262	(Hassan, Lemaire et al. 2002)
			AldoB1-2-R	CTCGTGGAAGAAGATGATCCCGCC			

2.4.2. Material collection and preparation

After morphological and meristic data were collected, fish were rinsed in deionised water and scraped clean of scales behind the dorsal fin, where a 1cm^2 section of muscle tissue was removed. Tissue samples were then individually stored in 80 % ethanol at room temperature. DNA was extracted from adult and whole larvae using a standard phenol-chloroform-isoamyl alcohol (PCI) DNA extraction protocol, according to Barker et al (1998). Tissue was first patted dry and rinsed in distilled water to remove any residual ethanol that can inhibit the enzymatic activity of proteinase-k. The dried muscle tissue was then added to a 1.5ml Eppendorf tube with 400µl of extraction buffer (0.01M Tris, 0.005M NaCl, 250µM sodium dodecyl sulphate (SDS); pH8.3) and 50µl of proteinase-k. Sample tissue was homogenised using an inoculation loop and incubated on a rocking tray at 57°C overnight.

A 25:24:1 phenol:choloroform:isoamyl alcohol (PCI) mix was added to the DNA solution and gently shaken for five minutes before being spun down at 14 000RPM. The supernatant was then pipetted off and an equal amount of PCI was added to the supernatant and gently inverted for five minutes and spun down a second time. If an interphase layer was still present, a third PCI treatment was undertaken. A 24:1 chloroform:Isoamyl alcohol (CI) mixture was then added in equal volume to the supernatant and inverted for two minutes before being centrifuged at 14 000 rpm for one minute. The top aqueous phase was then removed and 2 - 2.5 volumes of 100 % molecular grade, ice cold ethanol were added before being placed in a freezer at -20°C overnight. Samples were then spun down for 20 minutes at 14 000RPM. Ethanol was then poured off and the pellets were washed with 800µl of 75 % ethanol and spun down for 20 minutes at 14 000rpm. The ethanol was poured off and the DNA pellets were dried in a vacuum desiccator for twenty minutes. DNA pellets were then dissolved in 100µl of molecular grade water.

DNA was checked for quality and quantity using a Nanodrop 3 000 spectrophotometer and subsequently run on a 1 % agarose gel, using 1 µl of loading dye and 5 µl of PCR product per sample. DNA quantity varied greatly between individuals (1 – 1 000 ng.µg^{-1} DNA per sample), with varying levels of phenolic and protein contamination (via spectral absorbance ratios of

DNA and other contaminants) determined using a Nanodrop 3 000 spectrophotometer. DNA was stored at -20°C before being used in PCR.

2.4.3. EPIC encoded data collection and analysis

Specimen DNA was amplified via Polymerase Chain Reaction (PCR) using several EPIC primers designed to amplify various genes (see table 3) (Messing 1983, Ivanova*, et al.* 2007). PCR reactions for all markers contained 12.5 µl EconoTaq PLUS GREEN 2X Master Mix (Lucigen), 0.84 µl forward and reverse primer (10 µM) and 11 µl H_2O. The PCR thermal cycle was [94 °C for 3 min], 34x [(94 °C for 30 s) (annealing temperature for 45 s) (72 °C for45 s)] and [72 °C for 10 min], [4 °C for ∞]. PCR product for each sample was checked for quality and quantity using a Nanodrop 3 000 spectrophotometer. PCR amplicons were run on a 3 % agarose gel for 5 hours at 100 volts, adapted from Touriya*, et al.* (2003). O'GeneRuler plus DNA Ladder, Ready-to-Use, 100 - 3 000 bp (Thermoscientific) and a reference sample was run with every gel. Gels were viewed on a Biorad Molecular Imager, Gel Doc™ XR+. A representative sample was sequenced for each marker using an ABI 3730 capillary sequencer at Inqaba Biotechnical Industries (Pty) Ltd, Hatfield, South Africa.

Gel images were scored using Image lab 4.1 (Bio-Rad Laboratories Inc.). Gel warping and distortion of fragment distance after electrophoresis was accounted for and standardised using various programme tools in Image lab, to attain the most accurate account of the data. Five percent of samples were reamplified and scored to test for repeatability. Samples were scored according to their predicted focal allele band size ranges. Scored allele lengths were then categorized into between 5bp and 20bp groupings and assigned a numerical number, representing different alleles according to GenAlex data requirements (Peakall 2012).

Data were then exported to GenAlex 6.4.1 for analysis. Analysis of molecular variance (AMOVA) was performed on the data, with samples grouped by site and region, relative to the total population. Phi_{PT} (Φ_{PT}) was calculated for within and among the sites and Phi_{RT} (Φ_{RT}) between two regions (in essence a Phi_{PT} value calculated between grouped sites), compared to

the total genetic variance (Peakall 2012). A Mantel test was performed between genetic and geographic distances with 10 000 permutations, using GenAlex (Peakall 2012). Allele frequencies, heterozygosity, deviations from Hardy-Weinberg equilibrium, Nei's genetic distance, and a PCA of the genetic distance matrix were calculated in GenAlex. Finally, rates of migration between sites were calculated using Migrate 3.2 (Beerli 2008). Default settings for a Bayesian analysis, with a burn-in of 1 000 and 500 000 iterations (5 000 000 parameter values) with adaptive heating on were run for the data, based on an island model of evolution. GenAlEx was used to calculate immigration between regional groups (those used to calculate Φ_{RT}) using default settings.

Patterns in population structure were assessed using Bayesian clustering analysis in the program Structure 2.3.4 (Pritchard*, et al.* 2000). The analysis was run using five MCMC simulations per run, with 1 000 000 iterations and 500 000 burn in, to test possible population clusters between one and ten (K). Structure harvester (Earl 2012) was used to determine the most likely combination of homogenous clusters (Evanno*, et al.* 2005).

A standard random representative sample (from KZN, 2012) was sequenced (the first allele of each gene) for reference with an ABI 3730 capillary sequencer at Inqaba Biotechnical Industries (Pty) Ltd, Hatfield, South Africa. Sequence electropherograms were edited using BioEdit version 7.0.9 (Hall 1999). Ambiguities in certain loci were compared assessed and compared to their respective electropherograms, and changed according to the signal peaks. The final sequences were trimmed to a length where definitive bp signals were obtained. The NCBI data base did not have any sequences of the genes tested, thus a BLAST search did not return any similar sequences (Altschul 1990).

2.4.4. mtCOI data collection and analysis

Specimen DNA was amplified via Polymerase Chain Reaction (PCR) using primers designed to amplify a 652 base pair region of the mtCOI gene (see table 1) (Messing 1983, Ivanova*, et al.* 2007). The PCR reactions for both markers contained 21 µl H_2O, 4 µl 10x buffer, 1.8 µl $MgCl_2$

(25 mM), 0.84 µl forward and reverse COI primer (10 µM), 0.15 µl dNTP mix (10 mM), 1 µl BSA (10 mM) and 0.2 µl Supertherm Taq polymerase 5 $u.\mu l^{-1}$ (Lucigen), and 1 µl template DNA (2 – 200 $ng.\mu l^{-1}$). The PCR thermal cycle was [95 °C for5 min], 30x [(94 °C for 30 s) (55 °C for 45 s for the 3/550 primer and 58 °C for 45 s for the mtDNA primer) (72 °C for 45 s)] and [72 °C for 10 min], [4 °C for ∞]. PCR product for each sample was checked for quality and quantity (via spectral absorbance ratios of DNA and other contaminants) using a Nanodrop 3 000 spectrophotometer and amplicons were subsequently separated by electrophoresis on a 1 % agarose gel stained with 100 µl ethidium bromide. Gel electrophoresis was run for 60 minutes at 100 volts, using 1 µl of loading dye and 5 µl of PCR product per sample. Samples were viewed on a Biorad Molecular Imager, Gel Doc™ XR+. Samples were then sequenced using an ABI 3730 capillary sequencer at Inqaba Biotechnical Industries (Pty) Ltd, Hatfield, South Africa.

BioEdit version 7.0.9 (Hall 1999) was used to align and edit the mtDNA sequence electropherograms. A Clustal W multiple alignment algorithm was used to align sequences, which were subsequently checked and realigned by eye. Ambiguities in certain loci were then checked for by choosing the strongest signal on the chromatogram for each locus. The final alignments were trimmed to a length of 652 bp. A sequence search for highly similar sequences was conducted on a consensus sequence of all sequences, using BLAST (Altschul 1990). Sequences of *Sardinops sagax* and *Sardinops melanostictus* from various localities around the world were retrieved from the NCBI nucleotide collection and added to the sequence set for further analysis (see Table 4). Only sequences retrieved from BLAST of 652bp or more were included to preserve information in data when trimming sequences for analysis. All sample sequences were uploaded onto GenBank with accompanying information (submission # 1705716).

Table 4: COI sequence information of samples included for phylogenetic analyses, from the NCBI Genbank website.

Species	Number of sequences	Location (Co-ordinates)	Collection date	Accession number	Reference
Sardinops sagax	2	Durban, KZN, South Africa (29.945 S 31.00 E)	30-Jun-05	JF494411 JF4944112	(Steinke*, et al.* 2011)
	2	Agulhas Bank, Mossel Bay, South Africa (34.873 S 21.618 E)	26-Apr-08	JF494409 JF494410	(Steinke*, et al.* 2011)
	5	British Columbia, Canada (48.93 N 125.34 W)	15-Oct-05	FJ165120 FJ165121 FJ165123 FJ165126 FJ165128	(Steinke*, et al.* 2009)
Sardinops melanostictus	3	Yokosuka, Japan (35.00 N 139.50 E)	07-Jul-05	FJ952841 FJ952842 FJ952843	(Zhang and Hanner 2011)

DnaSP was used to construct haplotypes from the sequenced samples (Scheet and Stephens 2006, Librado and Rozas 2009). Gaps were ignored and all variable sites were used. MrModelTest 2.0 (Nylander 2004) and Modeltest 3.7 (Posada and Crandall 2005) were used to search for the best fit model of evolution that fitted each genetic marker dataset. The HKY model of Hasegawa*, et al.* (1985) was chosen to construct the neighbour joining (NJ) and maximum likelihood (ML) trees in PAUP 4.1b (Swofford 2002) as well as the Bayesian tree searches, using MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). The mtDNA Bayesian trees were rooted with *Pomacentrus baenchschi* (accession number:JF435106) and *Sardina pilchardus* (accession number: KC501229) sequences, retrieved from GenBank after a BLAST search of the NCBI nucleotide collection (Altschul 1990). Sequences of *Sardinops sagax* and *Sardinops melanostictus* of various localities from Genbank were also included (Table 4). The NJ and ML trees were bootstrapped for 1 000 iterations. The Bayesian trees constructed with haplotypes were created using 4 Markov chains of 10 000 000 generations each, with samples every 100 generations (average standard deviation of split frequencies less than 0.01). The first 2 500 000 trees were discarded as burn in (25% of the generations is considered sufficient for burn-in), with the rest of the genealogies used to construct a 50 % majority-rule consensus tree. Balance of chain length and how frequent the posterior was sampled during the run was calculated in Tracer (Rambaut*, et al.* 2014), and showed an effective sample size of 871.916, which was sufficient (generally an effective sample size of 100 - 200 is deemed sufficient). The

raw trace plot in Tracer showed that the chains in the MrBayes analysis were long enough and sample frequency indicated that the MrBayes run was sufficiently sampled (Rambaut*, et al.* 2014). Estimations of population expansion times were calculated using Tau (τ) = 2μt, where μ is the mutation rate and t the time of expansion. Generation time was estimated to be 2 years (van der Lingen and Durholtz 2004) and a mutation rate of COI of 2 % per million years was employed, typical of most teleost fishes (Liu*, et al.* 2006). Tau (t) was calculated in Arlequin version 3.5.1.3 (Excoffier and Schneider 2005) and expansion time was estimated with the program http://www.uni-graz.at/zoowww/mismatchcalc/mmc1.php (Schenekar and Weiss 2011). The mismatch distribution together with its population expansion parameters were calculated in Arlequin and the validity of the expansion model was tested using a significance of 0.05 and run for 1000 bootstrap replicates.

A haplotype network was constructed using the median joining method (Bandelt*, et al.* 1999) for the COI sequence data, with gaps coded as missing, and run at a 95 % confidence interval. The number of specimens assigned to haplotypes in DNAsp 5.0 was also supplied for each branch on the minimum spanning tree network (Bandelt*, et al.* 1999).

Data were then exported to GenAlex 6.4.1 for analysis, and was coded as haploid. Analysis of molecular variance (AMOVA) was performed on the data, with samples grouped by region (RSA, Canada and *S. melanostictus*). Phi_{PT} was calculated for within and among the sites and compared to the total (Peakall 2012). A mantel test was performed between genetic and geographic distances with 10 000 permutations, using GenAlex (Peakall 2012). Nei's genetic distance and rates of immigration were also calculated using GenAlex. Finally a PCA for the genetic distance matrix was conducted in GenAlex 6.4.1.

Chapter three:

Morphological characteristics of the sardine (*Sardinops sagax*) in South Africa

3.1. Introduction

The South African sardine (*Sardinops sagax*) occurs throughout the year along the western coast of South Africa and Namibia, from Luderitz to the Agulhas bank as far as Port Elizabeth. Seasonal winter migration during the winter “sardine run” extends the range of *S. sagax* approximately 700km northward into the sub-tropical waters of KZN (van der Lingen, Hendricks*, et al.* 2010). The large biogeographic range of the South African *S. sagax* spans two oceans and a range of physical and ecological conditions and ecosystems. The diverse ecosystem parameters over this large spatial scale certainly play a role in shaping the biology, ecology and thus the population structure of *S. sagax* in South Africa (de Moor and Butterworth 2009) as they have globally (Parrish*, et al.* 1989, Agüero and Rodríguez 2004). Understanding population structure is important to developing accurate modelling of fish stocks and thus the effective management of any fishery (Agüero and Rodríguez 2004).

Ecological differences between biogeographic areas shape the phenotypic and behavioural characteristics of individuals (Cadrin 2000). Differences in mean ocean temperature, salinity, food type and availability, are some factors are important factors affecting the distribution of *S. sagax* along the western, south east and eastern coastlines of South Africa. Thus, phenotypic traits may be useful in testing population structuring between biogeographically separated *S. sagax,* and have been investigated previously (de Moor and Butterworth 2009, Wessels 2009, van der Lingen, Hendricks*, et al.* 2010).

Thus far, stock structure of *S. sagax* in South Africa has been successfully delineated using meristic, morphometric and ecological data (De Goede and Van der Lingen 2005, van der

Lingen*, et al.* 2009, Wessels 2009, van der Lingen, Hendricks*, et al.* 2010) and has been successful in delineating stock structure. The aim of these studies has focused mostly on the commercially important *S. sagax* west of Cape Agulhas and the eastern stock on the Agulhas bank. Nevertheless, van der Lingen, Hendricks*, et al.* (2010) have focused specifically on meristic characteristics of *S. sagax* of the KZN sardine run, and included KZN sardines in a morphometric study of the South African sardine.

Freon*, et al.* (2010) suggested the use of similar methods to determine whether sardines undertaking the migration up into KZN are morphologically disparate from the Western and Eastern Cape stock. An investigation of this nature would have important implications for the seasonal KZN beach-seine fishery and to its management, for the development of a multi-stock model for the South African sardine (Freon*, et al.* 2010).

Indeed, some investigation has been conducted on the KZN sardines. Sardines caught in KZN have been shown to exhibit a lower mean conditioning factor, different mean vertebrae frequencies and a shorter caudal length than sardines on the Agulhas Bank and on the Western Cape coast (van der Lingen, Hendricks*, et al.* 2010). Differences in morphology between western and south-eastern stocks (including KZN samples) have also been observed using multivariate morphometrics. Nevertheless, a morphometric study focusing specifically on KZN sardines has yet to be completed. These previous studies have suggested some form of structuring between KZN and southern Benguela sardines.

Therefore KZN sardines are expected to show morphological differences from the Western Cape stock, and possibly between biogeographic regions. To test this hypothesis, sardines were collected at four sites along the South African coast, including samples from the KZN sardine run and subjected to a range of morphometric and meristic analyses to determine whether phenotypic differences existed between sardines from the KZN component of the sardine run and the Western Cape or Agulhas bank.

3.2 Materials and methods

Refer to chapter 2, 2.2. Meristic data collection and analysis and 2.3. Morphological data collection and analysis.

3.3. Results

A total of 1040 sardines were collected for meristic analysis from the four sites along the South African coast, including a three year temporal scale for KZN (Table 2). Morphometric analysis included 540 sardines, with 500 individuals excluded as they were frozen for longer than three months, which influenced the shape of sardines (Wessels*, et al.* 2010), and a further 86 outliers excluded mostly due to the presence of distended stomachs or bent or torn flesh (95 % confidence interval).

3.3.1 Meristic results

No sexual dimorphism was detected, as a non-parametric Kruskal-Wallis test showed no significant sex based groupings (Kruskal-Wallis ANOVA: $H_{1, 150} = 22.04$, $p = 0.779$). Sex ratios for all sites were not significantly different from an expected 1:1 sex ratio (Table 5), except KZN for the years 2011, 2012 and the combined KZN grouped samples (2011 - 2013). The KZN sample for 2013 showed the largest male skewed sex ratio, although still not significantly so ($\chi^2 > 0.05$, Table 5). Nevertheless, the low significance may have been due to the small sample size (n = 17), thus limiting the confidence in this result. Cape Town was the only site with an equal sex ratio and Port Elizabeth showed a negatively skewed male sex ratio. All other sites showed male skewed sex ratios. The sex ratio was significantly different between KwaZulu-Natal and Port Elizabeth (Tukey HSD: $p < 0.001$) but not among the other sites (homogeneity subset test: $p < 0.080$).

Table 5: Sex ratio for the four *Sardinops sagax* sites and results of a Chi² test for deviation from an expected 1:1 sex ratio. A combined KZN result is also presented. Significant values are indicated in bold.

Site	Sex ratio (percent male)	*N*	*df*	Chi²	Asymptotic sig.
Cape Town	50	200	1	0.000	1.000

Mossel Bay	56	200	1	2.880	0.090
Port Elizabeth	44	201	1	2.632	0.105
KZN 2011	61	204	1	9.490	**0.002**
KZN 2012	60	217	1	9.332	**0.002**
KZN 2013	71	17	1	2.882	0.090
KZN (2011-2013)	61	438	1	21.041	**0.000**

Vertebral counts between the three years (2011, 2012 and 2013) for KZN were very similar according to a Kruskal-Wallis test ($H_{3, 49} = 0.03$, $p = 0.943$), and were thus combined for comparison with the other three sites. Vertebral counts between Cape Town and the other three sites (KZN 2011-2013, Mossel bay and Port Elizabeth) were significantly different from each other (Tukey's HSD: $p < 0.001$), however there were no differences between the other sites (Tukey's HSD homogenous subset: $p = 0.354$). Figure 3 illustrates the frequency distributions of vertebrae of sardines from the four sites. Individuals from Cape Town displayed a median of 50 vertebrae, whilst those from other sites exhibited a median vertebra number of 49, accounting for more than 60 % of each population (Figure 3). Cape Town had a mean vertebra count of n = 49.695, Mossel Bay n = 49.005, Port Elizabeth n = 49.080, and KZN n = 49.000. No sardines from KZN (all years) or Mossel bay were recorded with 51 vertebrae (Figure 3). A graded decrease in individuals with 49 vertebrae and an increase of individuals with 48 and 50 vertebrae was observed between Mossel Bay, Port Elizabeth and KZN.

Overall wet body mass and caudal length was significantly higher for sardines from Cape Town, Mossel Bay and Port Elizabeth, than those from KZN (Tukey's HSD: $p < 0.001$; Figure 4). There was no difference between the former three sites according to a Tukey's HSD *post hoc* test. A decrease in standard deviation was observed among all sites from Cape Town to KZN for both wet body mass and caudal length (Figure 4).

A non-parametric Spearman's rank correlation between wet body mass and caudal length was conducted after the assumptions of normality were not met using a 1 sample Kolmogorov-Smirnov test, and showed a strong positive correlation ($R^2 = 0.956$) for all data (Figure 5). Individual correlation coefficients were also calculated for each site. A decrease in the correlation coefficient was observed from Cape Town ($R^2 = 0.970$), to Mossel Bay ($R^2=0.856$), Port Elizabeth ($R^2=0.802$) and KZN (across all years), illustrated the lowest correlation coefficient ($R^2=0.773$).

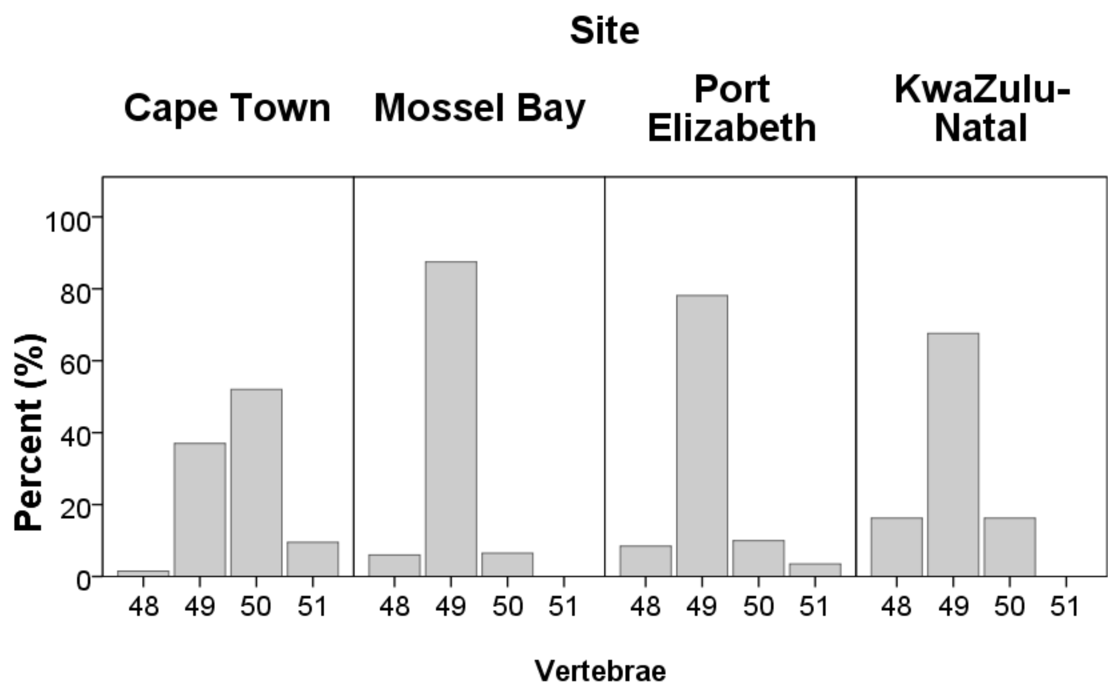


Figure 3: Frequency distributions (percentage) of vertebrae counts for sardines from four sites along the South African coastline.

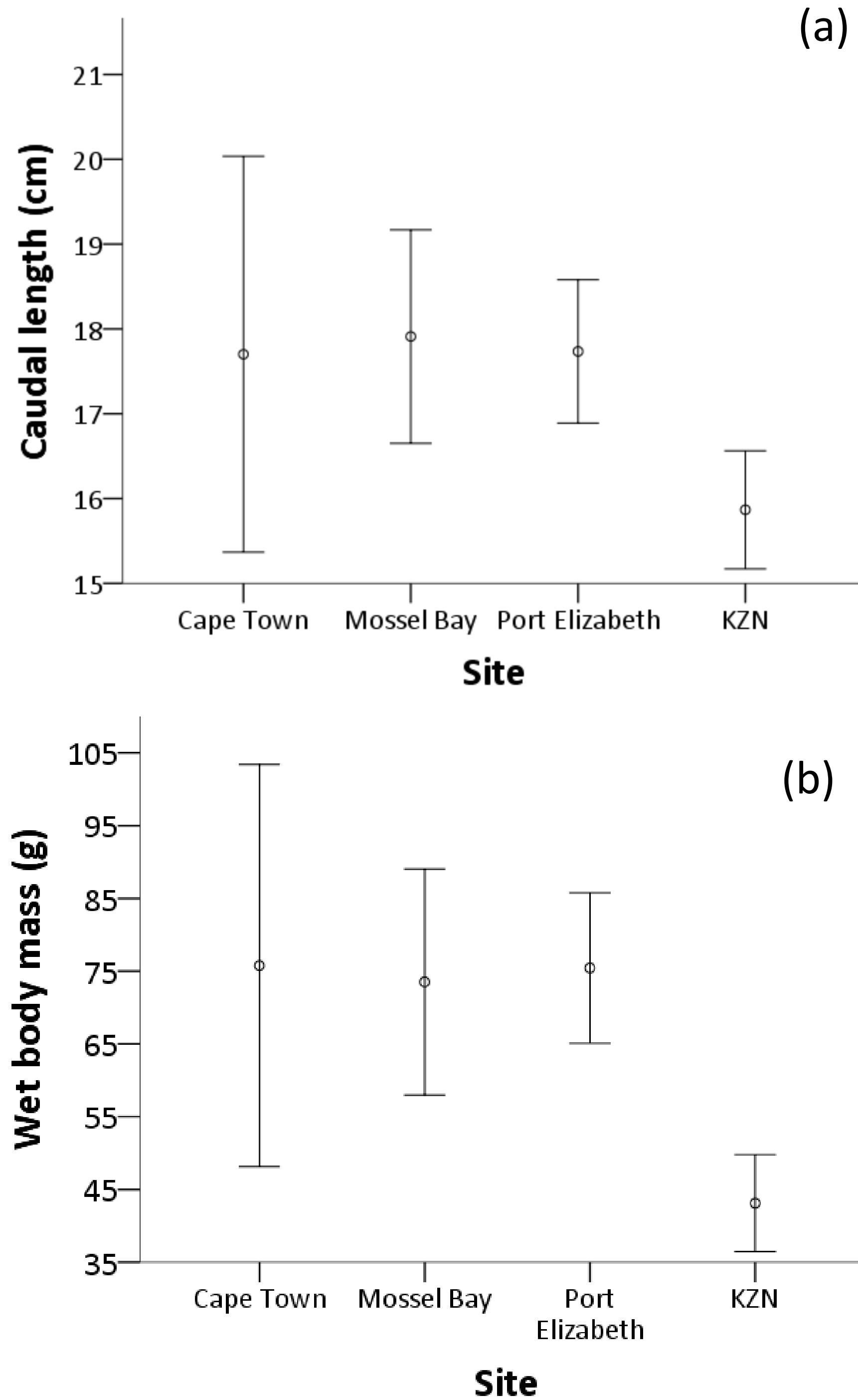


Figure 4: (a) mean caudal length and (b) Mean wet body mass of *Sardinops sagax* at four sites along the South African coastline. The bars indicate the standard deviation.

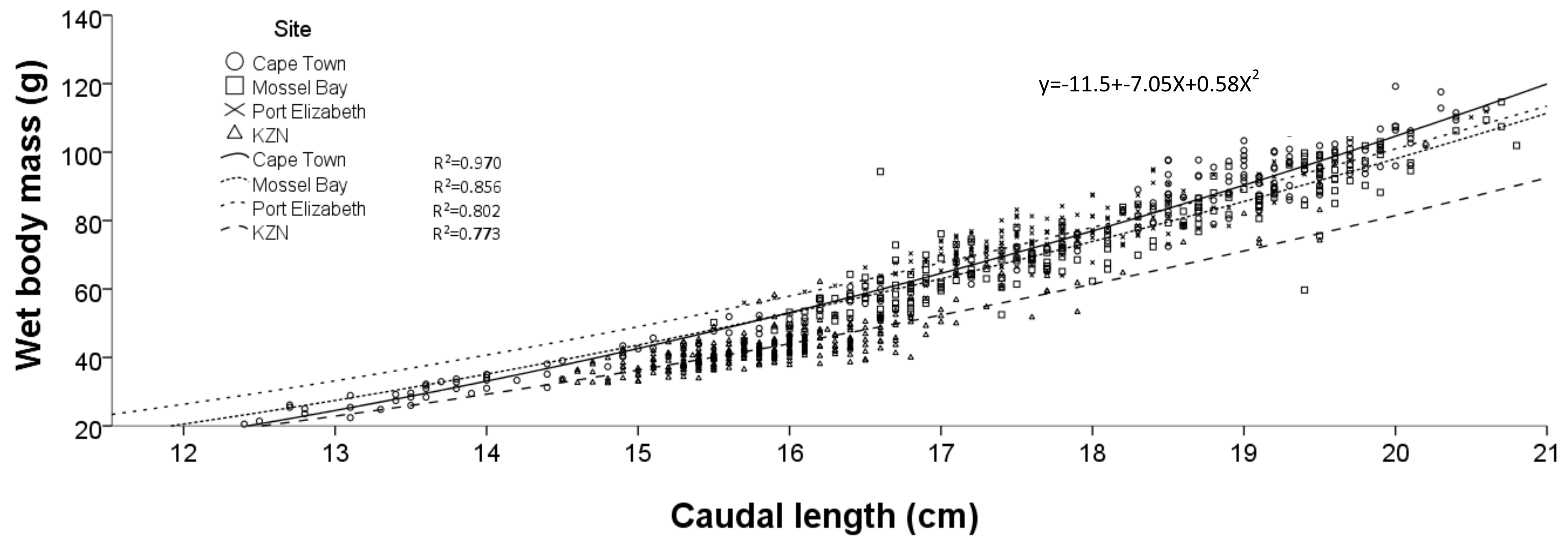


Figure 5: Mass at length scatter plot for *Sardinops sagax* collected at four sites along the South African coastline between January and November 2011 to 2013, with fitted power curves representing mean weight at length of fish. Power equation shown for all data combined.

3.3.2. Morphometric results

Pairwise population morphometric distances (the 21 truss measurements described in Figure 2) differed between the four populations ($F_{5.201}$ = 16.372 – 130.241, $p < 0.05$) except t19 ($F_{3, 1.571}$ = 16.7108, p =0.372), according to a one way ANOVA (table 8). Truss variables that did not differ between sites were the distance between the snout and the point where the jaw hinge meets the operculum and the body area between the front and back insertion points on the dorsal fin (t8), and the ventral fin and front insertion point on the anal fin (t19). All other truss measurements were significantly different between the sites tested (Figure 6).

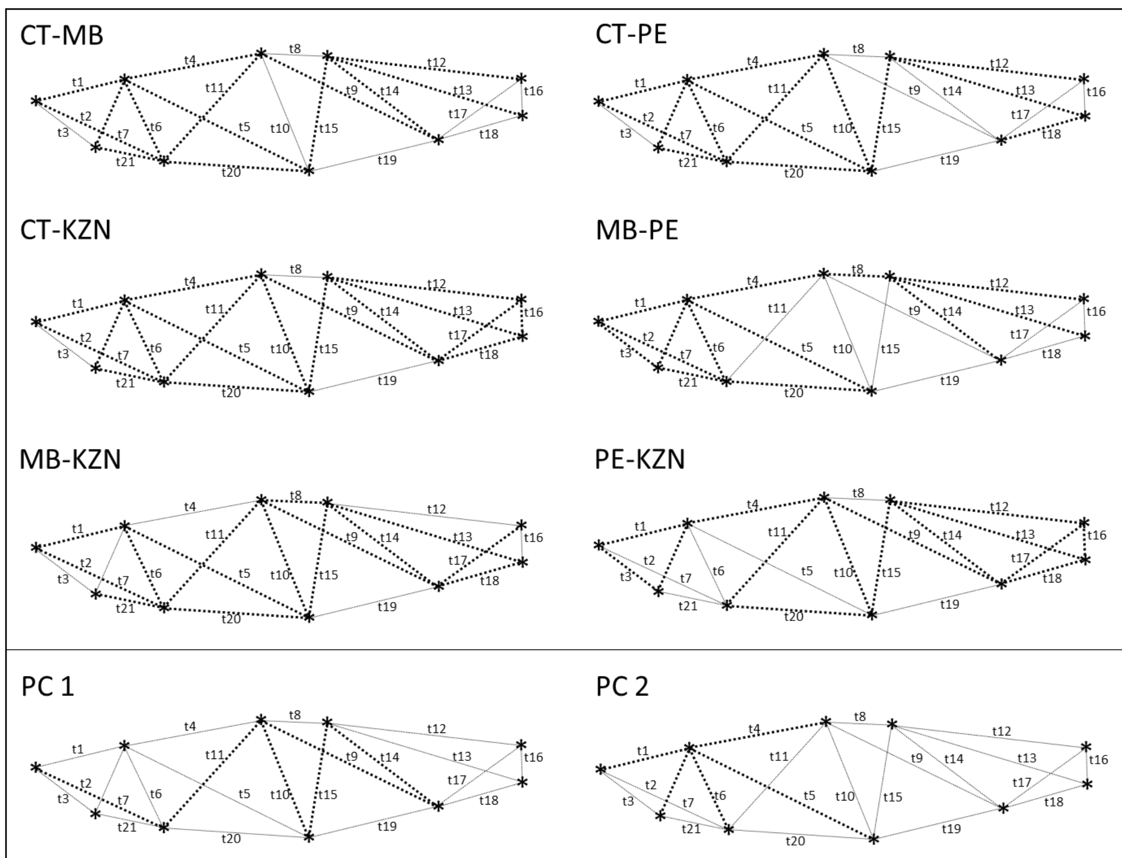


Figure 6: Tukey's HSD *post hoc* test of size corrected truss measurements for *Sardinops sagax* at four sites along the South African coastline. Significant differences in truss length are represented as bold dotted lines. PC1 and PC2 show truss measurements that showed the greatest contribution to variance among samples (bold dotted lines). CT = Cape Town, MB = Mossel Bay, PE = Port Elizabeth, KZN = KwaZulu-Natal.

The difference in truss size between sites indicates an increase or decrease in height and length of sardines. The largest variability in truss size between sites (F statistic, table 8) were the distances within the area of the head (t1, t2, t3, t6 and t7), and the lowest were the vertical and horizontal distances behind the operculum (pre dorsal: t4, t5 and t11) and the horizontal and perimeter distances behind the front dorsal insertion and the front anal insertion (post dorsal: t8, t9, t10, t12, t13 and t19).

Table 6: One way ANOVA illustrating significance (*= Significant difference; P<0.05) of size corrected truss measurements for *Sardinops sagax* at four sites along the South African coastline. Results include PCA eigenvectors with eigenvalues and cumulative variance for the first four principal components (PCs) of a Principal Coordinates Analysis.

One way ANOVA			PCA eigenvectors	
Truss measurement	Sig.	F	PC1	PC2
t1	*	112.622	-0.094	-0.184
t2	*	102.780	-0.140	0.053
t3	*	73.105	-0.064	0.102
t4	*	31.472	0.061	0.222
t5	*	46.112	0.071	0.220
t6	*	94.275	-0.043	0.185
t7	*	130.241	-0.070	-0.117
t8	*	24.427	0.019	0.134
t9	*	30.592	0.124	-0.024
t10	*	39.983	0.114	-0.036
t11	*	42.655	0.109	0.042
t12	*	41.975	0.089	-0.137
t13	*	43.630	0.102	-0.138
t14	*	25.879	0.114	-0.113
t15	*	34.421	0.109	0.000
t16	*	28.925	0.065	-0.052
t17	*	41.671	0.050	-0.059
t18	*	46.476	0.027	-0.038
t19	0.372	16.711	0.077	0.038
t20	*	40.810	0.097	0.067
t21	*	79.277	-0.093	-0.047
Eigenvalue			0.0212	0.0026
Cumulative variation (%)			71.457	80.261

The first two principal components (PC) of the PCA for the size corrected truss measurements accounted for 80.26 % of the total variance explained and were thus a good approximation of the data (Figure 6). The remaining PCs accounted for approximately 20 % of the variance each.

PC3 accounted for 4.75 %, PC4: 3.27 %, PC5: 2.66 and PC6 2.20 %. All eigenvectors were between negative one and one, indicating the effect of size had been removed from the data (Table 6). The greatest contributor to variance in PC 1 was t2 followed by t9, t10 and t11, and t14 and t15 (Table 6). Truss variables associated with the second landmark (top front corner of operculum, below occipital ridge; Table 2) showed the greatest contribution to variance in PC 2 (Table 6). Truss measurements t3 and t16-t21, all lying on the ventral side of the fish perimeter from the tail to the snout, showed the least contribution to the variance between both PCs.

The PCA plot showed substantial overlap of the 95 % confidence intervals (Figure 7). Nevertheless, some differentiation was observed between sites. The KZN samples were mostly negative according to PC1, with Mossel Bay approximately neutral and Port Elizabeth and Cape Town positive (Figure 7). KwaZulu-Natal, Port Elizabeth, and Mossel Bay had a negative loading on PC2, whereas Cape Town showed a positive loading (Figure 7).

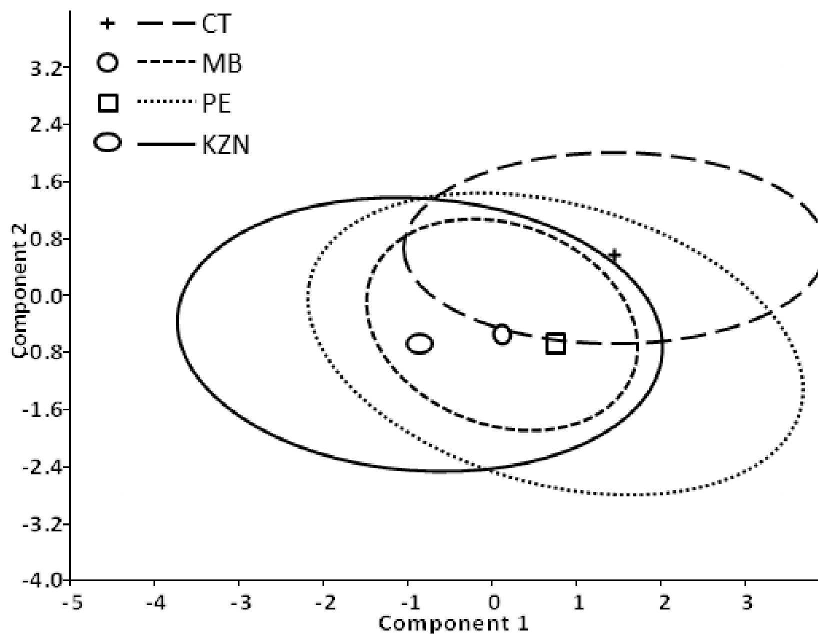


Figure 7: Principal Coordinates Analysis of size corrected truss measurements for PC1 and PC2, with 95 % confidence interval ellipses, for *Sardinops sagax* at four sites along the South African coastline. CT = Cape Town, MB = Mossel Bay, PE = Port Elizabeth, KZN = KwaZulu-Natal.

Seventeen truss measurements were included in the DFA, with t18-t21 removed subsequent to failing a tolerance test (a statistical version similar to the probability interval). Overall discrimination between sites was significant according to a DFA (Wilk's-Lambda = 0.088, F = 25.63, $p < 0.001$), thus validating the groupings (Table 7). The canonical analysis identified three discriminant functions although only the first two were significant (chi^2 = 1064.5, 572.9 and 148.5; df = 51, 32 and 15; p = 0.001, 0.001 and 0.09). DF1, DF2 and DF3 had eigenvalues of 2.037, 1.609 and 0.399. DF1, DF2 and DF3 accounted for 50.4 %, 39.8 %. and 9.9 % of discriminant capacity. The highest loadings for DF1 were three truss measurements found in the head of the fish and three in the body. For DF2 the highest loadings were for the tip of the snout to the front insertion of the dorsal fin and three measurements in the body (Figure 8). DF1 comprised negative values for truss variables that lay along the perimeter of samples from the snout to the tail (t1, t4-t6) and for the vertical trusses on the body behind the operculum to the back insertion of the dorsal fin to the ventral fin (t8, t10-t12). DF2 had negative trusses; these were trusses radiating from the top operculum insertion point (t1, t4-t6), and three trusses radiating diagonally across the body(t5, t14, t15).

Table 7: Summary results for the discriminant function analysis of size corrected truss measurements for function 1 and function 2, for *Sardinops sagax* sampled at four sites along the South African coastline. Truss measurements t18 - t21 were excluded after failing a tolerance test.

	Standard coefficient for canonical variables			Discriminant function analysis		
truss measurement	DF1	DF2	DF3	Wilk's Lambda	*F*	*P*
t1	-2.643	-1.151	-1.857	0.593	102.780	0.000
t2	1.779	0.176	1.469	0.672	73.105	0.000
t3	1.948	-0.146	-0.273	0.827	31.472	0.000
t4	-0.011	-1.948	0.065	0.765	46.112	0.000
t5	-0.736	-0.267	0.280	0.614	94.275	0.000
t6	-0.303	-0.462	-0.175	0.535	130.241	0.000
t7	0.423	0.729	-0.077	0.860	24.427	0.000
t8	-0.697	-0.958	-0.581	0.831	30.592	0.000
t9	0.027	1.553	1.020	0.790	39.983	0.000
t10	-1.904	1.377	0.190	0.779	42.655	0.000
t11	-0.508	0.131	-2.320	0.781	41.975	0.000
t12	-0.570	0.263	0.833	0.775	43.630	0.000
t13	1.524	1.009	0.551	0.853	25.879	0.000
t14	0.570	-0.825	0.290	0.813	34.421	0.000
t15	-0.022	-0.043	-0.569	0.838	28.925	0.000
t16	0.298	0.048	1.082	0.783	41.671	0.000
t17	1.232	0.654	1.062	0.763	46.476	0.000

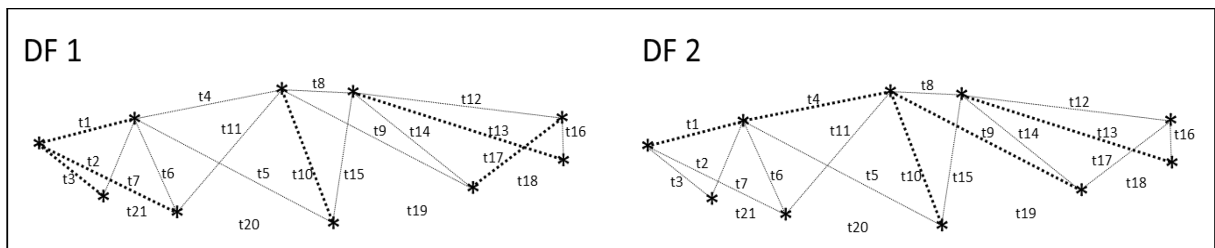


Figure 8: Truss measurements that loaded highest in the discriminant function analysis for each of the two significant discriminant functions DF1 and DF2 (bold dotted lines), for *Sardinops sagax* sampled at four sites along the South African coastline. Truss measurements t18-t20 were excluded after failing a tolerance test and are thus not depicted.

The 95 % ellipses showed substantial overlap between sites (Figure 9). The Mossel Bay, KZN, and Port Elizabeth centroids grouped relatively closely, whereas Cape Town was more distinct. Loadings for Cape Town were mostly negative for DF1, whereas the other three sites loaded

positively (Figure 9). The DF2group centroids were negative for KZN, Port Elizabeth and Cape Town, and positive for Mossel Bay (Figure 9). In total, 89.2 % of individuals were correctly reclassified into their original groups, using Mahalanobis distance (Table 8). Only 4.6 % Cape Town individuals were reclassified to different groups (Mossel Bay and Port Elizabeth, Table 8). Approximately 16% of Mossel Bay sardines were reclassified to Cape Town, Port Elizabeth and KZN. Only 5.9 % of the KZN sardines were reclassified, all to Mossel Bay (Figure 9).

Table 8: Reallocation of individuals in new groups, according to discriminant functions, for *Sardinops sagax* at four sites along the South African coastline.

Site	Predicted group membership				Total (%)
	Cape Town (%)	Mossel Bay (%)	Port Elizabeth (%)	KZN (%)	
Cape Town	95.4	1.9	2.8	0	100
Mossel Bay	3.3	84.2	10.4	2.2	100
Port Elizabeth	1.4	6.2	92.5	0	100
KZN	0	5.9	0	94.1	100

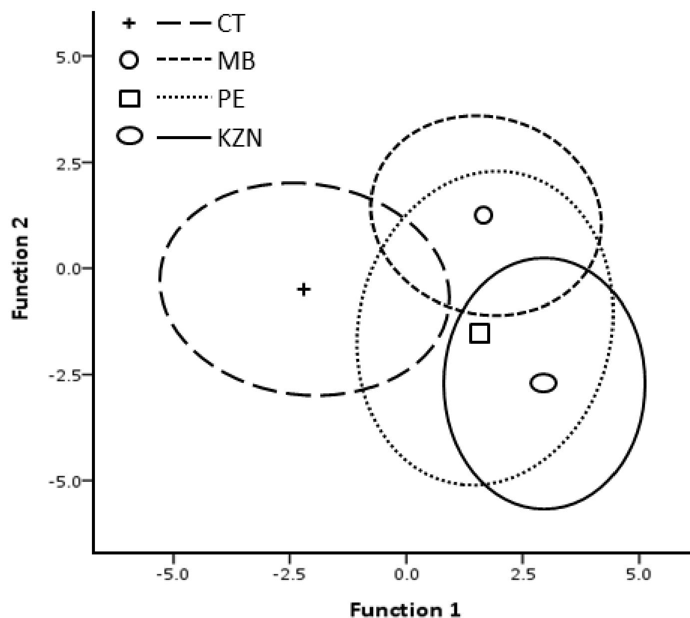


Figure 9: Discriminant function analysis of size corrected truss measurements for function 1 and function 2, with 95 % confidence interval ellipses, for *Sardinops sagax* at four sites along the South African coastline. Truss measurements t18 - t20 were excluded after failing a tolerance test.

Cluster analysis of mean sardine shape was conducted for four populations; Cape Town, Mossel bay, Port Elizabeth and KZN (Figure 10). The Cluster analysis supported the separation of the mean shapes of Cape Town, Mossel Bay and Port Elizabeth from KZN with 100 % confidence. The mean Cape Town shape separated from the Mossel Bay and Port Elizabeth mean shape with 94 % confidence. Mossel Bay and Port Elizabeth separated further, with 75 % confidence.

The difference in geometric shape was mainly focused on the head shape, attributed mostly to the increase of operculum size and the distribution of landmark 2 (Top front corner of operculum, below occipital ridge). Body size also explained variation between sites (Figure 11). Sardines from KZN and Port Elizabeth displayed a reduction in head size (both on the vertical and horizontal axes). However, this was less pronounced in the KZN samples due to the overall reduction in their absolute body shape (Figure 11). These observations were supported by the multivariate distances (Figure 6). Mossel Bay samples also showed a reduction in head size, relative to body size, however to a lesser degree than the latter two. Sardines from Cape Town showed a slight increase in head size from the mean shape (Figure 11). Sardines from Port Elizabeth fish showed a large increase in dorso-ventral depth and length, and exhibited the largest increase in dorsal height and overall body length, followed by Cape Town, Mossel Bay, and KZN (Figure 11). KZN and Port Elizabeth showed an increase in posterior size between the back dorsal fin and front anal fin insertion point to both caudal fin insertion points, whereas sardines from Cape Town and Mossel Bay showed a decrease in the same measurements, but an increase in the area between the ventral fin and front anal fin insertion points.

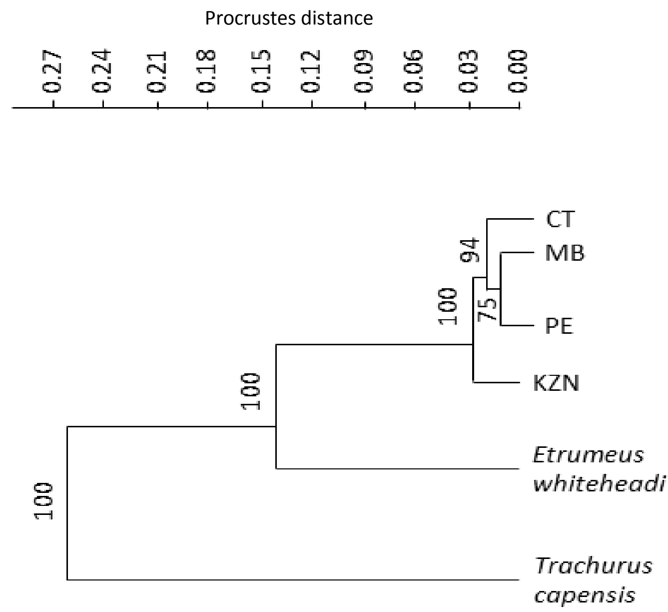


Figure 10: Cluster analysis of mean Procrustes distances separating clusters of sardines from each of the four sites and two outgroups, *Trachurus capensis* and *Etrumeus whiteheadi*. Numbers represent nodal support (Bootstrap percentages based on 100 000 iterations).

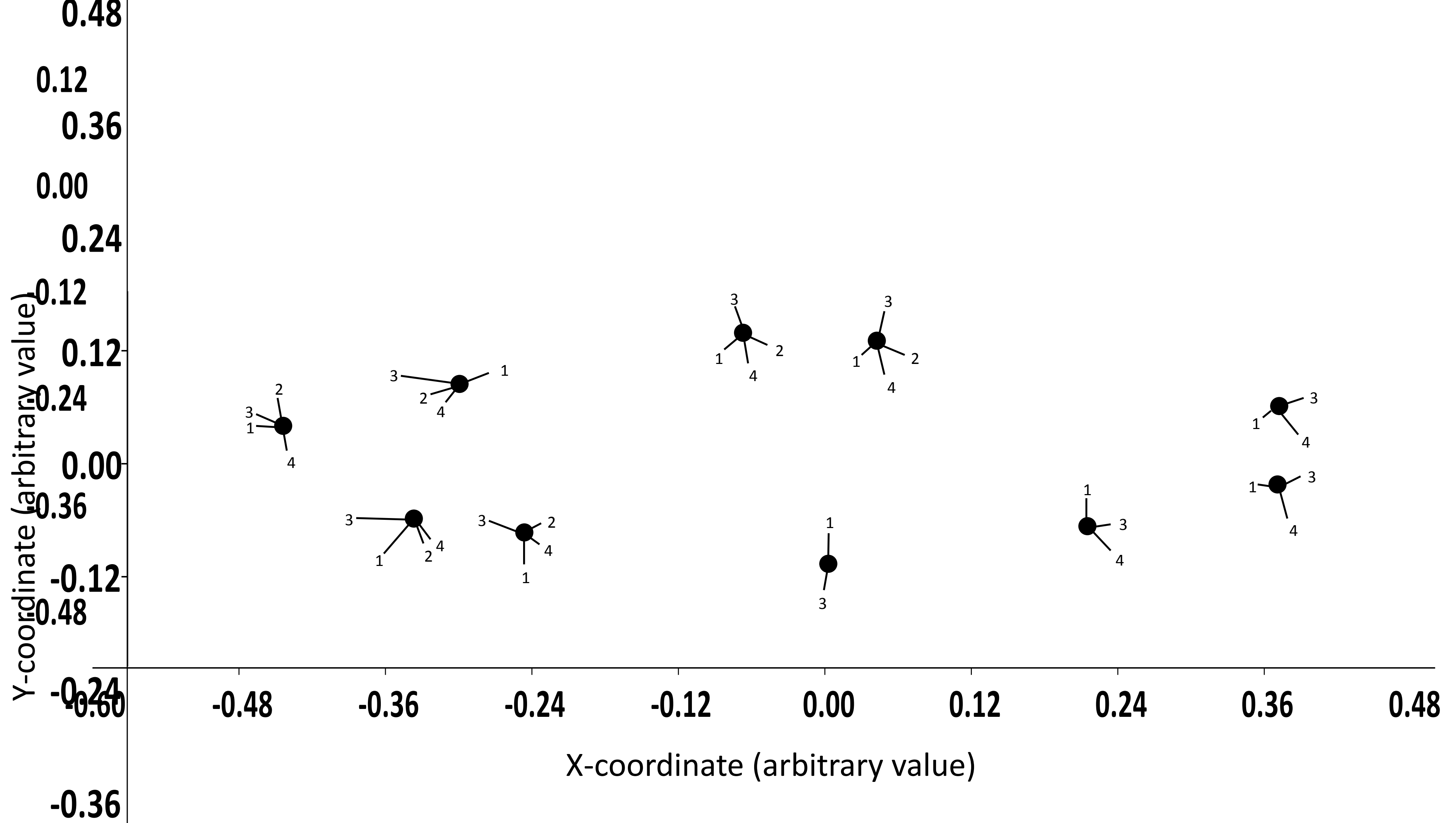


Figure 11: Relative warping (Lines represent relative warp of a population mean shape from an overall consensus shape) of ten landmarks (black dots) placed on individuals, according to discriminant functions, for *Sardinops sagax* at four sites along the South African coastline. Cape Town = 1, Mossel Bay = 2, Port Elizabeth = 3, KZN = 4.

3.4. Discussion

Analysis of morphometric and meristic data ascribed no sexual dimorphism to the South African sardines. However, the presence of heavily egg laden females should also be considered, and was not covered by the analyses, apart from the removal of seven females from the Cape Town sample as they had severely distended stomach cavities. Sex ratios were equal and not different from results previously reported (Akkers and Melo 1996; 1998, van der Lingen and Durholtz 2004, van der Lingen*, et al.* 2006, de Moor and Butterworth 2009, van der Lingen, Hendricks*, et al.* 2010, Wessels*, et al.* 2010).

The change in vertebra frequency in sardine sampled along the coastline from the Western Cape to KZN was significant. The mean and modal numbers of vertebrae concurred with those of Wessels (2009). These differences in number of vertebrae are attributed to differences in environment, especially temperature, during the larval stages when the number of vertebrae is set (Begg*, et al.* 1999, Florence*, et al.* 2002, McDowall 2008). Sardines spawned in warm waters tend to develop fewer vertebrae than those spawned in cooler water (Hulme 1995) and a similar response known as Jordan's rule; where a positive correlation between number of vertebrae and increasing latitude is often observed (McDowall 2008). The differences in number of vertebrae are consistent with those expected in various parts of the sardine's range across different biogeographical regions described by Teske*, et al.* (2011). Sardines on the Agulhas bank had fewer vertebrae and less variability in vertebral count than sardines caught in the Western Cape west of Cape Agulhas. This suggests that the sardines trapped between the two regions were spawned there. However, vertebral count is very loosely correlated to environmental conditions and often intra-population variance of vertebral counts can be large (Florence*, et al.* 2002). Added to this, variability in the spatial distribution of larvae and adults can mislead the interpretation of results (Wessels 2009). Thus, the use of vertebral counts for population discrimination should account for these factors (Begg*, et al.* 1999, Wessels 2009).

The reduction in the standard deviation of both mass and length of individuals from Cape Town to KZN may have been due to the fact that the shoals of sardines sampled in Cape Town

had greater differences in size class compared to the shoals sampled in the other regions. The weight to length relationship of sardines decreased from Cape Town to KZN. The reduction in weight at length observed in the KZN population has been ascribed to the energetic costs incurred in swimming from the Agulhas Bank to KZN during the sardine run (Freon*, et al.* 2010, van der Lingen, Hendricks*, et al.* 2010). The greater deviation (r^2 value) from the weight at length of individuals in KZN may also be related to the energetic costs of migration up the east coast, where fish may experience varied food availability and workload. These observations were similar to those based on other recent sardine runs, described by van der Lingen, Hendricks*, et al.* (2010). Differences in environmental conditions between shoals of sardine or among years of sardine run may also contribute to the lowered conformity of sardines to weight at length ratios.

The greatest difference of fish morphology, according to the PCA, was the size of the body and head (from snout to back of operculum) between the tested populations. The results of the DFA confirmed those measurements as significant to the variability in the fish morphology. The Western Cape samples grouped separately from those caught in the other regions due mostly to larger head size and the depth of body. The reclassification of most (91.5 %) individuals to the sites where they were collected indicated the statistical significance of the classification. Relative size provided some discrimination between populations; however population overlap was still significant. The geometric shape analysis supported the multivariate results with the Cape Town samples showing an increase in the size of the head, and the KZN, Mossel Bay and Port Elizabeth samples showing decreases in head size.

Silva (2003) and Wessels (2009) concluded that differences in head size and shape were discriminating factors in the delineation of stock structure in European *Sardinian pilchardus* and South African *Sardinops sagax*. In the latter case, Wessels (2009) found a similar reduction in head size in the south east samples, the greatest differences however were found between Western Cape samples and sardines sampled in Namibia. Cluster analysis supports these findings with all sample regions separating out with high confidence. Freon*, et al.* (2010) hypothesis (test 3) that sardines from the run are phenotypically distinct from the rest of the population was thus accepted, as multivariate and geometric morphometrics showed

differentiation among sites, including the KZN sardine run individuals. Although direct cause-effect relationships between meristic characters should not be drawn (van der Lingen, Hendricks*, et al.* 2010), association between vertebrae and length-body mass relationships between sites was observed.

It must be noted that it would be beneficial to be able to include sardines from further north, such as from Namibia (as in the study by Wessels (2009)) to gain better resolution and for relative data comparisons. The weak sardine run of the past years has not allowed for an inter-shoal comparison of sardine morphology (KZN experienced poor sardine run events in 2012 and 2013). Greater resolution would increase the strength of the conclusions of tests employed in this study. The small sample size of the KZN population relative to the other populations is also a factor to account for in comparing the results (Cadrin 2000) even though few sardines lay outside of the 95 % confidence interval.

The observed morphological differences between populations are most probably influenced heavily by environmental factors including temperature, food type and availability and others (Cadrin 2000). The heterogeneous environment that the sardines occupy in South Africa is probably the reason for the observed morphological structuring (Heydorn*, et al.* 1978) which seemed to follow the results of Wessels (2009) and van der Lingen, Hendricks*, et al.* (2010). The addition of genetic analyses can test connectivity and structure of these regions (Palumbi 1994, Cadrin 2000, Hellberg*, et al.* 2002). Morphological parameters can then be compared with the genetic structure of populations to determine whether phenotype is an appropriate estimator of population affinity. The genetic structuring in the South African sardine is discussed in the next chapter.

Chapter four:

Genetic population structure of the KZN sardine (*Sardinops sagax*) using mitochondrial and exon-primed, intron-crossing (EPIC) markers

4.1. Introduction

Assessment of genetic structure is important to fishery scientists for the delineation of fished resources at a genus, species, and population resolution and aids in developing rational conservation strategies (Palumbi 1994, Agüero and Rodríguez 2004, Palsbøll*, et al.* 2007). Population structuring of small pelagic teleosts such as *S. sagax* is influenced by an interaction of a multitude of variables which affect natural selection, genetic drift, mutation and gene flow (Palumbi 1994, Hellberg*, et al.* 2002). These effects can be measured by estimates of allele frequencies of genes among individuals, thus determining population structure (Palumbi 1994, Hellberg*, et al.* 2002, Hansen 2003). Contrary to this, morphological attributes are under strong selective pressure and may not reflect an individual's phylogeny (Palumbi 1994, Hellberg*, et al.* 2002, Zinetti*, et al.* 2013). Thus, the use of genetic analyses is important in determining population structure of fish stocks (Excoffier*, et al.* 1992, Carvalho and Hauser 1995, Punt and Hilborn 1997, Begg*, et al.* 1999, Palsbøll*, et al.* 2007, Hauser and Carvalho 2008).

Pelagic teleosts such as *S. sagax* have large range sizes, short generation turnover, larval and adult dispersal ability, large population size, and undergo mass breeding and spawning; these life history traits and others such as balancing selective pressures keep populations homogenous (Grant*, et al.* 1985, Palumbi 1992; 1994, Hellberg*, et al.* 2002, Gonzalez and Zardoya 2007). Conversely, factors contributing to genetic structuring include sexual and other selection pressures, larval retention, spawning migrations, local adaptations and geographic and biological barriers (Palumbi 1994, Martínez*, et al.* 2006, Gonzalez and Zardoya 2007). Historical factors such as previous climate and sea level fluctuation, such as that of the last glacial maximum 18000 years ago, may also influence genetic structure (Palumbi 1994, Martínez*, et al.* 2006).

In Southern Africa, *S. sagax* occurs across two heterogeneous ocean current systems and numerous biogeographic regions (Teske, et al. 2011). These regions grade from tropical, north of St. Lucia, and Sub Tropical between St. Lucia and the Wild Coast along the north Eastern Cape coast (Teske*, et al.* 2011). The biota is considered typical of warm temperate climates from Algoa Bay to Cape Agulhas, and cool temperate from Cape Point northward up the West coast of South Africa (Teske*, et al.* 2011). The sardine run, a seasonal migration of *S. sagax* up the east coast of South Africa, is a feature which has also been hypothesised to occur due to some form of genetic structure. Freon*, et al.* (2010) determined that there were three hypotheses relating to the sardine run that are amenable to genetic testing. These hypotheses include subpopulation spawning migration (Baird 1971), natal homing and imprinting (Carr 1967, Stabell 1984), and relic behaviour (Wyatt*, et al.* 1991). These hypotheses and the testing of larger scale population structure of the Southern African sardines are complementary, and can be tested simultaneously.

Nevertheless, studies have thus far investigated the population structure of *S. sagax* across these regions based on body morphology and meristic characters (van der Lingen*, et al.* 2009, Wessels*, et al.* 2010). As mentioned in chapter three, an investigation of this nature would have important implications for the seasonal KZN beach-seine fishery and its management, for the development of a multi-stock model for the South African sardine, and answer key questions relating to the nature of the sardine run (Freon*, et al.* 2010). Such a study would also corroborate or contradict the use of morphology in determining population structure in *S. sagax*.

Incorporation of both mtDNA and nuclear DNA data to population studies allows a thorough and robust investigation of genomic variation between individuals (Carvalho and Hauser 1995, Hauser and Ward 1998, Ward 2000, Ward*, et al.* 2009). Broad-scale markers, such as mitochondrially-encoded cytochrome oxidase 1 (mtCOI), have been used successfully on clupeids to elucidate phylogenetic relationships within and between species (Ward*, et al.* 2005, Ivanova*, et al.* 2007) and can be used to validate morphological species identification. Recently, universal primer cocktails designed to amplify a 650 base pair region of the mtCOI region, have been developed at the Canadian centre for DNA barcoding and have been

successfully used to barcode marine fishes. The barcoding of sardines has already been conducted on sardine larvae and adults in KZN and abroad (Ward*, et al.* 2005, Ward*, et al.* 2009, Zemlak*, et al.* 2009, Steinke*, et al.* 2011, Cawthorn*, et al.* 2012, Keskİn and Atar 2013).

The sardine run may show shallow but significant structuring due to large effective populations and balancing selection. This necessitates the use of a spectrum of highly, to moderately variable markers. Exon-primed, intron-crossing (EPIC) primers have been described by Touriya*, et al.* (2003) and used on the clupeid *Sardina pilchardus*, and other non-model teleost fish (Hassan*, et al.* 2002, Li*, et al.* 2010). These primers have the advantage of being applied to a wide array of organisms, and it is possible to analyse the exon or intron component of a sequence separately to determine homology (Li*, et al.* 2010). The exon and intron regions are often under different selective pressures, thus it is possible to analyse genetic material at separate scales of phylogeny (Li*, et al.* 2010). The use of genetic analyses, specifically EPIC markers and mtCOI data are thus appropriate for testing population structure and in testing the hypotheses of Freon*, et al.* (2010) about the sardine run.

The aim of this study is to search for both fine scale and broad scale structure in populations of the South African *Sardinops sagax*. A mediated interaction between life strategies and environmental factors in *S. sagax* is predicted to result in some form of structuring, similar to that found in other small pelagic teleosts. Low but significant genetic differentiation is expected between the Cape and KZN samples and individuals of the sardine run are expected to show little seasonal temporal differentiation over separate years if they are to be recognised as a separate subpopulation. To test these hypotheses, samples from Cape Town and KZN, including juveniles spawned in KZN were analysed using mtCOI data and nine EPIC markers, and investigated using standard population and phylogenetic analyses.

4.2. Materials and methods

Refer to chapter 2, 2.4. Genetic material, primer and data collection, and analysis.

4.3. Results

4.3.1. Population genetic analyses, using EPIC primers

Ten EPIC markers were tested for the presence of polymorphisms; opsin (Ops), actin (Act), chymo-trypsin B (Chym-B), myosin light chain (MLc), Calmodulin (Cam-3), glyceraldehyde-3-phosphate dehydrogenase (GPd), Aldolase B 4(Aldo-B 4), gonadotropin-releasing hormone (GnRH), alpha-tropomyosin (Tr-1) and Aldolase-B (aldo-B). Only the gonadotropin-releasing hormone 3 (GnRH) marker failed to amplify with *S. sagax* genomic DNA. The rest were amplifiable across all populations and possessed detectable polymorphisms. The representative samples sequenced ranged from 135bp - 1022bp.

Genetic diversity indices for each marker studied were number of individuals = 30.750 – 42.75, number of alleles = 2.750 – 6.500, effective number of alleles = 1.653 – 4.515, Shannon's information index = 0.658 – 1.637, expected heterozygosity 0.389 – 0.776, and observed heterozygosity 0.241 – 0.911 (Table 10). Amongst the markers, MLc showed the most genetic diversity with the highest number of alleles, effective number of alleles, Shannon's diversity index and effective heterozygotes. The largest number of individuals and observed heterozygotes was observed for the gen e Actin. The least genetic diversity with the lowest number of alleles, effective alleles, Shannon's diversity index, number of observed and number of effective heterozygotes was observed for the gene Aldo B-1. Tr-1 had the lowest number of genotyped individuals (Table 10).

Genetic diversity indices among all the markers, according to site were number of individuals = 31.111 – 42.556, number of alleles = 4.000 – 4.778, effective number of alleles = 2.686 – 3.178, Shannon's information index = 1.084 – 1.224, expected heterozygosity 0.598 – 0.657, and observed heterozygosity 0.513 – 0.637 (Table 10). Genetic diversity per site was highest in the KZN 2012 population, with the highest number of individuals, number of effective alleles, Shannon's diversity index and number of effective heterozygotes and observed heterozygotes. Sardines from Cape Town had the largest number of alleles. KZN larvae had the lowest number of individuals, number of alleles, number of effective alleles, Shannon's diversity index, and number of effective heterozygotes. Sardines from Cape Town had the lowest number of observed heterozygotes (Table 10).

Table 9: Reference sequences for each of the nine EPIC markers used for *Sardinops sagax* at four sites along the South African coastline. A random individual from the KZN 2012 population was used.

Name	Locus	Size (bp)	Reference sequence (5' - 3')
Opsin	Ops	1022	GAGTTCATTCTCACTCCATGCTGTAAGCACGCCCATTTCATGACTAGTGTTCCTTCAGGTTACGTTGCGGGGGGGAATTTGCAGTCTGATCGAAGGTCCTGGGTTGCAGATGCAAAGGCTCCTTATTGGAAAGGATTATGAACCCCGGTC TTTTCTGGGAAATAAGTCCTGACAGGCCGGACTGCACCCCGACGCCTCATTTTTTGGTTTTCCCCTATTTTTGGGCATGATGCTTATTTTTGTTTTTCCTATTGGGAACTTTTTCCGGTAAATGGACCCTTCTACTGGCAATGTGTTGAGCT GGATGGTAGCCTTTTTTGTTTCATACTGTTTTTTGGTTGGGGTTTCTTACAGTCGATTTTGCTATACGCCTATCAATGCGCCAAAGATGGTGACCCATCACACAGACAATTTGTTTAGGAATGGATGCCACTGGTTCTCTATAAACATAATA TTTGTTATGTAAGCTATAGGCTACATGCTTTTTGTTATGATACAAAAAAAGACAAATGCTTCTGAAATAGACTACCCTTTTTATCATTTTTTTTCAGAAAAAAAAACCCCACACCCGCTATGAGCTGCCTGACAATTTGCATTTTTTCCAATGG ACATTTTTTTCCAAAACACCTATATCCACCCATATAATTGTATTGGCTGTAAATAACCAGCGACAACAATCCCAACCAAATGTCCCAGTGGAATTCAAACAATTTTCCCACATCCTGCTCTGTAATTATAGCCCCGCTAGAGAAAACTTCTAC TCCTGGATCAACACATCAAAAAAAAGCACCCTGTCCTCCAAATCTCTGCACAGCCTTGGTTAACGTTTCCTGAAATAGACTCTGGCCCGGAAAAACAAAACAGAAAATCGACCCCAAGTCCCCAAGAAAGCCAAACATCCAGAACTGTCA CCGGACCCGGGACTGTCAGCGTACTGACCCTTCATTTTACACAGGGTCAGCCGGGGATTCCTCTATTCCTGTGCCATAGGCAGCCTAATGCCCACTGTCCCTTGGGATCAG
Actin	Act	382	GGGCACCTTTGATCCTACTCTCCGTGACTCCCACCGTCACCCCTGGCTGGGGGGATCCCCCAGGCCAGATTTAAACAAGCTCCCCCAGCTCCACCAAGGCCCCACTTCCCTCTCTCTTTCTCTCCCTCCCCTGTCTCTCTCTCTGAATCTCTCT CTTTCTCCCTCTCCCTCTCTCTGTCTTTCTCCCTCGCCCTCTCCCCCAGTCTCCCTCTTTCTCTCTCTCCTTCCTCTCTCTCTTGTGCTCTCTGGCCTGCGTGTTTGTGCAGAGACGGCTCGCTCTCGCACGTCCGAGGTTGGCTCTGGTGCTCTG GGGACAGCTGGTAAGAGCCAACCTGGTAAGAGGCACGGAGAGAGGGGCCCCTGTGTTGCAGTCAGCCTCGGA
chymo-trypsin B	ChymB	382	GGGCACCTTTGATCCTACTCTCCGTGACTCCCACCGTCACCCCTGGCTGGGGGGATCCCCCAGGCCAGATTTAAACAAGCTCCCCCAGCTCCACCAAGGCCCCACTTCCCTCTCTCTTTCTCTCCCTCCCCTGTCTCTCTCTCTGAATCTCTCT CTTTCTCCCTCTCCCTCTCTCTGTCTTTCTCCCTCGCCCTCTCCCCCAGTCTCCCTCTTTCTCTCTCTCCTTCCTCTCTCTCTTGTGCTCTCTGGCCTGCGTGTTTGTGCAGAGACGGCTCGCTCTCGCACGTCCGAGGTTGGCTCTGGTGCTCTG GGGACAGCTGGTAAGAGCCAACCTGGTAAGAGGCACGGAGAGAGGGGCCCCTGTGTTGCAGTCAGCCTCGGA
myosin light chain	MLc	674	GTCGTCCATCTGGGGGGAAGCGGCCCACAGGTTGGTCATCTGAGGAGGAGGAGGGGGAGGAAGAGGAATAGCAGTTAGATAGATAGATATATACTTTATTGTCCCCATAGGGAAATTTGTCTTGGACTTCTCAAAGTATATGACCCG GTACAAAAACATATAAAACAACCACCTCAGATACAACATGACATCCATATCCTCATGATAAATCCAAAATACATTTGACAGAAATATTAATTACAGCAATCTGTGTATGATGCTTCACATGTACATAAAACAAATATTAGTGCTGTTTTACT TTGTTGATTAGGACAAAGCTGGCTTGTCACACAACATATCATACCCTATTCATCTGAAATACACTACTTTGTTGGCTCTGAAGCCTACATCATCATTATCGTGTATGTGAAGTGGTTTTGGATTTTTAAAAAGTGCAGACATTTCCTGTGTA TAACCCACAGGACAGTGTTTTATGCAAGTTAAAAGAAACAAACCCATATTAATACCCTATTAACTGCCTCATAGCTGAAAAAATTTTAAGTCAATGTATACGCCATGGCTAATAGTTAGGAAATTACGGTATTTCTATGTCTTACAATCAT GAACATTATCATCAACGTAATCATGAACTTAATTCATGAACACTCACCTCCTCGAGAGGAAACCCCTGTCGA
Calmodulin	Cam-3	491	CATAGAGTAGGATGTGTATGAGAGAGAAGCTACAGCATGTGTGTGTGTATGTGTGTGTCGGCGTGTGTGTCATGCAAATGATCTTTCACCACATTTTGAATGTCTTGGATAGATCTACTCTGCCAAAGACGCTGAACGCTAACAAAATG ATGGCCCAACCACTGAAGGTCAAGAATGATGAGGGACTGCACATTCCCTTTCATAGACTGCTCCCATTTACTTAACCAGAGCAATCTAAAGACAGCCTGACTCATAATCTTATATGCACAAATGTGCCAACTCTCAGAGATAGGCCTATA ATCACATTCTAACATGGCTAACACCATGCAATCACATCAGTAGGAAGGCCATTATCTTGGCTATCCAAGTGATGGGAAAAAAAAAACCTTATCTGGTCTTGTCAATGTCCATGACCTCTCACCACATAGATATGCCCACACACACCCACACC AAAGTCTCACCTTGTCAAACACTCTGAAGGCCTCACGGAGC
glyceraldehyde-3-phosphate dehydrogenase	GPd	303	TCCGGCACTGTCTCTACTCTTTAGTGAATATTAAGAGATGACAACTGATTTTAAACACAAAATGATCAATTATGAACAAGTAACATTTGTGTGTTGTTGATAGAAAGACAAAACACTTTGATAATTTCATACTGTTAATTATTAACTGTAAA CATCTTCTGGCTGACAGGCTTACATGTTCAAGTATGACTCCACCCACGGCCGCTTCAAGGGTGAGGTCAAACCCCGCTCCTCGGGGGAGATGTCAAAGCCTCATACATCTACCTCCTGCCCGGGCGCTTCAAGGTAGAGGGAAAAGTCA AA
Aldolase B 4	AldoB-4	135	TGTTACTCGATCACTTTTACTCCGTTACTCAGTAGTTTTATGTGATGCATGGTGCAAGTTGGTGACTGTTTCTTGCTCCTTTCAGAATGGTCTCGTGCCCATTGTGGAGCCAGAGATCCTGCCTGATGGAACCCA
alpha-tropomyosin	TR-1	848	GCAATGCTCTTGCCATGGTTTAGGAGCAACAGTCGTGCCTCTTGAAGGGATCAGCGGCATAGTCTCAGGGTGTTTACAATAAAGTGTTTGACACACACACAACTTACTTGTTAGGGTCTGATTCTTTATGGTTATGGTTATGGTTATTGAA TTTGGCAGACGCTTTTGTCCAAAGTTTAAAAGTGTATGTATACTCCTCATACACAAACTCTACTAAATAAGTGTTATTCTTATAAAGGGATGTCTGTGCTGTCAACACAAAAAAAAAGAATGAGACCAGCATCGTTATGATTTAGATGACT TCCTCGTCAAGTTATTTTGTTGTTGTTATGATGTCATATAATGATGATAAAAAGCAGGTCACGATAGCAACCATATTCACCCCACATTTTTTATGGTTCAATTATCTGGTCTTGTTTCATGGTGCTGTGCAACGAAGAGAACATCACCCGGT GTTGTGAAAAAGATATTAAAAAAGTAGTATCCGGCTCGGAGGTGCCTCTTTTATAAAAACCACCCTGTGTAGCAGGAGGGGTGTGACTCTGCTGATGTAGCACCACTCTCTATTCAGAGTGGAAGGGAGATCTCACTTTGCTCCGATTCT GGATCTTGGGGTGCTCTCCTCGCATAGAACACACCTAACGCAGAGAAACAAAATATCTTAAAATCTTCTTATCTCTTCTTTTTGTTATAAAATTTTGTTGTGGCGTGTGTATGCCCAATACCGAAGATAAAATAATCGAACACAGTATGAT GTGGAGCGAGGTACGATACTTCACCCCCCTCCCCCCCTTTTTTTTTTAAAAAAAATAATTTCTCCACACCGCTCTACAAACTTTTTTTTGTCC
Aldolase-B	AldoB-1	262	GCTACGGCATCAACAGCTATTGTGCCTCTTGGGTGTTGGCCATGACCTGATTGCATTTGTGTTGTCCTTGCTTATTTGTGTTGAAATGCGTCGGTCAACTTTTTTATTTATTATTTTTATTTCCTAGGTACCATGGGCAACCGTCTTCAGAAG ATCAACGTGGAGAACACCGAGGAGAACCGCCGCTACTTCCGTGACCTCCTCTTCTCCGTTGACCAATCCATCAACCAGTGTGTGGGCGGGATCATCTTCTTCCACGAGA

Table 10: Genetic diversity indices for *Sardinops sagax* at four sites along the South African coastline, based on frequencies of nine EPIC markers. Standard errors are presented in parentheses. N = number of samples, N_a = number of alleles, N_e = effective number of alleles; I = Shannon's information index, H_e = expected heterozygosity, H_o = observed heterozygosity. KZN = KwaZulu-Natal, CT = Cape Town.

Locus	N	N_a	N_e	I	H_o	H_e
MLc	38.500 (3.476)	6.500 (0.289)	4.515 (0.273)	1.637 (0.057)	0.836 (0.038)	0.776 (0.015)
ChymB	32.750 (3.881)	3.750 (0.250)	2.526 (0.154)	1.057 (0.037)	0.392 (0.014)	0.599 (0.027)
aldoB-1	39.500 (2.754)	2.750 (0.250)	1.653 (0.092)	0.658 (0.047)	0.241 (0.103)	0.389 (0.036)
Ops-1	36.250 (4.589)	3.000 (0.000)	2.313 (0.068)	0.928 (0.018)	0.605 (0.106)	0.566 (0.013)
Tr-1	30.750 (5.513)	6.500 (0.646)	4.069 (0.604)	1.536 (0.125)	0.401 (0.083)	0.737 (0.038)
AldoB-4	40.750 (3.400)	4.250 (0.479)	2.796 (0.404)	1.110 (0.157)	0.596 (0.148)	0.619 (0.054)
Cam-3	41.000 (2.345)	5.500 (0.289)	3.051 (0.248)	1.275 (0.065)	0.563 (0.108)	0.666 (0.027)
Act	42.750 (2.657)	4.500 (0.500)	3.469 (0.309)	1.324 (0.108)	0.911 (0.025)	0.704 (0.029)
GPd	35.750 (1.250)	3.000 (0.000)	2.392 (0.224)	0.944 (0.079)	0.649 (0.066)	0.568 (0.049)
Population	N	N_a	N_e	I	H_o	H_e
KZN 2011	35.889 (1.822)	4.444 (0.475)	2.983 (0.239)	1.169 (0.105)	0.586 (0.086)	0.635 (0.036)
KZN 2012	42.556 (1.215)	4.444 (0.530)	3.178 (0.303)	1.224 (0.110)	0.637 (0.088)	0.657 (0.037)
KZN Larvae 2012	31.111 (2.220)	4.000 (0.441)	2.686 (0.282)	1.084 (0.094)	0.573 (0.079)	0.598 (0.038)
CT	40.667 (2.483)	4.778 (0.641)	3.058 (0.459)	1.175 (0.145)	0.513 (0.096)	0.610 (0.057)

Allele frequencies were significantly different from those expected under Hardy-Weinberg equilibrium. All populations showed a heterozygote deficiency (Table 10). The genetic markers showed a heterozygote excess, except for ChymB, aldoB-1, Aldo B-4, Cam-3, and Tr-1, which had a heterozygote deficit (Table 10). Expected heterozygosity, a measure of gene diversity, was highest for MLc, and lowest for aldoB-1, among the genetic markers (Table 10). An AMOVA indicated 89 % intra-population variance which was significant (Φ_{PT} = 0.105, p < 0.001), and 11 % inter-population variance also significant (Φ_{RT} = 0.111, p < 0.001). Genetic difference between the two regions (All KZN sardines combined, compared with Cape Town) was negligible and not significant (Φ_{PR} = -0.003, P=0.736; Table 11). Total genetic difference among all populations was moderate and significant (Φ_{PT} = 0.091, p = 0.010; Table 11 and Table 12). AMOVA revealed the two most genetically different groupings to be KZN 2012 and the KZN larvae collected in the same year (Φ_{PT} = 0.127, p = 0.001; Table 11). The most similar were KZN 2011 and KZN 2012 (Φ_{PT} = 0.073, p = 0.001; Table 11).

An AMOVA was conducted for the data derived from each EPIC marker (Table 12). Tr-1 had the lowest Φ_{RT} value which was significant, while Mlc had the highest (Table 12). Cam-3 showed

the lowest Φ_{PR} value, while AldoB-5 had the highest (Table 12). ChymB had the lowest Φ_{PT} value although not significant, and AldoB-5 had the highest (Table 12). A Mantel test revealed no measurable isolation by distance ($p < 0.001$, $r^2 = 0.1039$). Levels of migration among the four populations differed among sites and years (Table 11) with mean Nm values ranging from 20.6 – 311.1. Migration rate between the two tested regions was 2.355 (Table 11). A PCA of the EPIC data failed to resolve any observable groupings (Figure 12).

Table 11: Summary of spatial and temporal distances, Φ_{PT} values and their level of significance, and Nei genetic distance (NeiP) among four populations and two regions, for *S. sagax* at two sites and over two years along the South African coastline, based on nine EPIC markers. Probabilities for Φ_{PT} values and confidence intervals (2.5–97.5 %) for Nm are represented in parentheses. Significant Nm values are shown in bold. KZN = KwaZulu-Natal.

Population 1	Population 2	Spatial and temporal separation	Φ_{PT}	NeiP	Nm (Migrate)	
					Pop 1 to pop 2	Pop 2 to pop 1
KZN 2011	KZN 2012	0km, 1 year	0.073 (P=0.001)	0.127	111.9 (54.0-106.7)	**122.8 (106.6-192.0)**
KZN 2011	KZN larvae 2012	0km, 1 year	0.091 (P=0.001)	0.108	57.9 (712.0-1000.0)	203.8 (0-72.6)
KZN 2011	CT	1400km, 16 months	0.092 (P=0.001)	0.108	**37.7 (12.6-62.0)**	160.4 (0-38.0)
KZN 2012	KZN larvae 2012	0km, N/A	0.127 (P=0.001)	0.189	**42.6 (18.6-66.0)**	**254.4 (182.0-379.3)**
KZN 2012	CT	1400km, 4 months	0.091 (P=0.001)	0.121	311.1 (4.0-44.0)	51.9 (7.3-51.3)
KZN larvae 2012	CT	1400km, N/A	0.090 (P=0.001)	0.180	64.2 (0-38.0)	**20.6 (2.6-38.6)**
Region 1	Region 2	Geographic distance (km)	Φ_{RT}		Nm (Genalex)	
All KZN	CT	1400km, N/A	-0.003 (P=0.736)		2.355	

Table 12: $\Phi_{PT/RT}$ values and their significances for the nine EPIC markers, for *S. sagax* in KwaZulu-Natal compared to those of Cape Town along the South African coastline. Φ_{RT} = Phi region versus total, Φ_{PR} =

	EPIC markers									
	MLc	Chymb	aldob1	Tr1	Ops1	AldoB5	Cam-3	Act	GPd	Total
Φ_{RT}	0.099 (P=0.010)	-0.060 (P=1.000)	-0.028 (P=0.950)	-0.099 (P=1.000)	0.090 (P=0.010)	0.026 (P=0.020)	0.022 (P=0.100)	-0.065 (P=1.000)	0.009 (P=0.210)	-0.003 (P=0.740)
Φ_{PR}	0.055 (P=0.010)	0.069 (P=0.020)	0.056 (P=0.030)	0.159 (P=0.010)	0.041 (P=0.040)	0.174 (P=0.010)	0.035 (P=0.020)	0.135 (P=0.010)	0.097 (P=0.010)	0.094 (P=0.010)
Φ_{PT}	0.149 (P=0.010)	0.013 (P=0.120)	0.029 (P=0.070)	0.076 (P=0.010)	0.127 (P=0.010)	0.195 (P=0.010)	0.056 (P=0.010)	0.076 (P=0.010)	0.105 (P=0.010)	0.091 (P=0.010)

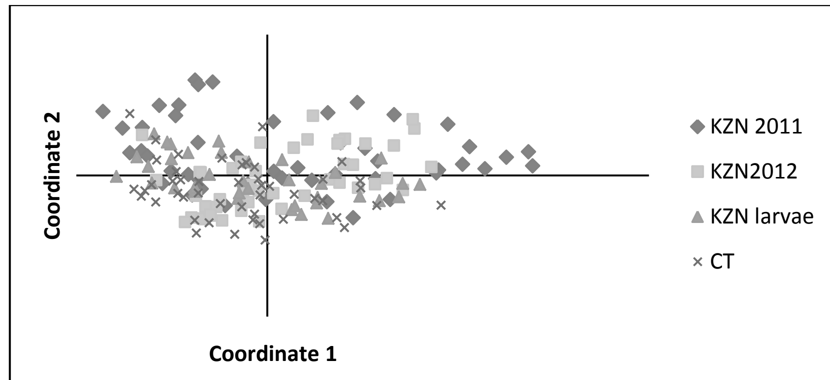


Figure 12: Principal coordinates analysis depicting the first two principal components of genetic distance among four populations, for *S. sagax* at two sites and over two years along the South African coastline, based on nine EPIC markers.

Bayesian analysis of population structure revealed the most likely number of population clusters for the sardines sampled were three (mean LnP(K) = -3214.780, Delta K = 155.644; Figure 13). KZN 2011 and the KZN larvae of 2012 were most likely to have originated from two separate clusters, mostly blue and approximately 30 % red (Figure 14). KZN 2012 adults and the Cape Town samples were composed of similar clusters (Figure 14). KZN 2012 comprised about 45 % each of the red and green and about 10 % blue population clusters (Figure 14). Cape Town comprised of approximately 40 % green population cluster, 35 % red and about 15 % of the blue population cluster.

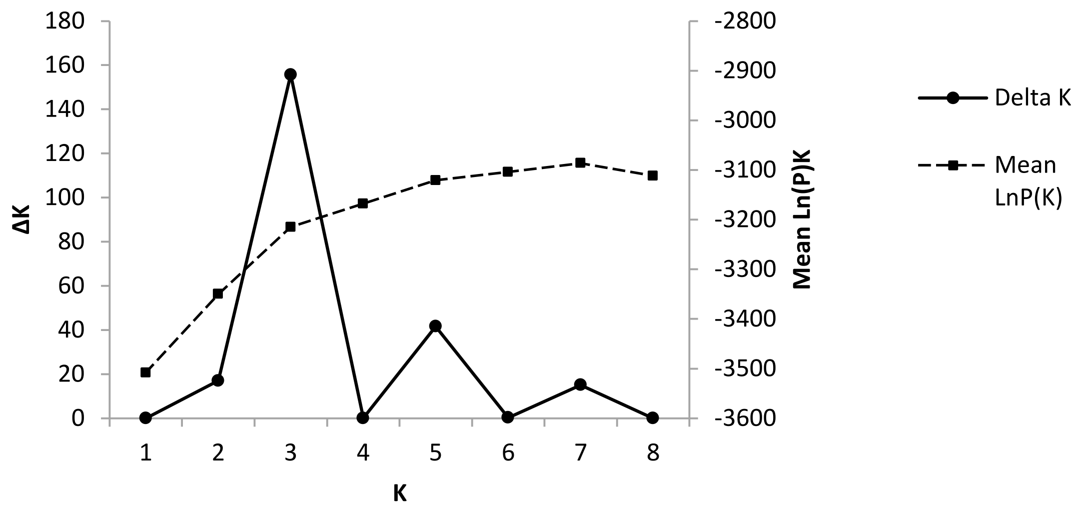


Figure 13: Delta (Δ) K (circles, solid line) and mean likelihood (squares, dashed line) versus corresponding values for the for *S. sagax* in KZN between 2011 and 2012, and Cape Town along the South African coastline, based on nine EPIC markers (as derived from Structure harvester) (Earl 2012).

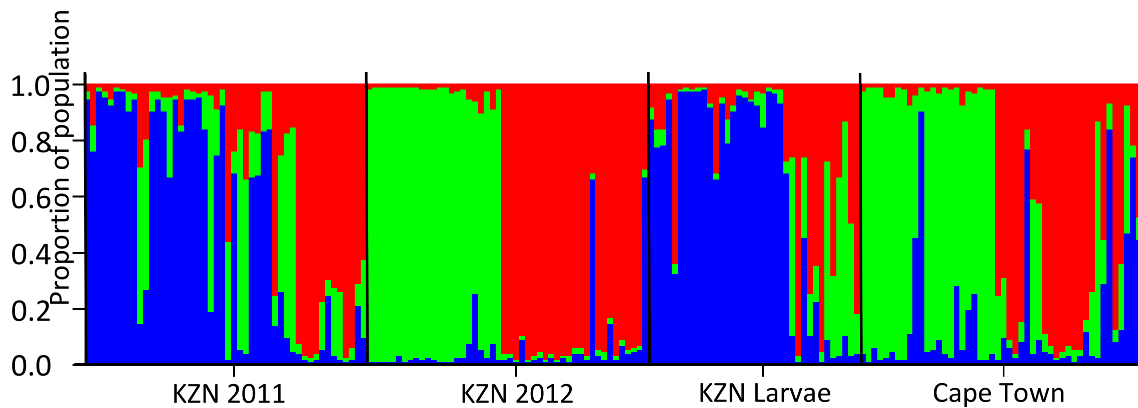


Figure 14: Bayesian analysis of population structure with three clusters (shown in red, green and blue), for *S. sagax* in KZN between 2011 and 2012, and Cape Town along the South African coastline, based on nine EPIC markers.

4.3.2. Analysis of Sardinops sagax using mtCOI genetic data

Between 665pb and 680bp of the mtCOI gene was amplified for 53 *S. sagax* individuals, collected at Cape Town, Port Elizabeth, and from the KZN sardine run for 2011, 2012. Juveniles collected in KZN in the months following the sardine run in 2012 were also included (Table 1).

A further two *S. sagax* individuals from the Agulhas Bank near Mossel Bay and two individuals sampled near Durban, were included from two previous studies, retrieved from Genbank (see Table 4 for information of sequences included from genbank (Steinke*, et al.* 2009, Steinke*, et al.* 2011)). Three sequences of another species, *S. melanostictus* were also included. Results of the reconstructed mitochondrial COI data yielded 41 haplotypes with posterior probabilities all over 0.9 (The probability that the calculated haplotypes were accurate). The trimmed, aligned sequence set was 652 bp in length with a G+C content of 49 %. There were 51 variable sites and 61 mutations. The mitochondrial COI gene generated a haplotype diversity of 0.916 (SD = 0.030) and an average number of nucleotide differences between sequences of 5.941. The most common haplotype was haplotype two which comprised 19 individuals from all the sites sampled. The next most common haplotype was haplotype 10, comprising two individuals from KZN 2012 and one of KZN 2011. Haplotype three comprised one KZN 2011 and a KZN 2012 individual. Haplotype 19 consisted of two larvae. The three *S. melanostictus* samples were assigned to two haplotypes. All other samples were allocated as individual haplotypes.

Grant and Bowen (1998) concluded that the *Sardinops* genus is monotypic. Based on this, all *Sardinops sp.* can then theoretically be analysed via population analyses. AMOVA revealed that 52 % of genetic difference between South African populations was due to inter-population variation. The Φ_{PT} values between the Cape Town and PE, Mossel Bay and Cape Town populations were lowest and were not significant (Table 13). The highest pairwise Φ_{PT} values were for KZN 2011 and Mossel Bay, and KZN 2012 and Mossel Bay, compared to all the other sites, which was significant (Table 13). The Nei's pairwise genetic distance (NeiP) were all lower than 0.03. Rates of migration were all between 1.51 and 3.22.

AMOVA revealed 74 % of genetic difference between three global regions (South Africa, Canada and Japan was due to inter-regional genetic variation. Genetic difference among the three regions tested was high, with all Φ_{RT} values above 0.750 (Table 14). Genetic distance (NeiP) ranged 0.015 – 0.021. A Mantel test revealed a positive correlation of isolation by distance ($p < 0.001$, $r^2 = 0.6383$). A PCA resolved strong groupings, with all individuals representing the three regions, separating (Figure 15).

Table 13: Summary of spatial and temporal distances separating populations, Φ_{PT} values and their level of significance in parentheses, NeiP and Nm, among six South African populations based on the mtCOI marker. KZN = KwaZulu-Natal.

Population 1	Population 2	Spatial and temporal separation	Φ_{PT}	NeiP	Nm (GenAlex)
KZN 2011	KZN 2012	0km, 1 year	0.124 (P=0.001)	0.001	3.516
KZN 2011	KZN larvae	0km, 1 year	0.195 (P=0.001)	0.001	2.055
KZN 2011	Port Elizabeth	540km, 16 months	0.196 (P=0.001)	0.001	2.048
KZN 2011	Mossel Bay	890km, 16 months	0.345 (P=0.001)	0.002	0.950
KZN 2011	CapeTown	1400km, 16 months	0.165 (P=0.001)	0.002	2.519
KZN 2012	KZN larvae	0km, N/A	0.238 (P=0.001)	0.002	1.599
KZN 2012	Port Elizabeth	540km, 16 months	0.216 (P=0.001)	0.002	1.804
KZN 2012	Mossel Bay	890km, 16 months	0.340 (P=0.001)	0.003	0.970
KZN 2012	CapeTown	1400km, 16 months	0.202 (P=0.001)	0.003	1.974
KZN larvae	Port Elizabeth	540km, 16 months	0.054 (P=0.105)	0.002	8.632
KZN larvae	Mossel Bay	890km, 16 months	0.055 (P=0.281)	0.001	8.537
KZN larvae	CapeTown	1400km, 16 months	0.011 (P=0.253)	0.001	44.757
Port Elizabeth	Mossel Bay	350km, 0 months	0.010 (P=0.430)	0.001	48.643
Port Elizabeth	CapeTown	970km, 2 months	0.000 (P=0.364)	0.002	-
Mossel Bay	CapeTown	620km, 2 months	0.000 (P=0.001)	0.001	-
Region 1	Region 2	Geographic distance (km)	Φ_{RT}		Nm
All KZN	CT	1400km, N/A	-0.003 (P=0.736)		2.355

Table 14: Summary of geographic distances between three regions (RSA, Canada and Japan), Φ_{RT} values and their level of significance, and Nei genetic distance (NeiP) among three regions, based on the mtCOI marker. Probabilities for Φ_{RT} values are represented in parentheses. RSA = South Africa.

Region 1	Region 2	Spatial distance	Φ_{RT}	NeiP	Nm (GenAlex)
RSA (*S. sagax*)	Canada (*S. sagax*)	17300 km	0.752 (P=0.001)	0.015	0.083
RSA (*S. sagax*)	Japan (*S. melanostictus*)	14500 km	0.750 (P=0.001)	0.018	0.083
Canada (*S. sagax*)	Japan (*S. melanostictus*)	6700 km	0.881 (P=0.014)	0.021	0.034

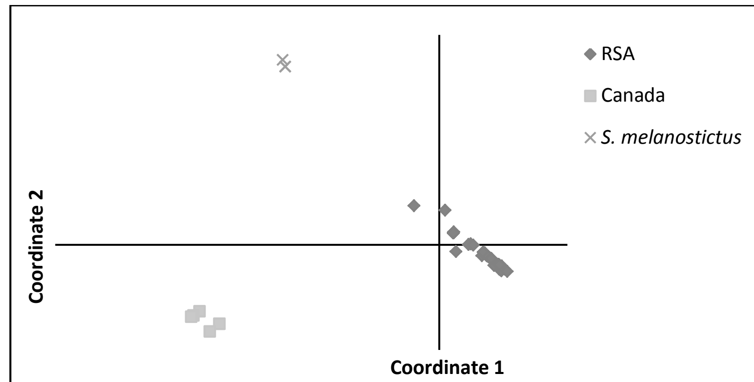


Figure 15: Principal coordinates analysis depicting the first two principal components genetic distance among three regions, for *S. sagax* and the outgroup *S. melanostictus*, based on the mtCOI marker.

The 50 % majority-rule consensus Bayesian phylogram, based on the mtCOI data, illustrated the phylogeny among 38 *S. sagax* haplotypes and, two *S. melanostictus* haplotypes, rooted with *Pomacentrus baenschi* and *Sardina pilchardus* individuals (Figure 16 a). The Japanese *S. melanostictus* and Canadian *S. sagax* formed strongly supported monophyletic clades, correlating with accepted species boundaries (Figure 16 a). Clade C of South African *S. sagax* demonstrated moderate nodal support, (Figure 16 a). Clade B representing hap 23, 24, 29 and 31 showed poor nodal support, but comprised two Cape Town and Port Elizabeth samples each (Figure 16 a) which showed moderate support. Clade C representing hap 7, 10, 11, 14 and 15 comprised seven KZN individuals from 2011, 2012 and juveniles from 2012 (Figure 16). The polytomy of 44 *S. sagax* individuals (and 23 haplotypes), of Clade A (which also nests clade B), included individuals of both the putative western and southern stock sardines. Nevertheless, the grouping of clade A exhibited poor nodal support (Figure 16 a). A minimum spanning haplotype network of the mtCOI data suggested three homogenous clades, namely South African *S. sagax*, Canadian *S. sagax*, and *S. melanostictus* (Figure 16 b). These clades represented the species level boundaries observed in Figure 16 a, with no of reticulation based on previous studies (Parish, et al. 1988, Grant and Bowen, 1998). A larger genetic distance, according to the Nei P and Φ_{PT} values (Table 13) and the PCA (Figure 15), was observed between *S. melanostictus* and *S. sagax* from Canada (Fig. 15, 16b), than those of Canada and Japan compared to the South African *S. sagax* clade (Fig. 16 a, b).

Divergence times between populations were calculated between the three regions based on mismatch distributions which were RSA: 230061 years before present, Canada: 180904 years before present, and *Sardinops melanostictus*: 80751 years before present.

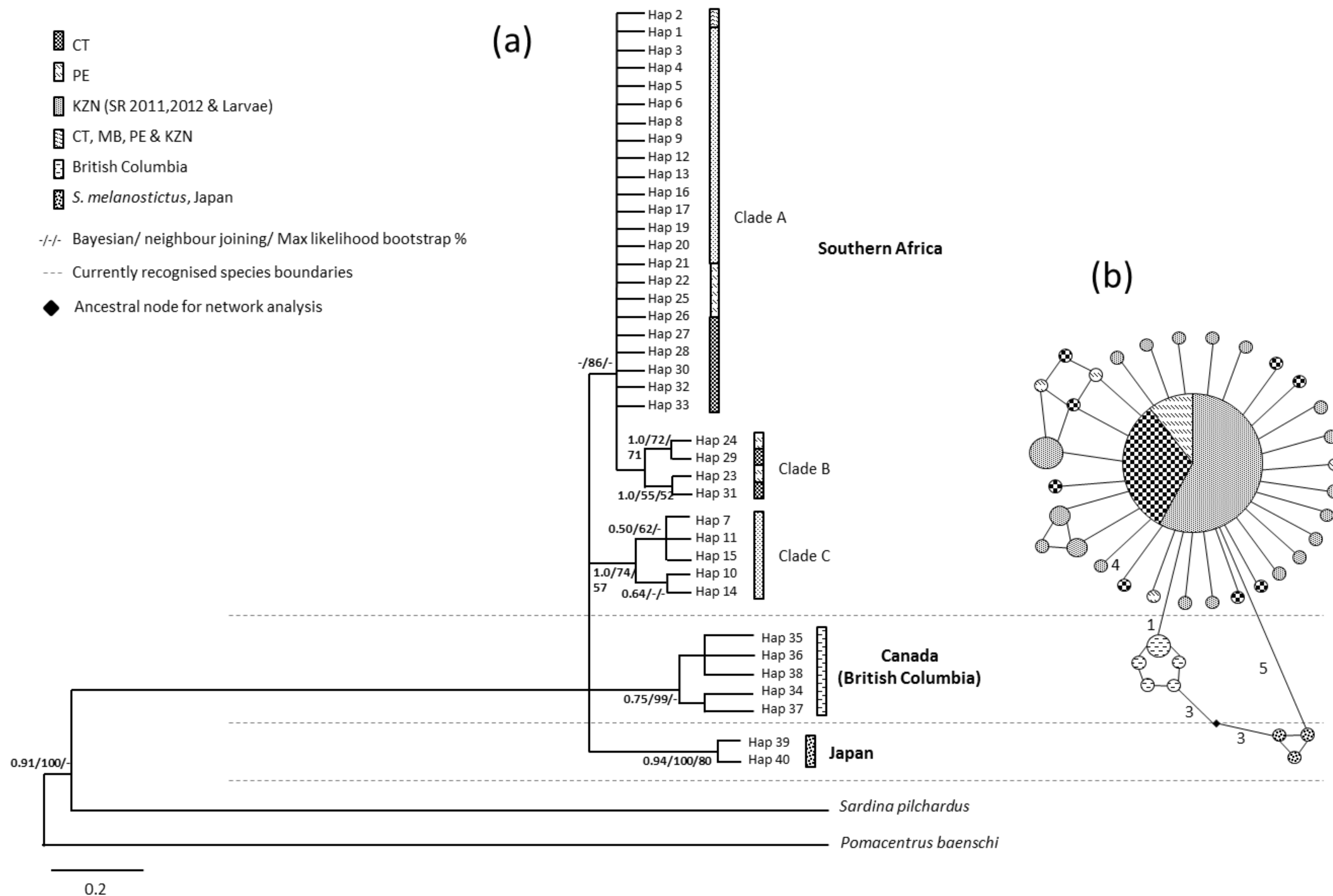


Figure 16: (a) Rooted (*Pomacentrus baenschi* and *Sardina pilchardus*) phylogram indicating evolutionary species level relationships and distribution for the mitochondrially encoded COI sequence data for *Sardinops sagax* in Southern Africa, with additional sequences for S. sagax from British Columbia and *Sardinops melanostictus* sequences from Japan; (b) a minimum spanning network of the COI sequence data using the median joining method, showing reticulation and haplotype distribution and number of sequences per haplotype (relative size of circles). Numbers represent mutational steps on branches.

4.4. Discussion

4.4.1 Population structure of KZN Sardinops sagax using EPIC genetic data

The nine EPIC markers used showed similar allele diversities to those observed in the original publications (Hassan*, et al.* 2002, Touriya*, et al.* 2003), and showed polymorphism and gene diversity between populations (H_e, Table 10). Alleles were found within similar size ranges to indices found in the studies of Hassan, *et al.* (2002) and Touriya, *et al.* (2003). The increased diversity of these markers allow for increased power of inference of the marker set (Hedgecock*, et al.* 1989). The relative ease of PCR amplification illustrated the universality of these primers for use among non-model teleosts, and confirmed that these primers are suitable for fine scale population genetic analyses.

The deficit of heterozygosity observed among five markers and all populations is indicative of mild inbreeding among populations (Palumbi 1994, Hellberg*, et al.* 2002). The higher levels of intra-population genetic difference over inter-population difference, and overall moderately low Φ_{PT} values indicate low to moderate levels of population genetic structure are also an indication of inbreeding (Weir and Cockerham 1984). Most clupeid species occupy large ranges and have large effective population sizes (Parrish*, et al.* 1989) and generally exhibit a heterozygote excess (Palumbi 1994, Hellberg*, et al.* 2002). However, in the case of the South African sardine, a heterozygote deficit exists. This observation is similar to the case of the European sardine (*Sardina pilchardus*) (Gonzalez, 2007) and the same species from Morocco (Atarhouch, *et al.* 2007). In both these cases, complex ocean hydrodynamics are thought to be responsible for varying degrees of population isolation and thus lowered heterozygosity within populations.

The heterogeneous marine environment along the South African coastline (Hutchings, 2009) coincides with three biogeographic regions, namely the sub-tropical, warm temperate, and cool temperate according to Teske*, et al.* (2011). These environmental factors may play a role in causing the observed genetic isolation. Possible genetic isolation due to the sardine run may also be a reason for the observed heterozygote deficit.

Sardine run individuals between 2011 and 2012 showed the least genetic differentiation, and almost equally high migration rates between populations. They share shallow genetic structuring and the same lineages of sardines undertaking the sardine run appear to mate and spawn in KZN annually. The rates of migration and genetic differentiation supported test four; that sardines of the sardine run are genetically distinct from the rest of the population.

Sardines of the sardine run and larvae caught in KZN in the months following the run showed the most genetic difference (considered moderate). In contrast, KZN larvae showed the greatest genetic similarity to the Western Cape population. Although the Phi_{PT} results show low structure between Western Cape and larval samples, there is relatively low recruitment of the KZN larvae to the adult Western Cape population (Figure 14).

Disparity in the relationships between genetically different populations may be due in part to the timing of the sampling of the populations in this study. If genetically distinct shoals occur during the run, then this could explain the moderate genetic difference between adults and larvae of the 2012 sardine run. Overall, genetic differences in the population's tests showed little structure (Figure 12).

Based on the moderate to high rates of immigration and emigration among the KZN Larvae and those from Cape Town, sardines spawned in KZN are likely recruited to the adult breeding populations in the Western Cape and Agulhas Bank for the COI gene (Test 7). Test eight could not be directly answered as the study had only analysed larvae from a single year class (2012). Nevertheless, the fact that Cape Town samples and KZN larvae showed low levels of structure and moderate levels of migration between themselves, suggests sardine individuals do not only spawn in KZN, but breed and spawn on the Agulhas bank and on the western coast too. This is supported by the ability of sardine to spawn year round (Akkers and Melo 1996). It is likely that the breeding and spawning of sardines multiple times a year would have the effect of lowering any structuring that a separate spawning event such as the sardine run would have on genetic isolation.

It seems then, that *Sardinops sagax* in South Africa exhibits genetic structuring not unlike other clupeids globally (Whitehead and Wongratana 1986, Parrish*, et al.* 1989, Tinti*, et al.* 2002, Croft*, et al.* 2003, Silva 2003, Gonzalez and Zardoya 2007, Baibai*, et al.* 2012). The slightly higher but significant gene diversity (H_e) of fishes in KZN, compared to Cape Town, may be suggestive of the heavy fishing pressures on the western coast sardines, the target of the South African industrial pilchard fishery. Selective pressure on fished stocks has a biological effect of not only reducing genetic diversity, but also includes the selective targeting of certain phenotypic characters (Law 2000, Coetzee, Merkle*, et al.* 2008). Fishing induced phenotypic selection, such as a minimum size of individuals in a targeted fishery, often results in a selection pressure for smaller fish at sexual maturity, as reproduction at a smaller size is favoured (Law 2000; 2007). Nevertheless, factors such as decadal-scale variance in average size of fish has been linked to density dependence (Coetzee, Merkle*, et al.* 2008).

Although the census population size and total population size (see biomass estimates for Sardine, by Coetzee*, et al.* (2010)), of the sardines undertaking the sardine run is much smaller than those of the Western Cape, they exhibit slightly higher but significant gene diversity, and in the case of the KZN 2012 sardines, have a larger effective population size than the sardines of the Western stock. Wright (1938) defined effective population size as "the number of breeding individuals in an idealised population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration". The smaller effective population size can be indicative of heavily fished stocks (Carvalho and Hauser 1995, Turner*, et al.* 2002, Hauser and Carvalho 2008). The low levels of structure and moderate levels of migration between larvae and the Cape populations (illustrated by the Cape Town, Mossel Bay, and Port Elizabeth mtCOI data) is suggestive of interbreeding of populations, possibly along the whole South African coastline. Nevertheless, the lower expected heterozygosity shown by KZN larvae over KZN adults and Cape samples may not be indicative of this interbreeding, where recombination and the sharing of alleles between populations is expected (Durand*, et al.* 2009). The slightly lower heterozygosity indices among the Cape Town population compared to the KZN population may be an example of the Wahlund effect, (where decreases in heterozygotes often represent a structured sub-population). Evolutionarily, higher migration rates would be beneficial in that it would serve to increase the transfer of genetic material between the Western Cape and

Agulhas Bank population (Nielsen*, et al.* 2003, Durand*, et al.* 2009). This would have the effect of reducing genetic drift and allele loss, thus increasing genetic resilience. These observations were mirrored by the COI data, where the KZN larvae illustrated low genetic difference from samples from the Western Cape (Cape Town) and Agulhas Bank (Cape Town, Mossel Bay and Port Elizabeth).

4.4.2 Analysis of Sardinops sagax in South African waters using mtCOI genetic data

Analysis of *S. sagax* and *S. melanostictus* between South Africa, Canada and Japan has yielded distinct clades for each (Figure 16 a) using the COI gene. Indeed, sardines from these regions are reproductively isolated due probably to their separation among three ocean basins, and as illustrated by the high genetic difference ($\Phi_{RT}>0.25$, Table 14) and low rates of migration ($Nm < 1$, Table 14) and for which there is strongly supported monophyly observed on the phylogram (Figure 16 a). This suggests that *S. sagax* between Southern Africa and Canada may represent separate sub-species, based on reproductive isolation under the biological and evolutionary species concept (Claridge, et al. 1997, De queiroz 2007) and the sardines from these regions may constitute evolutionary stable units based on this reproductive isolation and genetic difference (Φ_{RT} values, Table 14) (Crandall*, et al.* 2000). The observations are indicative of fish from each region that are no longer interbreeding, and are thus reproductively isolated (Palumbi 1994). Indeed, the three regions are positioned in separate ocean basins (south east Atlantic, western north Pacific, and the eastern north Pacific) over thousands of kilometres. Thus, even larval dispersal among regions is highly improbable, as the chance of random larval transport is further minimised through the diffusion effect and physiological traits (planktonic) of larvae (Palumbi 1994, Hellberg*, et al.* 2002). The strong genetic structuring among these regions are based on shallow evolutionary divergences (based on the genetic distance on the phylogram, Figure 16 a) and calculated divergence times. These early divergences concur with broad estimates from mitochondrial DNA by Grant and Bowen (1998) for *Sardinops*, and other Clupeids such as Engraulis (Liu*, et al.* 2006) in the Indo-Pacific, and the Mediterranean (Keskiˈn and Atar 2012). These estimates of radiation coincide with the interglacial periods during the late Pleistocene epoch (approximately 2 500 000 years ago to 11 700 years ago (Gradstein*, et al.* 2004).

As S. sagax from South Africa and Canada are so genetically disparate and reproductively isolated, the species (*sagax*) proposed by Parrish*, et al.* (1989) could in some respects be considered separate species or subspecies under the biological-isolation species concept, and the genetic species concept (Claridge*, et al.* 1997). The use of mtCOI data to delineate species is extensive and accepted (Frézal and Leblois 2008). Nevertheless, the above findings and other research such as that of Grant*, et al.* (1985), Parrish*, et al.* (1989), and Grant and Bowen (1998) argued that although genetic differences were significant, the genetic distances and relatively recent radiation (approximately 250 000 years before present) observed did not warrant further taxonomic delineation.

Nevertheless, data generated from the mtCOI data for the three regional populations suggests genetic difference (NeiP, Table 14) and distance (Φ_{RT}, Table 14) among regions is approximately equal. Thus logically, *S. melanostictus* should either be considered monotypic and assigned as *S. sagax*, or the *S. sagax* between RSA and Canada should be considered separate sub-species or species. The lack of data from other populations, known as "shadow" populations since they were not sampled, such as the Peruvian (*Sardinops sagax sagax*) and Australian (*Sardinops neopilchardus*) sardines may confound these results.

Phylogenetic analysis of *S. sagax* indicates well supported structuring of two clades (Clade B and C) amongst a polytomy (Clade A) of the rest of the South African samples (Figure 16 a). These clades displayed relatively low Nei's genetic distance from the rest of the South African samples. These groups indicate genetic structuring between a fragment of the Western Cape and Agulhas Bank (PE) samples (Clade B, Figure16 a), and some KZN samples (Clade C, Figure 16 a) from a relatively homogenous South African population. The genetic distances calculated for each region was similar to those of Grant and Bowen (1998). The branch lengths indicate that these clades show shallow genetic distances. These clades occur within the spatial ranges of the greater South African population, and thus represent possible sympatric population differentiation (clade A, B and C; Figure 16 a, b). Pairwise Φ_{PT} values showed the smallest genetic differences to be between the KZN larvae and the Western Cape (Cape Town) and

Agulhas Bank samples (Mossel Bay and Port Elizabeth). This followed the EPIC data results, and is discussed in detail in chapter 4.4.1 and chapter 5.

Freon*, et al.* (2010) predicted that sardines from the Western Cape and KZN may show genetic population structuring based on their various hypotheses tested. It is likely that the observed structuring is due to recent divergence in sardine populations, where a component of Clade B and C have an intra-population breeding preference. The grouping of Cape Town and Port Elizabeth samples together in clade B was not supported. The separation of KZN individuals of clade C may suggest that KZN sardines may exhibit some form of current or historical breeding isolation from Western and Southern stock sardines (represented in clade A and B). van der Lingen*, et al.* (2009) hypothesised that the western and southern putative stocks of sardines share a gene pool, owing to the heterogeneous environment, their motility, and the fact that there are no impermeable barriers to movement. Based on these results, it is probable that there is movement of juvenile and adult migrants into each of the populations, allowing for gene flow to persist between the areas each side of Cape Agulhas (Figure 1). This appears to correlate with the mtCOI data, although this relationship may also represent remnants of ancient population structure or some other life history traits or ecological factors (Palumbi 1994, Hellberg*, et al.* 2002).

To test Baird's (1971) hypothesis (H_U1) Freon*, et al.* (2010) listed the following questions that may be answered using genetic techniques:

- (Test 4) Are sardines from the sardine run genetically distinct from the rest of the population?
- (Test 7) Is there successful recruitment arising from individuals spawning in KZN waters?
- (Test 15) Are sardine run sardines mostly spawned in KZN?

Relating to test four, KZN 2012 sardines appear to be genetically different with relatively low migration rates between the populations tested, according to mtCOI population indices (Nm rates, Table 13) and separation of clade C on the phylogram (Figure 16 a). Nevertheless, the rest of the individuals caught in KZN showed groupings with all the other populations. KZN

2011 also demonstrated less genetic difference to the other populations (Φ_{PT}, Table 13). Thus, the data shows that some of the sardines undertaking the sardine run are genetically distinct, and show structuring from the Cape Town, Mossel Bay and Port Elizabeth samples. Nevertheless, the grouping of clade A (Figure 16 a) which included all sites does not seem to correlate with the hypothesis of a sub population spawning migration. Reasons for this observation include possible entrainment of “non sardine run” fish and the mixing of fish during the summer months when KZN sardine run fish return to the Cape.

Relating to test seven, the mtCOI data suggests successful recruitment based on the genetic similarity of larvae to adults (Φ_{PT} and Nm values, Table 13), across the populations tested, based on all samples collected during the sardine run. Rates of migration among the KZN larvae and all the other sites, especially the high rates of migration according to the EPIC data, to the Cape (Nm of 311.11, Table 11) support the hypothesis that sardines spawned in KZN may be recruited to the adult breeding populations in the Western Cape and on the Agulhas Bank. These results concurred with the rates of migration calculated from the COI data (Nm of 45.757, Table 13). The hypothesis is supported by these results as it shows the estimated transfer of genetic material between the populations; which by the migration results, is significant and considerably more than amongst other populations and the KZN juveniles. Test eight was not answered as the study only analysed larvae from a single year class (2012).

Freon, et al’s (2010) natal homing hypothesis (H_p1), based on previous work by Carr (1967) and Stabell (1984), requires that sardine run sardines are mostly spawned in KZN (Test 15). The mtCOI data suggested that most sardines belong to low structured interbreeding population where most of the larvae grouped with all other populations (Clade A, Figure 16 a). Most larvae showed little genetic difference and moderate migration rates between the Cape populations. Thus, most sardines undertaking the sardine run may not have been spawned in KZN.

It must be noted that mtCOI is primarily a species genetic barcoding marker, used to delineate species level genetic differences (Ward*, et al.* 2005). Thus, although its use in population level

studies is common, it may fail to resolve some fine scale structuring, which should be taken into account. The sample sizes used are adequate for phylogenetic techniques used in this study, however, more samples would aid resolution for population level techniques, thus caution must be exercised when interpreting this data.

According to the EPIC nuclear DNA data (Freon*, et al.* (2010) test 4), sardines of the sardine run did show structuring from the other populations, however this did not accord with the KZN larvae, which illustrated less genetic differentiation relative to the Cape populations than to KZN 2012. The rates of migration of sardines spawned in KZN, to those of the Western Cape and KZN adult populations suggest there was successful recruitment of individuals spawned in KZN (Freon*, et al.* (2010) test 7). Nevertheless, migration among the Cape and KZN larvae samples was reduced, compared with KZN 2011 sardine run adults. The Bayesian analysis of population structure, confirmed the findings that KZN larvae showed a population demographic similar to the KZN 2011 sardine run. Test eight (Freon*, et al.* 2010) was not answered as the study had only analysed larvae from a single year class and only in KZN, and comparisons of larval structure from the Cape are also needed (2012). The EPIC nuclear data suggested that most sardines belong to a finely structured interbreeding population, where most of the larvae grouped with all other populations, and most closely to the Western Cape stock. Thus, most sardines undertaking the sardine run may have not been spawned in KZN according to Freon*, et al.* (2010) test 15.

Chapter five:

Conclusions

The Sardine run is an important ecological and economic phenomenon (van der Lingen, Coetzee, et al. 2010). Recent investigation has allowed for a better understanding about why and how sardines undergo this migration up the east coast of South Africa. The hypotheses about the nature of the sardine run can allow for insights into these questions. Information collected in this study could allow scientists to plan management strategies based on morphological and genetic diversity, connectivity, and breeding patterns within the KZN sardines. Ultimately, understanding genetic relationships and breeding patterns of the migrating fish stocks will allow for the conservation of genetic diversity of the fish. Genetic diversity allows for genetic resilience to the South African *Sardinops sagax* and the preservation of the KZN sardine run. In this study, these questions have been tested and the structure of the KZN sardines based on genetic and morphological techniques has been investigated.

The global genetic structure of *Sardinops sagax* suggests that the South African stock has diverged enough to be considered a separate subspecies to the Canadian *S. sagax*. These findings concur with previous allozymes and mtDNA genetic analysis by Grant and Bowen (1998). Based on the findings of this study, *S. sagax* from Canada shows a similar degree of genetic divergence with the South African sardine as the South African sardine shares with *S. melanostictus* of Japan. However, it is argued that the shallow divergence observed in this study (during the Pleistocene less than 230 000 years ago) of these species does not validate a separate species grouping. Thus, the evidence presented is suggestive that *S. sagax* from Canada and South Africa could be delineated into two separate subspecies. However, more appropriate evidence (such as the inclusion of a larger mtCOI sample set, and more regions) is needed before a reclassification is warranted. Further study into defining the species *Sardinops* worldwide would ultimately allow for a final solution to the above dilemma in taxonomy.

Genetic and morphological analyses both described differences in population structure between regions. Although there seemed to be larger population differences observed according to morphology, genetic structuring was considered to be moderately low but significant, especially considering the nature of clupeid life history and ecology. Nevertheless, morphology and genetic analyses seemed to show a positive association of difference among the populations tested in this study. Mitochondrial (mtCOI) and nuclear (EPIC) markers reflected the observed genetic structuring.

Tests for the ultimate and proximate factors of the sardine run (Freon*, et al.* 2010) are summarised below (Table 15). Test 3 (Are sardines from the run phenotypically distinct from the rest of the population?), the only test not related to genetics, was supported by the morphology of the sardines. KZN sardines are genetically different from the rest of the population, with an overall smaller head and body size. This formed part of the questions that were used to test whether sardines undertaking the sardine run form part of a subpopulation spawning migration (H_{U1}). Test 3, 4 and 7 were investigated in this study, and all could not be rejected, thus supporting the notion that sardines undergoing the sardine run are a sub population spawning migration. Test 4 and 7 were also used to test whether sardines undergo the sardine run because they exhibit some form of relic behaviour (H_{U7}), which was also supported. The hypothesis that sardines undergo natal homing and imprinting (H_{P1}) was rejected with the rejection of test 15, where the data showed that sardine run sardines did not only spawn in KZN. Genetic analysis has shown that most sardines belong to a finely structured interbreeding population, where most of the larvae grouped with other populations, not only the sardine run adults. Test eight could not be assessed using the scope of the analyses in this study. Thus, all the tests for H_{U1} and H_{U7} except test eight were supported.

Table 15: Conclusions of tests from Freon et al. (2010), based on morphometric and genetic data in this study, used to test the ultimate and proximate hypotheses relating to the sardine run, including caveats to the analyses and proposed further analysis. Strength and reliability of the tests were based on the statistical analyses undertaken and their results.

Test	Question	Result	Strength of result and reliability	Analyses used in test	Caveats	Further analysis
3	Are sardines from the run phenotypically distinct from the rest of the population?	Yes, there are significant morphological differences between the sardine run fish compared with other regions. Sardine run individuals are more phenotypically related to Port Elizabeth and Mossel Bay samples than the Cape Town individuals.	++ ***	Meristic, multivariate and geometric morphometric analyses.	Small sample size for KZN morphometric data.	Larger sample sizes, including more west coast and KZN samples.
4	Are sardines from the sardine run genetically distinct from the rest of the population?	Yes, KZN adults illustrate some isolation from Cape samples. However, KZN larvae illustrate significant genetic affiliation to Cape Town, Mossel Bay and Port Elizabeth adults.	+ *	mtCOI and EPIC marker phylogenetic and population analyses	EPIC data analyses compared sardine run sardines to Cape samples only.	More temporal samples for KZN and addition of Agulhas bank samples for EPIC data analysis.
7	Is there successful recruitment arising from individuals spawning in KZN waters?	Yes, migration rates and structure between larvae and all populations suggest this.	+ **	mtCOI and EPIC marker phylogenetic and population analyses.	-	Collection and genetic analysis of more larvae samples from all regions is needed. Multiple temporal samples during and after the sardine run should be taken.
8	Do all individuals of the sardine run subpopulation only spawn in KZN?	Inconclusive. Structuring suggests not.	0 *	Comparisons with Cape larvae needed for test, thus reducing the null effects of a shadow population.	N/A	"
15	Are sardine run sardines mostly spawned in KZN?	Yes, according to EPIC data. Sardines of the sardine run from 2011 belong to a structured interbreeding population, where most of the larvae grouped with the KZN larvae population. This was, however, not true for the KZN 2012 sardine run sardines, which showed migration and genetic affiliation to the Cape Town population.	+ ***	mtCOI and EPIC marker phylogenetic and population analyses.	These results represent the sardine population over two years only. Within sardine run differences in shoals may yield more accurate results. mtCOI data however may have been too broad scale in scope to be able to delineate population level structure, thus failing to show genetic structure.	"

Strength of Result: - negative; - - strongly negative; 0 inconclusive; + positive; ++ strongly positive. Strength and reliability of test: * low; ** medium; *** high.

Although the results of genetic analysis supported the ultimate hypotheses, the observed structure may have also been an indication that KZN is a zone of further interbreeding for the sardine populations of the Western Cape and Agulhas Bank. To test this hypothesis, a more comprehensive study of population structure, by including larvae samples from all regions and better spatial sampling coverage of the sardine run is needed. These tests serve to answer the final question (test 8: Do all individuals of the sardine run subpopulation only spawn in KZN?) that remains to be *answered* and including the notion of interbreeding of the western and southern stock populations in KZN.

Having said this, the current study has raised important points for consideration by fisheries managers. Most importantly, fish taking part in the sardine run are genetically diverse and distinct at the population level from the Western Cape population. Thus, fisheries managers should take this into account when developing fishery models of *S. sagax* in South Africa. This study has also raised the importance of utilising many samples from the sardine run as possible. The usage of the nine EPIC markers was considered successful.

References

Adams DC, Rohlf FJ, Slice DE. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. Italian Journal of Zoology, 71: 5-16.

Agüero J, Rodríguez F. 2004. *Morphometric stock structure of the Pacific sardine Sardinops sagax (Jenyns, 1842) off Baja California, Mexico*. In: Ashraf M, Elewa T (eds). *Morphometrics*: Springer. p. 115-127.

Akkers TR, Melo YC 1996. Spawning seasonality in the South African pilchard. In: Barange M, C vdL editors. WOSAS - Workshop on South African sardine: Proceedings and Recommendations.: Benguela Ecology Programme Report No. 29. p. 22-25.

Akkers TR, Melo YC. 1998. Aspects of pilchard reproduction. Sea Fisheries Research Institute Report No. WG/AUG98/PEL/2. Cape Town. p. 1-8.

Altschul SF, Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. . 1990. Basic local alignment search tool Journal of Molecular Biology, 215: 403-410.

Armstrong MJ, Chapman P, Dudley SFJ, Hampton I, Malan PE. 1991. Occurrence and population structure of pilchard, *Sardinops ocellatus*, round herring, *Etrumeus whiteheadi* and anchovy, *Engraulis capensis* off the east coast of southern Africa. South African Journal of Marine Science, 11: 227-249.

Atarhouch T, Rüber L, Gonzalez EG, Albert EM, Rami M, Dakkak A, Zardoya R. 2006. Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). Molecular phylogenetics and evolution, 39: 373-383.

Baibai T, Oukhattar L, Quinteiro JV, Mesfioui A, Rey-Mendez M. 2012. First global approach: morphological and biological variability in a genetically homogeneous population of the European pilchard, *Sardina pilchardus* (Walbaum, 1792) in the North Atlantic coast. Reviews in Fish Biology and Fisheries, 22: 63-80.

Baird D 1971. Seasonal occurrence of the pilchard *Sardinops ocellata* on the east coast of South Africa. Investigational Report, Division of Sea Fisheries. South Africa. p. 1-19.

Bandelt H, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biogical Evolution, 16: 37-48.

Beckley L, Naidoo A. 2003. Exploratory trials with light-traps to investigate settlement stage fishes in subtropical, coastal waters off South Africa. African Zoology, 38: 333-342.

Beckley LE, Hewitson J. 1994. Distribution and abundance of clupeoid larvae along the east coast of South Africa in 1990/91. South African Journal of Marine Science, 14: 205-212.

Beckley LE, van der Lingen CD. 1999. Biology, fishery and management of sardines (*Sardinops sagax*) in southern African waters. Marine and Freshwater Research, 50: 955-978.

Beerli P. 2008. Migrate version 3.2 a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available at http://popgen.scs.edu/migrate.html [accessed 10 April 2010].

Begg GA, Friedland KD, Pearce JB. 1999. Stock identification and its role in stock assessment and fisheries management: an overview. Fisheries Research, 43: 1-8.

Brander K. 2010. Impacts of climate change on fisheries. Journal of Marine Systems, 79: 389-402.

Broekhuysen GJ, Liversidge R, Rand RW. 1961. The South African gannet *Morus capensis* . 1. Distribution and movements. Ostrich, 32: 1-19.

Burnaby T. 1966. Growth-invariant discriminant functions and generalized distances. Biometrics: 96-110.

Cadrin SX. 2000. Advances in morphometric identification of fishery stocks. Reviews in Fish Biology and Fisheries, 10: 91-112.

Carr AF. 1967. *So excellent a fish: a natural history of sea turtles.* New York: Schribner. p. 1-256.

Carvalho GR, Hauser L. 1995. *Molecular genetics and the stock concept in fisheries*, vol. 4. London: Chapman & Hall. p. 326-350.

Cawthorn DM, Steinman HA, Witthuhn RC. 2012. DNA barcoding reveals a high incidence of fish species misrepresentation and substitution on the South African market. Food Research International, 46: 30-40.

Chenuil A, Hoareau T, Egea E, Penant G, Rocher C, Aurelle D, Mokhtar-Jamai K, Bishop J, Boissin E, Diaz A*, et al.* 2010. An efficient method to find potentially universal population genetic markers, applied to metazoans. BMC Evolutionary Biology, 10: 276.

Claridge MF, Dawah HA, Wilson MR. 1997. *Species: the units of biodiversity*. London: Chapman & Hall Ltd. p. 1-442.

Coetzee JC 1996. Environnmental parameters affecting sardine distribution and microscale studies of sardine aggregative behaviour. In: M B, Van der Lingen CD editors. WOSAS - Workshop on South African sardine: Proceedings and Recommendations Benguela Ecology Programme Report No. 29. p. 38-43.

Coetzee JC, Merkle D, de Moor CL, Twatwa NM, Barange M, Butterworth DS. 2008. Refined estimates of South African pelagic fish biomass from hydro-acoustic surveys: quantifying the effects of target strength, signal attenuation and receiver saturation. African Journal of Marine Science, 30: 205-217.

Coetzee JC, Merkle D, Hutchings L, van der Lingen CD, van den Berg M, Durholtz MD. 2010. The 2005 KwaZulu-Natal sardine run survey sheds new light on the ecology of small pelagic fish off the east coast of South Africa. African Journal of Marine Science, 32: 337-360.

Coetzee JC, Van der Lingen CD, Hutchings L, Fairweather TP. 2008. Has the fishery contributed to a major shift in the distribution of South African sardine? ICES Journal of Marince Science, 65: 1676-1688.

Connell AD 1996. Seasonal trends in sardine spawning at Park Rynie, Natal south coast. In: Barange M, C vdL editors. WOSAS - Workshop on Southern African Sardine: Proceedings and Recommendations. Benguela Ecology Programme Report No. 29.: Benguela Ecology Programme Report No. 29. p. 29-33.

Connell AD. 2010. A 21-year ichthyoplankton collection confirms sardine spawning in KwaZulu-Natal waters. African Journal of Marine Science, 32: 331-336.

Connell AD. 2012. Marine Fish Eggs and Larvae From the East Coast of South Africa. Available at http://www.dnabarcodes2011.org/documents/presentations/12-1/vertebrates/1000-Steinke.pdf [accessed 10th January 2014].

Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. Trends in Ecology & Evolution, 15: 290-295.

Croft DP, Krause J, Couzin ID, Pitcher TJ. 2003. When fish shoals meet: outcomes for evolution and fisheries. Fish and Fisheries, 4: 138-146.

De Goede J, Van der Lingen C. 2005. Application of a fat staging technique to examine temporal and spatial variability in lipid content of southern Benguela sardine Sardinops sagax. African Journal of Marine Science, 27: 671-675.

de Moor C, Butterworth D. 2009. A 2-stock hypothesis for South African sardine: two discrete stocks. Unpublished report, MCM/2009/SWG-PEL/23. Cape Town: Marine and Coastal Management, Department of Environmental Affairs and Tourism.

de Moor C, Butterworth D, Coetzee J. 2008. Revised estimates of abundance of South African sardine and anchovy from acoustic surveys adjusting for echosounder saturation in earlier surveys and attenuation effects for sardine. African Journal of Marine Science, 30: 219-232.

Dicken M. 2010. Socio-economic aspects of boat-based ecotourism during the sardine run within the Pondoland MPA, South Africa. African Journal of Marine Science, 32: 405-411.

Dingle H. 1996. *Migration: the biology of life on the move*. London: Oxford University Press on Demand. p. 1-336.

Dudley SFJ, Cliff G. 2010. Influence of the annual sardine run on catches of large sharks in the protective gill nets off KwaZulu-Natal, South Africa, and the occurrence of *Sardinops sagax* in shark diet. African Journal of Marine Science, 32: 383-397.

Durand E, Jay F, Gaggiotti OE, François O. 2009. Spatial inference of admixture proportions and secondary contact zones. Molecular Biology and Evolution, 26: 1963-1973.

Earl D. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4: 359-361.

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

Excoffier L, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1: 47-50.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131: 479-491.

Fennessy ST, Pradervant P, de Bruyn PA. 2010. Influence of the sardine run of selected nearshore predatory teleosts in KwaZulu-Natal. African Journal of Marine Science, 32: 375-382.

Florence W, Hulley P, Stewart B, Gibbons M. 2002. Genetic and morphological variation of the lanternfish Lampanyctodes hectoris (Myctophiformes: Myctophidae) off southern Africa. South African Journal of Marine Science, 24: 193-203.

Freon P, Coetzee JC, van der Lingen CD, Connell AD, O'Donoghue SH, Roberts HR, Demarcq H, Attwood C, Lamberth SJ, Hutchings L. 2010. A review and tests of hypotheses about causes of the KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 449-479.

Frézal L, Leblois R. 2008. Four years of DNA barcoding: current advances and prospects. Infection, Genetics and Evolution, 8: 727-736.

Gonzalez EG, Zardoya R. 2007. Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). BMC Evolutionary Biology, 7: 197.

Gradstein FM, Ogg JG, Smith AG. 2004. *A geologic time scale 2004*, vol. 86. Cambridge University Press.

Grant W, Bowen B. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. Journal of Heredity, 89: 415-426.

Grant WS, Bas C, Margalef R, Rubies P. 1985. *Population genetics of the southern African pilchard, Sardinops ocellata, in the Benguela upwelling system*. *International Symposium on the Most Important Upwelling Areas off Western Africa (Cape Blanco and Benguela)*. Barcelona: Instituto de Investigaciones Pesqueras. p. 551-562.

Hall T 1999. BioEdit computer program, version 7.0. 9. Computer program and documentation, Ibis Biosciences Carlsbad, CA.

Hammer O, Harper D, Ryan P 2012. PAST: paleontological statistics software package for education and data analysis. Palaeontologia electronica 4 (art. 4): 9.

Hansen MM. 2003. Application of molecular markers in population and conservation genetics, with special emphasis on fishes. DSc, University of Aarhus, Denmark.

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

Hassan M, Lemaire C, Fauvelot C, Bonhomme F. 2002. Seventeen new exon-primed intron-crossing polymerase chain reaction amplifiable introns in fish. Molecular Ecology Notes, 2: 334-340.

Hauser L, Carvalho GR. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9: 333-362.

Hauser L, Ward RD. 1998. Population identification in pelagic fish: the limits of molecular markers. NATO ASI Series a Life Sciences, 306: 191-224.

Hedgecock D, Hutchinson ES, Li G, Sly FL, Nelson K. 1989. Genetic and morphometric variation in the Pacific sardine, *Sardinops sagax caerulea*: comparisons and contrasts with historical data and with variability in the northern anchovy, *Engraulis mordax*. Fishery Bulletin, 87: 653-671.

Hellberg ME, Burton RS, Neigel JE, Palumbi SR. 2002. Genetic assessment of connectivity among marine populations. Bulletin of Marine Science, 70: 273-290.

Heydorn AEF, Bang ND, Pearce AF, Flemming BW, Carter RA, Schleyer MH, Berry PF, Hughes GR, Bass AJ, Wallace JH*, et al.* 1978. Ecology of the Agulhas Current region: an assessment of biological responses to environmental parameters in the south-west Indian ocean. Transactions of the Royal Society of South Africa, 43: 151-190.

Huelsenbeck J, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. Version 3. Bioinformatics, 17: 754-755.

Hulme T. 1995. The use of vertebral counts to discriminate between North Sea herring stocks. ICES Journal of Marine Science, 52: 775-779.

Hutchings L, Morris T, van der Lingen CD, Lamberth SJ, Connell AD, Taljaard S, van Niekerk L. 2010. Ecosystem considerations of the KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 413-421.

IBM Corp. 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, New York.

Ivanova NV, Zemlak TS, Hanner RH, Herbert PD. 2007. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes, 7: 544-548.

Keski̇n E, Atar HH. 2012. Genetic structuring of European anchovy (*Engraulis encrasicolus*) populations through mitochondrial DNA sequences. Mitochondrial DNA, 23: 62-69.

Keskİn E, Atar HH. 2013. DNA barcoding commercially important fish species of Turkey. Molecular Ecology Resources: 788–797.

Law R. 2000. Fishing, selection, and phenotypic evolution. ICES Journal of Marine Science, 57: 659-668.

Law R. 2007. Fisheries-induced evolution: present status and future directions. Marine Ecology Progress Series, 335: 271-277.

Li C, Riethoven J-JM, Ma L. 2010. Exon-primed intron-crossing (EPIC) markers for non-model teleost fishes. BMC Evolutionary Biology, 10: 90.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25: 1451-1452.

Liu J-X, Gao T-X, Zhuang Z-M, Jin X-S, Yokogawa K, Zhang Y-P. 2006. Late Pleistocene divergence and subsequent population expansion of two closely related fish species, Japanese anchovy (Engraulis japonicus) and Australian anchovy (Engraulis australis). Molecular phylogenetics and evolution, 40: 712-723.

Lutjeharms JRE, Steele JH, Turekian KK, Thorpe SA. 2001. *Agulhas Current*. *Encyclopaeida of Ocean Sciences 1*: Academic Press. p. 104-113.

Martínez P, González EG, Castilho R, Zardoya R. 2006. Genetic diversity and historical demography of Atlantic bigeye tuna (*Thunnus obesus*). Molecular phylogenetics and evolution, 39: 404-416.

McDowall RM. 2008. Jordan's and other ecogeographical rules, and the vertebral number in fishes. Journal of Biogeography, 35: 501-508.

McGarigal K 2000. Multivariate Statistics for Wildlife and Ecology Research. New York: Springer-Verlag. p. 283.

Messing J. 1983. New M 13 vectors for cloning. Methods in enzymology, 101: 20-78.

Miller DCM, Moloney C, van der Lingen CD, Lett C, Mulon C, Field JG. 2006. Modelling the effects of physical-biological interactions and spatial variability in spawning and nursery areas on recruitment of sardine in the southern Benguela ecosystem. Journal of Marine Systems, 61: 212-229.

Myeza J, Mason RB, Peddemors VM. 2010. Socio-economic implications of the KwaZulu-Natal sardine run for local indigenous communities. African Journal of Marine Science, 32: 399-404.

Natoli A, Peddemors VM, Hoelzel AR. 2008. Population structure of bottlenose dolphins (*Tursiops aduncus*) impacted by bycatch along the east coast of South Africa. Conservation Genetics, 9: 627-636.

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Grønkjær P. 2003. Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. Molecular ecology, 12: 1497-1508.

Nylander J. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, 2.

O'Donoghue SH. 2009. The Sardine Run: Investigating sardine, *Sardinops sagax,* and predator distributions in relation to environmental conditions using GIS and remotely sensed products. PhD, University of KwaZulu-Natal, Durban.

O'Donoghue SH, Drapeau L, Dudley SFJ, Peddemors VM. 2010. The KwaZulu-Natal sardine run: shoal distribution in relation to near-shore environmental conditions, 1997 to 2007. African Journal of Marine Science, 32: 293-307.

O'Donoghue SH, Drapeau L, Peddemors VM. 2010. An investigation of broad-scale distribution patterns of sardines and predators in relation to remotely sensed environmental conditions. African Journal of Marine Science, 32: 279-291.

O'Donoghue SH, Whittington PA, Dyer BM, Peddemors VM. 2010. Abundance and distribution of avian and marine mammal predators of sardine observed during the 2005 KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 361-374.

Olofsson H, Ripa J, Jonzén N. 2009. Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proceedings of the Royal Society B: Biological Sciences, 276: 2963-2969.

Palsbøll PJ, Berube M, Allendorf FW. 2007. Identification of management units using population genetic data. Trends in Ecology & Evolution, 22: 11-16.

Palumbi SR. 1992. Marine speciation on a small planet. Trends in Ecology & Evolution, 7: 114-118.

Palumbi SR. 1994. Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics, 25: 547-572.

Parrish R, Serra R, Grant W. 1989. The monotypic sardines, *Sardina* and *Sardinops*: their taxonomy, distribution, stock structure, and zoogeography. Canadian Journal of Fisheries and Aquatic Sciences, 46: 2019-2036.

PCN. vdB. 1978. The Natal sardine run. Annual Report No. 1. Sea Fisheries Branch, Department of Industries Report No. 100. Cape Town. p. 30.

Peakall RSPE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics, 28: 2537-2539.

Peddemors VM. 1993. Bottlenose dolphin *Tursiops truncatus* research in Natal, South Africa: A review. Lammergeyer, 42: 24-34.

Peddemors VM. 1999. Delphinids of southern Africa: a review of their distribution, status and life history. Journal of Cetacean Research and Management, 1: 157-165.

Pereyra R, Saillant E, Pruett C, Rocha-Olivares A, Gold J. 2004. Characterization of polymorphic microsatellites in the Pacific sardine *Sardinops sagax sagax* (Clupeidae). Molecular Ecology Notes, 4: 739-741.

Posada D, Crandall K. 2005. Modeltest 3.7. Program distributed by the author. Universidad de Vigo, Spain.

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

Punt AE, Hilborn R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35-63.

Rambaut A, Suchard M, Xie D, Drummond A 2014. Tracer v1.6.

Reed C, MacKenzie K, van der Lingen C. 2012. Parasites of South African sardines, *Sardinops sagax*, and an assessment of their potential as biological tags. Bulletin of the European Association of Fish Pathologists, 32: 41-48.

Rees M, Jessica C, Metcalf E, Childs DZ. 2010. Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proceedings of the Royal Society B: Biological Sciences, 277: 1149-1151.

Ripa J, Olofsson H, Jonzén N. 2010. What is bet-hedging, really? Proceedings of the Royal Society B: Biological Sciences, 277: 1153-1154.

Roberts MJ, van der Lingen CD, Whittle C, van den Berg M. 2010. Shelf currents, lee-trapped and transient eddies on the inshore boundary of the Agulhas Current, South Africa: their relevance to the KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 423-447.

Rohlf FJ, Bookstein FL. 1987. A comment on shearing as a method for “size correction”. Systematic Biology, 36: 356-367.

Rouault M, Pohl B, Penven P. 2010. Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. African Journal of Marine Science, 32: 237-246.

Ryynanen HJ, Primmer CR. 2006. Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (*Salmo salar*) and other salmonid fishes. BMC Genomics, 7: 192.

Scheet P, Stephens M. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. The American Journal of Human Genetics, 78: 629-644.

Schenekar T, Weiss S. 2011. High rate of calculation errors in mismatch distribution analysis results in numerous false inferences of biological importance. Heredity, 107: 511-512.

Silva A. 2003. Morphometric variation among sardine (*Sardina pilchardus*) populations from the northeastern Atlantic and the western Mediterranean. ICES Journal of Marine Science, 60: 1352-1360.

Stabell OB. 1984. Homing and olfaction in salmonids: a critical review with special reference to the Atlantic salmon. Biological Reviews, 59: 333-388.

Steinke D, Zemlak TS, Boutillier JA, Hebert PD. 2009. DNA barcoding of Pacific Canada's fishes. Marine Biology, 156: 2641-2647.

Steinke D, Zemlak TS, Connell AD, Heemstra PC, Hebert PDN. 2011. DNA Barcodes: Linking adults and immatures of marine fishes. Unpublished.

Stevens JP. 2012. *Applied multivariate statistics for the social sciences*. New York: Routledge Academic. p. 235-413.

Sunnucks P, Wilson A, Beheregaray L, Zenger K, French J, Taylor A. 2000. SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. Molecular ecology, 9: 1699-1710.

Swofford D 2002. PAUP*. Version 4.1. Sunderland, MA: Sinauer Associates.

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

Tinti F, Di Nunno C, Guarniero I, Talenti M, Tommasini S, Fabbri E, Piccinetti C. 2002. Mitochondrial DNA sequence variation suggests the lack of genetic heterogeneity in the Adriatic and Ionian stocks of Sardina pilchardus. Marine Biotechnology, 4: 163-172.

Touriya A, Rami M, Cattaneo-Berrebi G, Ibanez C, Augros S, Boissin E, Dakkak A, Berrebi P. 2003. Primers for EPIC amplification of intron sequences for fish and other vertebrate population genetic studies. Biotechniques, 35: 676-681.

Turner TF, Wares JP, Gold JR. 2002. Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). Genetics, 162: 1329-1339.

van der Lingen CD. 2002. Diet of sardine *Sardinops sagax* in the Southern Benguela upwelling ecosystem. South African Journal of Marine Science, 24: 301-316.

van der Lingen CD, Coetzee JC, Demarcq H, Drapeau L, Fairweather TP, Hutchings L. 2005. An eastward shift in the distribution of southern Benguela sardine. Globec International Newsletter, 11: 17-22.

van der Lingen CD, Coetzee JC, Hutchings L. 2010. Overview of the KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 271-277.

van der Lingen CD, Durholtz MD 2004. Review of variability in biomass, growth, and other characteristics of anchovy Engraulis encrasicolusi and sardine Sardinops sagax in the Benguela current upwelling ecosystem. Cape Town: FRA/APN/IAI/GLOBEC/PICES Workshop "Global comparison of sardine, and anchovy and other small pelagics: building towards a multi-species model,

Tokyo, Japan"; see www.myg.affrc.go.jp/kaiyo/sardine-anchovy/index-sardine-anchovy.html.

van der Lingen CD, Durholtz MD, Fairweather TP, Melo Y. 2009. Spatial variability in biological characteristics of southern Benguela sardine and the possible existence of two stocks. Unpublished report, MCM/2009/SWG-PEL/39. Cape Town: Marine and Coastal Management, Department of Environmental Affairs and Tourism.

van der Lingen CD, Freon P, Fairweather TP, van der Westhuizen JJ. 2000. Density-dependent changes in reproductive parameters and condition of southern Benguela sardine *Sardinops sagax*. African Journal of Marine Science.

van der Lingen CD, Freon P, Fairweather TP, van der Westhuizen JJ. 2006. Density-dependent changes in reproductive parameters and condition of southern Benguela sardine *Sardinops sagax*. African Journal of Marine Science, 28: 625-636.

van der Lingen CD, Hendricks M, Durholtz MD, Wessels G, Mtengwane C. 2010. Biological characteristics of sardine caught by the beach-seine fishery during the KwaZulu-Natal sardine run. African Journal of Marine Science, 32: 309-330.

Ward R. 2000. Genetics in fisheries management. Hydrobiologia, 420: 191-201.

Ward RD, Hanner R, Hebert PDN. 2009. The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biology, 74: 329-356.

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

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution: 1358-1370.

Wessels G. 2009. Meristic and morphometric variation among sardine *Sardinops sagax* around the coast of southern Africa. . MSc Dissertation, University of Cape Town.

Wessels G, Moloney CL, van der Lingen CD. 2010. The effects of freezing on the morphometrics of sardine *Sardinops sagax* (Jenyns, 1842). Fisheries Research, 106: 528-534.

Whitehead J, Nelson G, Wongratana T. 1988. FAO. Species Catalogue. Clupeoid fishes of the world (Suborder Clupeioidei), An annotated and illustrated catalogue of the Herrings, Sardines, Pilchards, Sprats, Shads, Anchovies and Wolf-herrings Part 2. Engraulidae. FAO Fish. Synop, 125.

Whitehead PJP, Wongratana T. 1986. *Clupeidae*. In: Smith MM, Heemstra PC (eds). *Smith's Sea Fishes*. Berlin: Springer-Verlag. p. 199-204.

Wyatt T, Cushing DH, Junquera S. 1991. *Stock distinctions and evolution of European sardine.Long term variability of pelagic fish populations and their environment.* . In: Kawasaki T, Tanaka S, Toba Y, Taniguchi A (eds). *Long term variability of pelagic fish populations and their environment.* Oxford: Pergamon Press. p. 229-238.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN. 2009. DNA barcoding reveals overlooked marine fishes. Molecular Ecology Resources, 9: 237-242.

Zhang J, Hanner R. 2011. DNA barcoding is a useful tool for the identification of marine fishes from Japan. Biochemical Systematics and Ecology, 39: 31-42.

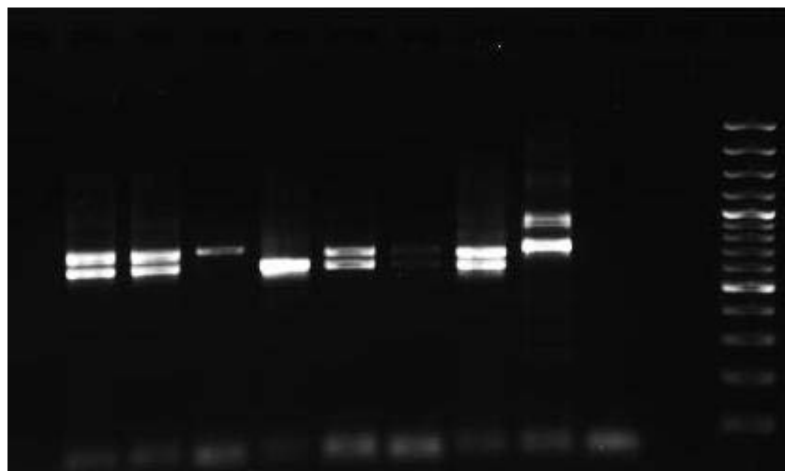
Zinetti F, Dapporto L, Vanni S, Magrini P, Bartolozzi L, Chelazzi G, Ciofi C. 2013. Application of molecular genetics and geometric morphometrics to taxonomy and conservation of cave beetles in central Italy. Journal of Insect Conservation, 17: 921-932.

Appendix 1

Typical photograph used in the morphometric analysis of *Sardinops sagax* in this study.

Appendix 2

A 3 % agarose gel electropherogram run for 5 hours at 100 volts. The gel electrophoresis depicts polymorphisms used to genotype sardines for each of the nine EPIC markers in this study.



Appendix 3

Hardy Weinberg Equilibrium test for sardines, including the significances of heterozygosity for each EPIC marker between the sites.

Population	Locus	DF	Chi^2	Probability	Significance
KZN 2011	MLc	21	36.168	0.021	*
	Chymb	6	13.296	0.039	*
	aldob1	3	6.158	0.104	Ns
	Tr1	15	70.348	0.000	***
	Ops1	3	22.427	0.000	***
	AldoB5	6	17.149	0.009	**
	Cam-3	10	31.241	0.001	***
	Act	10	37.106	0.000	***
	GPd	3	23.802	0.000	***
KZN 2012	MLc	15	49.384	0.000	***
	Chymb	6	51.069	0.000	***
	aldob1	1	37.349	0.000	***
	Tr1	21	61.331	0.000	***
	Ops1	3	13.358	0.004	**
	AldoB5	10	12.350	0.262	ns
	Cam-3	10	35.134	0.000	***
	Act	10	30.071	0.001	***
	GPd	3	8.374	0.039	*
KZN Larvae	MLc	15	30.853	0.009	**
	Chymb	6	10.026	0.124	ns
	aldob1	3	4.112	0.250	ns
	Tr1	10	28.189	0.002	**
	Ops1	3	4.301	0.231	ns
	AldoB5	3	5.383	0.146	ns
	Cam-3	15	27.508	0.025	*
	Act	3	27.590	0.000	***
	GPd	3	9.029	0.029	*
CT	MLc	21	55.947	0.000	***
	Chymb	3	12.600	0.006	**
	aldob1	3	33.757	0.000	***
	Tr1	28	118.110	0.000	***
	Ops1	3	5.996	0.112	ns
	AldoB5	10	21.555	0.018	*
	Cam-3	15	50.704	0.000	***
	Act	10	15.909	0.102	ns
	GPd	3	1.359	0.715	ns

Appendix 4

Haplotype list with individual assignment to haplotypes in DNAsp, for the mtCOI gene.

Haplotype number	Number of individuals representing haplotype	Individual identification
1	3	B1, B9, C13
2	19	B2, C18, F1, F4, H3, H4, H6, H7,PE3,PE8, CT1, CT2, CT8, CT9, CT11, JF494407.1, JF494410.1, JF494409.1, JF494412.1
3	1	B3
4	1	B4
5	2	B5, C12
6	1	B6
7	1	B7
8	1	B8
9	1	B11
10	3	C2, C3, C6
11	1	C4
12	1	C14
13	1	C15
14	1	C19
15	1	F2
16	1	F3
17	1	F6
18	1	H5
19	2	H8, H10
20	1	H9
21	1	PE1
22	1	PE6
23	1	PE7
24	1	PE11
25	1	CT3
26	1	CT4
27	1	CT5
28	1	CT6
29	1	CT7
30	1	CT10
31	1	CT12
32	1	CT13
33	1	JF494411.1
34	2	FJ165126.1, FJ165128.1
35	1	FJ165121.1
36	1	FJ165125.1
37	1	FJ165123.1
38	1	FJ165120.1
39	2	JF952843.1, JF952841.1
40	1	JF952842.1