

Genetic stock assessment of the Blue Bream, *Pachymetopon aeneum* (Pisces: Sparidae), a South African linefishery species, using microsatellite markers

By

**Sibusiso Yokwana
(200807479)**

A dissertation submitted in fulfilment of the requirements for the degree of
Master of Science (Zoology) in the Faculty of Science and Agriculture
at the University of Fort Hare



University of Fort Hare

Together in Excellence

University of Fort Hare

Together in Excellence

Department of Zoology and Entomology

Faculty of Science and Agriculture

University of Fort Hare, Alice, South Africa

Supervisors:

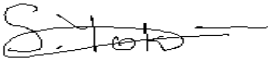
Dr Gavin Gouws, South African Institute for Aquatic Biodiversity

Prof. Niall G. Vine, University of Fort Hare

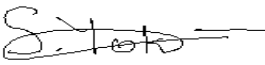
July 2024

Declaration

I, **Sibusiso Yokwana**, student number **200807479**, hereby certify that this dissertation is my original work and has not been submitted in this or any other form to any University other than the University of Fort Hare, Alice, South Africa. Where use has been made of the research of others, it has been duly acknowledged in this text.

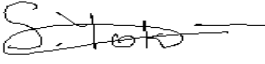
Signed: 

I, **Sibusiso Yokwana**, student number **200807479**, hereby declare that I am fully aware of the University of Fort Hare's policy on plagiarism and I have taken every precaution to comply with the regulations.

Signature: 



I, **Sibusiso Yokwana**, student number **200807479**, declare that I am fully aware of the University of Fort Hare's policy on research ethics and I have taken every precaution to comply with the regulations. I have obtained an ethical clearance certificate from the University of Fort Hare's Research Ethics Committee and my reference number is the following: **GOU011SYOK01**

Signature: 

Dedication

I would like to dedicate this body of work to the memory of my late father and my late grandmothers, who have been a constant source of inspiration and guidance throughout my life. I am honoured to have had their love and support, and I will always cherish the memories we shared. Secondly, I would like to express my gratitude to my kids, who have been a source of motivation and inspiration throughout this journey. I am grateful for their love and support, and I am honoured to have them in my life. Thank you for everything, and I love you always.



University of Fort Hare
Together in Excellence

Acknowledgements

First, I would like to express my sincerest gratitude to my supervisors, Dr Gavin Gouws and Prof. Niall G. Vine, for their invaluable guidance, advice and support throughout my research journey. Their expertise and encouragement have been instrumental in the completion of this work. I am forever indebted to them for their unwavering support and for providing me with the opportunity to grow as a researcher. Thank you both for everything

I would also like to extend my heartfelt thanks to NRF-ACEP Phuhlisa Programme for sponsoring this project. Garth van Heerden and Zipho Canda deserve special recognition for steering the NRF-ACEP Phulisa programme at SAIAB, ensuring that we received the necessary academic support throughout the project. Their dedication and commitment to the success of this project is greatly appreciated. Nangomso!

My deepest appreciation goes to my family, particularly my mom Nopinkie Yokwana, my Dabs Mrs NP Ntloko, Mrs N Boko, Miss N Maha, and Miss M Nolisi, for their unwavering support and constant prayers. Thank you to the Mputa and Yokwana families for their love and encouragement. Ndibulela ku-MaTshawe namaNkabane. Makube chosi, kube hele! I couldn't have done this without all of you.

I would also like to extend my sincerest thanks to my friends, who have been by my side throughout this research journey, offering a helping hand whenever needed. There are too many of you to mention here, it would take a whole thesis, but I appreciate your friendship through the good and bad times. Your unwavering support and encouragement has been priceless to me. Thank you for being a constant source of fun, positivity and inspiration.

Last but certainly not least, I would like to express my deepest gratitude to SAIAB. The entire institution has been amazing and welcoming. The staff and students have been like a second family to me, making my time in Makhanda an enjoyable experience. I will forever

cherish the relationships I have built there and the support they provided me during my research journey. I am grateful for the wealth of knowledge and resources that SAIAB has provided me, and I will always be proud to have been a part of such a wonderful institution.

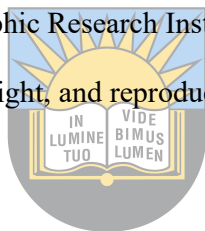


University of Fort Hare
Together in Excellence

Abstract

The blue bream (*Pachymetopon aeneum*) is a valuable species in South Africa's line fishery sector, necessitating effective management to ensure its preservation. This study utilised eight microsatellite markers to analyse the genetic diversity and structure of blue bream populations, providing insights to inform better management strategies. Microsatellites were selected for their effectiveness in assessing genetic diversity within and between populations below the species level.

Given the species' vulnerability to overexploitation due to slow growth, late maturity, and protogynous hermaphroditism a comprehensive genetic stock assessment is critical. Samples were collected from 27 locations across the species' range between 2009 and 2015 by DAFF, fishing observers, and the Oceanographic Research Institute. These samples, recorded with details such as location, date, size, weight, and reproductive stage, were grouped into seven regions for robust analyses.



This study aimed to determine the genetic diversity and structure of *P. aeneum* populations. DNA was extracted from samples, and eight microsatellite loci were amplified using fluorescently-dyed primers. Fragment analysis was conducted using the SANGER sequencer, and the fragment lengths were scored using the GeneMapper software, with quality control performed in Microchecker. Hardy-Weinberg equilibrium was assessed using Arlequin 3.5. Genetic diversity and population structure analyses were performed with GenAlex 6.5, Arlequin, and STRUCTURE 2.3.4.

The results indicated high levels of genetic diversity and low inbreeding, suggesting random mating across the distribution range. AMOVA, PCoA, and Bayesian assignment identified two subpopulations. Relatedness and population assignment analyses supported these

findings. Isolation by distance analysis showed no significant association between pairwise genetic differentiation and geographic distance.

The evidence of two subpopulations suggests that management strategies for *P. aeneum* should be reconsidered. Further investigation with a larger sample size is recommended.

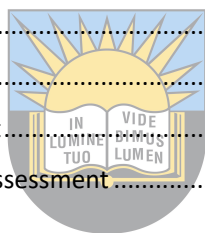
Additionally, fishing trends should be monitored, as excessive local harvesting could impact the entire stock.



University of Fort Hare
Together in Excellence

Contents

Declaration.....	Error! Bookmark not defined.
Dedication.....	Error! Bookmark not defined.
Acknowledgements.....	Error! Bookmark not defined.
Abstract.....	Error! Bookmark not defined.
Chapter 1.....	Error! Bookmark not defined.
Literature review.....	Error! Bookmark not defined.
General introduction.....	5
1.1 South African line fisheries sector	5
1.1.1 Description of the line fisheries sector	5
1.1.2 History of the South African line fisheries sector	5
1.1.3 Management of South African line fisheries sector	6
1.1.4 Status of the South African line fisheries sector	7
1.1.5 Challenges in the South African line fisheries sector	7
1.2 Stock assessment	8
1.2.1 Stock assessment description	8
1.2.2 Genetics in fish stock assessment.....	10
1.2.3 Genetic methods used in stock assessment.....	11
1.2.4 Selected Markers for this study	12
1.3 The study species: The blue bream (<i>Pachymetopon deneum</i>).....	Error! Bookmark not defined.3
1.3.1 Taxonomy and description.....	Error! Bookmark not defined.3
1.3.2 Distribution	Error! Bookmark not defined.4
1.3.3 Habitat	Error! Bookmark not defined.5
1.3.4 Diet.....	Error! Bookmark not defined.6
1.3.5 Reproduction	Error! Bookmark not defined.6
1.3.6 Movement.....	Error! Bookmark not defined.7
1.3.7 Management.....	Error! Bookmark not defined.7
1.4 Rationale	Error! Bookmark not defined.9
1.5 Aims and objectives	22
Chapter 2.....	23
Materials and Methods.....	23
2.1 Sampling.....	23
2.2 DNA extraction	24
2.3 Microsatellites data processing	Error! Bookmark not defined.



University of Fort Hare
Together in Excellence

2.3.1 PCR amplification and genotyping	Error! Bookmark not defined.
2.3.3 Microsatellite Analysis	Error! Bookmark not defined.
2.3.4 Hardy Weinberg Equilibrium (HWE)	Error! Bookmark not defined.
2.3.5 Linkage disequilibrium	Error! Bookmark not defined.
2.3.6 Genetic diversity	Error! Bookmark not defined.
2.3.7 Population differentiation	31
2.3.8 Analysis of Molecular Variance (AMOVA).....	32
2.3.9 Principal Coordinate Analysis (PCoA).....	33
2.3.10 Structure	34
2.3.11 Isolation by distance	Error! Bookmark not defined.
2.3.12 Relatedness	Error! Bookmark not defined.
2.3.13 Population assignment	Error! Bookmark not defined.
Chapter 3.....	Error! Bookmark not defined.
Results.....	Error! Bookmark not defined.
3.1 Microchecker	Error! Bookmark not defined.
3.2 Linkage disequilibrium	40
3.3 Genetic diversity	41
3.4 Population comparison.....	43
3.5 Bayesian clustering/assignment	Error! Bookmark not defined.
3.6 Principal Coordinate Analysis (PCoA).....	Error! Bookmark not defined.
3.7 Analysis of molecular variance (AMOVA)	Error! Bookmark not defined.
3.9 Isolation by distance	Error! Bookmark not defined.
3.10 Relatedness	50
3.11.1 Population Assignment (“As is” method)	51
3.11.2 Population Assignment (“Leave-one-out” method)	52
Chapter 4.....	Error! Bookmark not defined.
Discussion	Error! Bookmark not defined.
4.1 Comparison with other species	Error! Bookmark not defined.
4.2 Biological interpretation Fundamental data suitability	Error! Bookmark not defined.
Genetic diversity	Error! Bookmark not defined.
Overall patterns of population structure.....	Error! Bookmark not defined.0
4.6 Conclusion	Error! Bookmark not defined.
References	Error! Bookmark not defined.



University of Fort Hare
Together in Excellence



University of Fort Hare
Together in Excellence

Chapter 1

Literature review

The management of linefish resources in southern Africa is increasingly worrisome, driven by heightened fishing effort and technological advancements in fishing gear, techniques, and fish-finding technology (Mann 2013). These advancements make fishing more efficient but also increase the risk of overexploitation by enabling fishers to catch more fish in less time and target species that were previously harder to locate. Compounding this concern is the inadequacy of existing fisheries data, which rely on subjective perceptions rather than rigorous quantitative evaluations (Mann 2013).

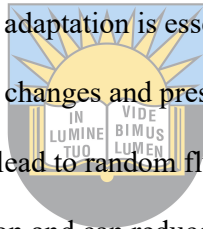
In the realm of fisheries management, the central concept of stock delimitation revolves around ensuring sustainable harvests for each managed stock unit (Carvalho & Hauser 1994). This process entails a comprehensive understanding of ecological and genetic processes, facilitating the equitable distribution of fishing effort and mortality among geographically-separated populations (Grimes et al. 1987, Begg & Waldman 1999).

The importance of genetic information becomes particularly evident in the identification of reproductively-isolated stocks or those with limited gene flow. This insight is crucial for crafting effective management strategies, given that smaller isolated stocks are notably more susceptible to the adverse impacts of fishing efforts and the subsequent risk of collapse but also to processes like genetic drift, which can further reduce genetic diversity and adaptive potential (Wright 1945, Tremblay & Ackerman 2001, Waples et al. 2008).

Examining connectivity and gene flow among marine species provides additional layers of understanding. Species with pelagic larvae exhibit a high potential for connectivity among local populations, and the study of gene flow through larval dispersal offers valuable insights

into the intricate connections among these populations (Sale 1991, Caley et al. 1996, Palumbi 2003, Levin 2006, Cowen & Sponaugle 2009). The intricacies of the population structure of marine fishes are fundamentally tied to genetic or genomic information. Understanding the movements of individuals, such as philopatry (return to birth location) or spawning philopatry, emerges as a critical factor in effective stock management (Hueter et al. 2005).

Local selection plays a significant role in shaping genetic structures within populations. In southern Africa, coastal species often exhibit regionally confined genetic lineages that are adapted to specific environmental conditions. For example, studies have shown that species such as *Perna perna* display genetic divergence linked to different marine biogeographic provinces, where varying selective pressures can lead to distinct evolutionary units (Teske et al. 2011, Zardi et al. 2007). This local adaptation is essential for understanding how populations respond to environmental changes and pressures. Furthermore, genetic drift, particularly in small populations, can lead to random fluctuations in allele frequencies. This process enhances genetic differentiation and can reduce overall genetic variation within populations. In the context of South African marine species, limited dispersal capabilities due to physical barriers, such as ocean currents and upwelling zones can exacerbate the effects of genetic drift (Teske et al. 2011, Zardi et al. 2007). For instance, the Benguela Current acts as a formidable barrier that influences gene flow between populations leading to pronounced genetic differences (Teske et al. 2011). Physical barriers like upwelling cells and coastal currents restrict the movement of organisms and their gametes, leading to isolated populations that evolve independently over time. This isolation can result in significant genetic divergence among populations that are geographically close but genetically distinct (Teske et al. 2011, Madeira 2024). The presence of such barriers highlights the importance of understanding oceanographic conditions in marine phylogeography.



University of Fort Hare
Together in Excellence

The South African coastline, extending from the Gariep (Orange) River on the west coast to Kosi Bay on the east coast, is divided into distinct biogeographic provinces based on species composition and hydrological conditions (Teske et al. 2006). These provinces are broadly categorized into three main regions: the cool-temperate west coast, the warm-temperate south coast, and the subtropical east coast (Harrison 2002, Teske et al. 2006). However, the precise boundaries of these regions remain a subject of debate, with various studies proposing differing delineations (Emanuel et al. 1992, Bolton et al. 2004). This variability arises from differences in the taxa studied, reliance on historical distribution records, variation in sampling depths, and inconsistent coverage of the coastline (Beckley & van Ballegooyen 1992, Harrison 2002, Bolton et al. 2004; Sink et al. 2004). Most research efforts have concentrated on specific taxonomic groups, such as fish larvae (Beckley & van Ballegooyen 1992), estuarine fish (Harrison 2002), intertidal invertebrates (Emanuel et al. 1992), and seaweeds (Bolton et al. 2004). Given the differential responses of these groups to environmental factors, establishing uniform biogeographic boundaries for all taxa is challenging (Harrison 2002).



University of Fort Hare
Together in Excellence

Both biotic and abiotic factors play a crucial role in shaping marine population structures. Even closely related species occupying similar habitats may exhibit significant differences in population structure (Todd et al. 1998, Riginos & Victor 2001). For instance, invertebrates with shorter larval durations tend to show higher levels of genetic differentiation compared to those with prolonged planktonic larval stages (Kyle & Boulding 2000, Collin 2001). Similarly, sessile marine species lacking dispersive larvae often exhibit greater genetic subdivision compared to their planktotrophic counterparts (Todd et al. 1998, Riginos & Victor 2001). While some studies suggest an inverse relationship between dispersal potential and genetic differentiation, others indicate that recruitment processes and local environmental conditions may override dispersal capabilities (Jiang et al. 1995, Knowlton 2000, Dawson &

Jacobs 2001). Certain species with high dispersal potential may have restricted distributions due to localized recruitment dynamics, whereas others thought to have wide ranges may actually consist of cryptic species complexes (Knowlton 2000, Dawson & Jacobs 2001). Despite these findings, dispersal mechanisms remain a key factor influencing the distribution of marine organisms (Avise et al. 1987).

Biotic changes along the South African coastline can be attributed to the limited dispersal ability of some species, which may result in genetic differentiation due to restricted gene flow across biogeographic barriers (Teske et al., 2011). Estuarine species provide a good example for this, while many estuarine species tolerate high salinity levels (de Villiers et al. 1999, Teske & Wooldridge 2003, 2004) and have extensive distribution ranges (Day 1974, Griffiths 1976, Steyn & Lussi 1998), populations may still experience isolation due to geographical distance and the ecological distinctions between estuarine and marine habitats (Teske & Wooldridge, 2003). These factors contribute to unique genetic structure within estuarine populations, highlighting the interplay between environmental and biological influences in shaping population dynamics along the South African coast (Teske et al. 2011).

In conclusion, the complexity of genetic differentiation in marine species reflects the interplay of historical and contemporary factors shaping population structure. Historical climatic events, such as glaciations and sea-level changes, alongside ocean currents and geographic barriers, have driven population isolation and divergence over evolutionary timescales (Pontes et al. 2024). More recently, human-induced pressures like overfishing and habitat destruction, coupled with environmental changes such as ocean warming and acidification, continue to influence gene flow and adaptation (Mendenhall 2020). These factors highlight the need for a nuanced understanding of genetic structure to inform the sustainable conservation and management of Linefish resources in southern Africa.

General introduction

1.1 South African line fisheries sector

1.1.1 Description of the line fisheries sector

The line fisheries sector involves the use of hook and line to catch fish (DAFF 2014). It encompasses a range of fishing methods, including modern longline fishing (Gammage 2017) and traditional hand-held, baited-hook fishing (DAFF 2013). The sector operates both inshore and offshore (Mann 2000), and includes commercial, recreational, and subsistence catches (Mann 2000). Longline fishing is a capital-intensive method that uses specialized vessels and technology to target high-value offshore species (DAFF 2013), while hand-held rod fishing is a more economical method often used for subsistence purposes in rural areas (DAFF 2013). The line fisheries sector in South Africa is diverse and plays a significant role in the country's economy and food security (DAFF 2013).



1.1.2 History of the South African line fisheries sector

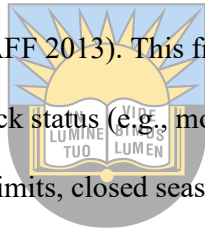
University of Fort Hare
Together in Excellence

The line fisheries sector in South Africa has a long and varied history, with evidence of fishing activities dating back to the 1500s and the arrival of Dutch seafarers and the Khoi people (Thompson 1913, Sauer 2003, Palmer 2008). Initially, the sector was concentrated in the coastal regions of the Eastern Cape (EC) and Western Cape (WC), later expanding to include the KwaZulu-Natal (KZN) region (Sauer 2003). The sector faced several challenges in its early history, including restrictions imposed by Dutch colonizers of the Cape that were lifted in 1795 after the British captured the region (Thompson 1913, Griffiths 2000, DAFF 2013). The traditional line fishing sector was initially dominated by row and sailboats (Sauer 2003), but the post-World War II period saw significant technological advancements and industrialization, including the rise of motorized vessels, small boat harbours, and echo-

locating technology (Sauer 2003, DAFF 2013). These developments significantly increased fishing efforts and catches, which resulted in the dramatic expansion of the line fisheries industry (Sauer 2003, DAFF 2013). This period of growth has continued to shape the sector to the present day.

1.1.3 Management of South African line fisheries sector

The management of the line fisheries sector in South Africa has undergone several changes over time in an effort to ensure the sustainability of the sector and the resources it relies upon (DAFF 2013). Initial management measures, including minimum size limits for certain species (Sauer 2003), were introduced in 1940. However, a decline in catch trends and an increase in fishing effort led to the implementation of a comprehensive management framework in 1985 (Penney 1989, DAFF 2013). This framework included revised size limits, categorization of species based on stock status (e.g., more vulnerable reef species versus more resilient shoaling species), bag limits, closed seasons for some species, and efforts to maintain 1985 levels (Sauer 2003, DAFF 2013).



University of Fort Hare
Together in Excellence

Despite these measures, catch rates for many species continued to drop and stocks of supposedly-resilient species were reported to be over-exploited (Sauer 2003, DAFF 2013). In response, a new Linefish Management Protocol (LMP) was created, which included a 70% reduction of effort and a limit of 450 vessels (DAFF 2016). This protocol remains the foundation of South African line fish management today, with the goal of rebuilding collapsed stocks, and achieving and maintaining sustainable levels of operation (DAFF 2016). In December 2000, the South African government along and with the aid of the Marine Linefish Research Group declared the line fish resource to be under a state of emergency, as provided for by the Marine Living Resources Act (MLRA, Act 18 of 1998) (DAFF 2013).

1.1.4 Status of the South African line fisheries sector

The status of the line fisheries sector in South Africa is currently a subject of concern due to the significant growth and expansion of the sector in recent years as a result of technological advancements and increased participation from line fishers (Sauer 2003, DAFF 2013, 2016). To address the sustainability of the resources being harvested, the commercial line fish sector is regulated through the use of Total Allowable Effort (TAE), which is allocated based on boat size and crew numbers (DAFF 2016). The TAE is also applied to shore-based and recreational line fishers in the form of a fish per person limit, for example, a limit of five fish per person per day for *Pachymetopon aeneum* (*P. aeneum*) (DAFF 2016).

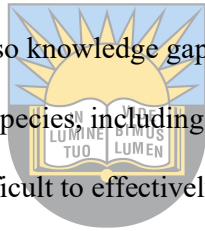
Despite these regulatory measures, a 2003 stock assessment report indicated that some fish stocks in the sector are still partially exploited or over-exploited (DAFF 2016). The introduction of the Linefish Management Protocol (LMP) in 1999 and the declaration of a state of emergency in 2000 demonstrate a commitment to improving the sustainability of the sector, but more work is needed to ensure the long-term viability of the line fisheries sector in South Africa (DAFF 2016). It is important for the sector to continue to monitor and assess the status of fish stocks in order to ensure the sustainable use of these resources.

1.1.5 Challenges in the South African line fisheries sector

The South African line fisheries sector faces several challenges that may impact its sustainability. These challenges include:

1. Vulnerability of line fishery-targeted species to overfishing: Many of the species targeted by line fishers in South Africa, such as sparids, are vulnerable to overfishing due to their life-history traits, including sex change, late maturity, aggregating behaviour, long life spans, and estuarine dependency (Cowley 2008, DAFF 2016).

2. Occurrence of line fishery-targeted species as by-catch: In addition to being targeted directly by line fishers, many of the species in this sector are also caught as by-catch in other fisheries, which hinders the implementation of effective management measures in the line fish sector (DAFF 2016).
3. Compliance and enforcement: Despite the implementation of the Linefish Management Protocol (MLP) and the presence of active Fisheries Compliance Officers (FCOs), compliance and enforcement in the line fishery sector remain significant issues (SASSI 2016). Indications of ineffective monitoring of landings and a lack of an observer program for the line fish sector have contributed to these challenges (SASSI 2016).
4. Knowledge gaps: There are also knowledge gaps regarding the movement and genetics of many line fishery species, including sparids (Mann 2013). This lack of understanding can make it difficult to effectively manage these species and ensure their sustainable use.



University of Fort Hare
Together in Excellence

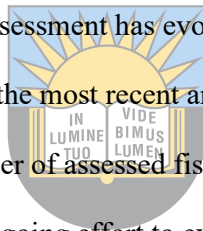
1.2 Stock assessment

1.2.1 Stock assessment description

Stock assessment is a crucial aspect of sustainable fishery management, as it involves the analysis of biological, monitoring and harvest data to predict the response of a fishery to potential changes in management policies (Hilborn & Walters 1992, Quinn & Deriso 1999). This is accomplished by using statistical models that are best suited to the available information to evaluate the status of the stock (Cadrin & Dickey-Collas 2014). There is often confusion surrounding the concept of a stock, as the terms "stock" and "population" are often used interchangeably and have overlapping definitions. A population is generally defined as a

large group of individuals of the same species that are capable of reproducing to sustain the species in a particular habitat or area (May 1984). On the other hand, a stock is often described as a group of fish with shared life-history traits that are capable of self-reproduction (Hilborn & Walters 1992). Some definitions of a stock also include genetic information, recognizing the importance of genetic diversity in maintaining the resilience of a species (Ovenden 1990). Regardless of the specific definition, it is clear that understanding the stock structure of a particular species is essential for effective management. Ignoring this information can result in the depletion of fish stocks and the loss of genetic diversity, increasing the risk of inbreeding and regional extinction (Nelson & Soule' 1987, Smith 1991).

South Africa's marine fishery stock assessment has evolved significantly, with the 2023 DEFF report (DEFF 2023) providing the most recent and comprehensive evaluation of the country's fishery resources. The number of assessed fish stocks increased from 43 in 2012 to 77 in the latest report, reflecting an ongoing effort to expand coverage and improve management strategies (DEFF 2023). The report indicates notable improvement in stock status, with 66% of the assessed stocks categorized as not of concern—classified as unknown, abundant, or optimal—compared to only 46% in 2012. However, 34% of stocks remain of concern, being classified as either depleted or heavily depleted. Notable changes in stock status since the 2020 report include an increase in the number of depleted stocks from nine to 13 (DEFF 2023). The report also highlights an increase in the number of stocks with unknown status and an increase in fishing pressure, emphasizing the challenges associated with data availability and resource monitoring. The findings underscore the complexity of South Africa's multi-species fisheries and the importance of adaptive management strategies to ensure sustainable utilization and long-term resource resilience (DEFF, 2023).



University of Fort Hare
Together in Excellence

1.2.2 Genetics in fish stock assessment

In stock assessment, genetic data can provide valuable insights into the genetic diversity, structure and dynamics of a fishery resource. Genetic variation within a species can be influenced by a range of factors, including demographic history, gene flow, and adaptation to different environments (Campana et al. 1995, Hewitt 1996, Laikre & Ryman 1996, Allendorf et al. 2008). Genetic diversity is important for the survival and adaptability of a species, as it allows the population to adapt to changing environments and helps to maintain the overall health of the species (Reusch et al. 2005, Schindler et al. 2010, Hellmair & Kinziger 2014).

In stock assessment, genetics can be used to identify the relationships between different populations, assess the degree of gene flow between populations, and detect the presence of genetically distinct subpopulations within a larger stock (Ovenden 1990, Allendorf & Lundrigan 1996). This information can be used to inform management decisions, such as setting harvest limits and establishing fishing quotas, to ensure the sustainable use of fishery resources. In addition, genetic data can be used in conjunction with other types of data, such as biological and ecological information, to improve the accuracy and precision of stock assessment models (Cadrin & Dickey-Collas 2014).

Ignoring genetic data in stock assessment may lead to inadequate management and have potentially negative impacts on genetic diversity. For example, a decline in an unidentified population may go unnoticed, which could negatively affect the overall genetic diversity of the species (Allendorf et al. 2008, Pinsky & Palumbi, 2014). A loss of genetic diversity within a gene pool can adversely affect the survival and adaptability of a species (McGinnity et al. 2009, Pinsky & Palumbi 2014). Therefore, the incorporation of genetic data into stock assessment is essential for the effective management and conservation of fishery resources.

1.2.3 Genetic methods used in stock assessment

Genetic approaches have become increasingly important in stock assessment and management in fisheries (Allendorf et al. 2008, Pinsky & Palumbi, 2014). These methods can provide valuable insights into the structure and dynamics of fish populations, which can inform the sustainable use of fishery resources (McGinnity et al. 2009, Pinsky & Palumbi, 2014). Some common genetic methods used in stock assessment include:

1. Genetic tagging: This approach involves the use of molecular markers, such as microsatellites or single nucleotide polymorphisms (SNPs), to identify individual fish and track their movements within and between different populations (Black et al. 2001, Campana et al. 1995). This can provide information on the degree of gene flow and connectivity between populations, as well as the potential impacts of fishing on population structure (Reusch et al. 2005, Schindler et al. 2010, Hellmair & Kinzinger, 2014).
2. Genetic population structure analysis: This method involves the use of molecular markers to identify the presence of genetically distinct subpopulations within a larger stock (Hewitt 1996, Laikre & Ryman 1996, Allendorf et al. 2008, 2013). This can be used to inform management decisions on harvest limits and quotas, as well as to detect shifts in population structure due to environmental or anthropogenic factors (McGinnity et al. 2009, Pinsky & Palumbi, 2014).
3. Genetic stock identification: This approach involves the use of genetic markers to identify the geographic origin of fish (Campana et al. 1999, Black et al. 2001), which can be used to inform management decisions on harvest limits, delineation of fishing boundaries, and quotas, as well as to detect illegal fishing or misreporting of catch data (Reusch et al. 2005, Schindler et al. 2010, Hellmair & Kinzinger, 2014).



University of Fort Hare
Together in Excellence

4. Genetic monitoring: This method involves the regular collection and analysis of genetic data to track changes in genetic diversity and population structure over time and inform management decisions (Hewitt 1996, Laikre & Ryman 1996, Allendorf et al. 2008, 2013).

Overall, the integration of genetic data in stock assessment and management can provide valuable insights into the structure and dynamics of fish populations, which can help ensure the sustainable use of fishery resources (McGinnity et al. 2009, Pinsky & Palumbi, 2014).

1.2.4 Selected genetic markers for this study

When selecting markers for a population genetics study, it is important to consider factors such as their informative power (Allendorf, 2008), cost and feasibility (Kalinowski et al. 2007), suitability for the specific type of study (Black et al. 2001), compatibility with other markers (Laikre & Ryman, 1996), and the availability of reference materials (McGinnity et al. 2009). The chosen markers should also be able to distinguish between different genotypes or alleles within the population (Campana et al. 1995), and be polymorphic in order to detect variation within the population (Reusch et al. 2005). It is also beneficial to use markers that can be easily compared with those used in similar studies (Allendorf et al. 2013), and to have reference materials available for comparison (Schindler et al. 2010). Historically, genetic markers such as allozyme electrophoresis, mitochondrial DNA sequencing, and microsatellite analysis were widely used in stock identification (Schlötterer 2004). In recent years, advances in genomics have introduced high-throughput approaches such as next-generation sequencing (NGS) (Piredda et al. 2022), single nucleotide polymorphism (SNP) genotyping (Wenne 2023), and whole-genome sequencing (WGS), which provide higher resolution and greater accuracy in stock discrimination. This study made use of microsatellite markers.

Microsatellites are simple short tandem repeats of genetic sequences, and they usually have a

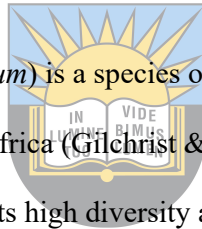
length of one to six base pairs, repeated up to 100 times at a locus (Tautz 1989). These markers have become popular in population genetics (Bamung et al. 2004, Schlötterer 2004). These genetic markers are highly polymorphic, have a high mutation rate, are distributed evenly throughout the genome, and are located in non-coding regions (Bamung et al. 2004, Schlötterer 2004), making them suitable for detecting discrete populations (Waples 1998). However, the development of microsatellites is expensive, and isolating the repeat units and designing primers are time-consuming (Mueller & Wolfenbarger 1999). Nonetheless, these markers are cost-effective once developed (Bhargava & Fuentes 2010).

1.3 The study species: The blue bream (*Pachymetopon aeneum*)

1.3.1 Taxonomy and description

The blue bream (*Pachymetopon aeneum*) is a species of fish (Figure 1.1) that is commonly caught in the line fisheries of South Africa (Gilchrist & Thompson 1908). It belongs to the family Sparidae, which is known for its high diversity and includes around 110 species in 33 genera (Smith 1938, Akazaki 1962, Smith & Smith 1986, Orrell & Carpenter 2004).

Pachymetopon aeneum is currently listed as a species of least concern by IUCN (2025). The Sparidae family plays a significant role in the South African line fisheries sector (Nelson 2006).



University of Fort Hare
Together in Excellence



Figure 1.1: An adult individual of the blue bream, *Pachymetopon aeneum* (Source: www.aquarium.co.za).

1.3.2 Distribution

The blue bream is a fish species native to South African waters (Heemstra & Heemstra 2004), with a main distribution extending from False Bay in the WC to Sodwana in KZN (Figure 1.2). The blue bream is most commonly found in the EC and southern KZN (Garratt 1988, Heemstra & Heemstra 2004).



University of Fort Hare
Together in Excellence

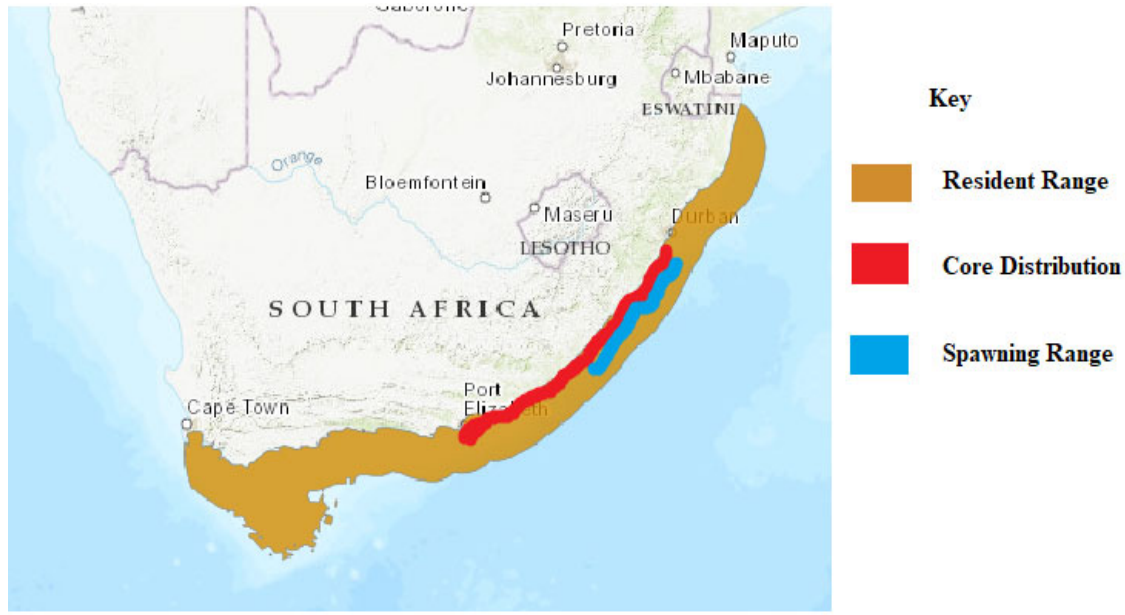


Figure 1.2: Distribution of the blue bream (*Pachymetopon aeneum*) in South Africa. The core distribution is indicated by the red line, the main spawning range is indicated by the blue line, and the full geographical range in South Africa is indicated by the brown line. (Source: IUCN 2020).



University of Fort Hare
Together in Excellence

1.3.3 Habitat

The blue bream is a benthic species, adults of the species are commonly found along inshore and offshore reefs, at depths ranging from 20 to 80 meters (Buxton & Clarke 1984, 1986).

Juvenile *P. aeneum* tend to inhabit shallow subtidal reefs, at depths of around 10 to 20 meters (Buxton & Clarke 1984, 1986, Fennessy 2013). The eggs of the *P. aeneum* are pelagic, meaning that they are found in the open waters of the ocean rather than near the bottom.

These eggs have been observed in shelf waters off Park Rynie in KZN (Connell 2012) and in the Tsitsikamma Marine Protected Area (MPA) in the EC (Fennessy 2013).

1.3.4 Diet

The blue bream is a carnivorous benthic-feeding species that primarily feeds on hydrozoans, tunicates, octocorals and mysids (Buxton & Clarke 1986). Its dentition, characterized by rows of uniform cutting teeth on both the upper and lower jaws, is well adapted for this type of feeding behaviour. The gape, or widening of the jaw, is small in *P. aeneum*, allowing it to ingest its prey in small pieces rather than larger chunks (Buxton & Clarke 1986). Both adult and juvenile *P. aeneum* exhibit similar dietary habits (Buxton & Clarke 1986).

1.3.5 Reproduction

In most vertebrates, sexual dimorphism is the norm, where sexes of the same species exhibit distinct characteristics and sex change is considered atypical (Chan & Yeung 1983).

However, in the family Sparidae, sex change is a commonly-observed phenomenon, with hermaphroditism being prevalent and functional (Buxton & Garratt 1990). Buxton and

Garratt (1990) modified Garratt's (1986) classification of reproduction behaviour in sparids into five categories: protogynous hermaphrodite (sex change from female to male),

protandrous hermaphrodite (male to female), simultaneous hermaphrodite (possessing both male and female sex organs), rudimentary hermaphrodite (having non-functional bisexual

ovotestis) and gonochorist (separate sexes). *Pachymetopon aeneum* is a protogynous

hermaphrodite, with sex change occurring between 200-350 millimeters (mm) fork length (FL), which translates to four to nine years, at the point when females reach sexual maturity

(Buxton & Clarke 1986). The peak spawning season for the *P. aeneum* was observed between September and March (summer) in the EC region and between August and November

(spring) in the KZN region. However, reproductively-active fish were caught throughout the year in both regions (Buxton & Clarke 1986). The northern EC and southern KZN region

appears to be the primary spawning area for the *P. aeneum* (Figure 1.2), as most cases of ripe fish were recorded in this region (Garratt 1988).

1.3.6 Movement

The movement patterns of many coastal fishery species, including *P. aeneum*, are not well understood. This lack of information was also highlighted by Mann (2013) and Fennessy (2013) who called for future studies to focus on the movement patterns of *P. aeneum*. As mentioned in Mann (2013) Oceanographic Research Institute (ORI) tagging project (unpublished data) conducted a study to investigate the movement patterns of the *P. aeneum*. The study included 171 tagged individuals, out of which only four were recovered, and none of these recaptures showed any significant movement from area of capture.

On the other hand, *P. aeneum* has been reported to undertake northward and eastward spawning migrations during summer and spring to form spawning aggregations in southern KZN and the northern EC (Garratt 1998). However, this movement pattern appears to be variable for this species, as the size of the fish migrating varies yearly (Garratt 1988). This could be because *P. aeneum* exhibits a limited migration reproductive strategy, where only a portion of the reproductively active stock migrates (Garratt 1988).

1.3.7 Management

Pachymetopon aeneum is a commercially-exploited species that is increasingly becoming an important component of the offshore commercial catches off the EC and KZN regions of South Africa (Smale & Buxton 1985, Garratt 1988, Brouwer & Buxton 2002, Fennessy et al. 2003, Donovan 2010, Dunlop & Mann 2013). According to studies, the maximum recorded age, weight and length for *P. aeneum* are 12 years (Buxton & Clarke 1986), 4.75 kilograms

(kg) (SADSAA 2012) and 60 centimetres (cm) total length (TL) (van der Elst 1993, Mann 2013), respectively.

A stable annual commercial catch of 10 tonnes per year was reported from 1985 to 2008 (Mann 2013). Targeted *P. aeneum* catch per unit effort (CPUE) has shown a significant increase between 1985 and 2007 in the KZN commercial ski boat fishery (Fennessy 2013). Two independent surveys of the KZN commercial line fishery have also shown an increase in CPUE for *P. aeneum*, from 0.43 kg/fisher/hr (1994 – 1996) to 1.03 kg/fisher/hr (2008 – 2009) (Mann et al. 1997a, Dunlop 2011). [Any new additional information about the updated stock status from the new section you have added above could go in here]

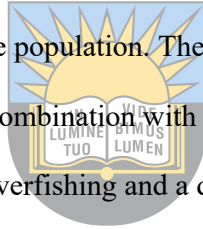
However, the percentage contribution of *P. aeneum* catches to the total linefish catch has remained relatively stable in the KZN commercial line fishery, at 2.95% for the period 1994 – 1996 and 2.6% for the period 2008 – 2009 (Mann et al. 1997a, Dunlop 2011). The Port Alfred (PA) commercial and recreational line fishery, on the other hand, has shown an increase over the period 1985 – 2008 (Donovan 2010). This increase in percentage contribution by mass in the PA line fishery sector was accompanied by a decreasing trend in mean size over the same period (Donovan 2010). This suggests overfishing of a particular size class, which could lead to a skewed sex ratio and early sex change, preventing the species from reaching a larger, desirable size (Donovan 2010).

Currently, the management measures for the *P. aeneum* include a daily bag limit of five fish per person per day for recreational fishers and unlimited catches for commercial fishers (DAFF 2012). This is of concern as commercial fishers have access to more sophisticated technology, which allows them to locate fish with ease, potentially contributing to localised overfishing (Fennessy 2013). Moreover, there are no other forms of regulation in place for

the management of this species, there is no closed season or minimum size limit (Fennessy 2013).

However, the species occurs throughout several no-take MPAs across its distributional range (such as the De Hoop Marine Reserve, Goukamma, Tsitsikamma National Park, Sardinia Bay Reserve, Bird Island, Dwesa-Cwebe and Pondoland Protected Areas), where it is receiving stringent enforcement of no-take regulations, preventing all forms of extractive activities like fishing and mining (Sale et al. 2005, Mann 2013). MPAs have been proven to be vital tools for the management of coastal fisheries and the protection of fish stocks from collapse (Sale et al. 2005).

Given the current management measures in place for *P. aeneum*, it is crucial to consider the potential impacts of overfishing on the population. The lack of regulations, such as closed seasons and minimum size limits, in combination with advanced fishing technology available to commercial fishers, could lead to overfishing and a decline in population numbers (Fennessy 2013).



University of Fort Hare
Together in Excellence

1.4 Rationale

A stock assessment is crucial for the effective management of the *P. aeneum* population. This assessment is considered a major information gap that needs to be filled (Mann 2013). The South African coastline is composed of different marine biogeographic provinces, including the tropical-subtropical KZN, warm temperate EC, and cool temperate WC (Teske 2006). Many species living in these biogeographic provinces exhibit genetic structuring due to adaptation to the specific environmental conditions of a particular biogeographic province (Teske 2010). However, not all species that reside in more than one marine biogeographic province exhibit this genetic structure (Teske 2006, Neethling 2008, Teske 2010).

Pachymetopon aeneum is found in all three of South Africa's marine biogeographic provinces, yet no genetic stock assessment has been conducted for this species.

Without a genetic stock assessment, it's challenging to understand the makeup of *P. aeneum* populations. This lack of information can hinder effective conservation and management strategies, as recognizing and preserving genetically-distinct subpopulations is crucial for maintaining diversity and ensuring the species' long-term survival (Mann 2013). The *P. aeneum* population is currently managed as a single stock throughout its distribution with no input from genetic studies. This approach to management may be compromised if genetic structuring exists within the *P. aeneum* population and this structure is associated with fishing grounds (Carvalho & Hauser 1995). Furthermore, Carvalho and Hauser (1995) suggest that, if present, structured and isolated units might respond negatively to harvesting, which is a concern as it can subsequently lead to genetic diversity loss, a rise in inbreeding in stocks, regional extinction, and socio-economic problems if the species cannot be fished anymore (Dudu 2015). To address this, molecular markers should be used to obtain more detailed information on the stock structure of the population (Carvalho & Hauser 1995).

The blue bream is a species that is characterized by its slow growth rate, late maturity, and protogynous hermaphroditism, which are traits that make it vulnerable to overexploitation (Buxton & Clarke 1986, Rhett 2012, Pinsky & Palumbi 2014). Overexploitation of this species is predicted to result in genetic changes within the population (Pinsky & Palumbi 2014), which can increase the risk of extinction and reduce the population's ability to recover (Walsh 2006).

One potential mechanism by which overfishing can affect the genetic makeup of a population is through size-selective fishing. This practice, which targets larger reproductively-contributing individuals, can result in reduced reproductive success and increased genetic

structuring (Buxton 1992, Griffiths 2000). However, the effects of overfishing on genetic changes are not limited to size-selective fishing, as increased mortality across all size classes can also lead to early maturation (Law 2007).

Additionally, overexploitation can lead to a decrease in effective population size (N_e), which can further affect the genetic diversity of a population. This can be observed in terms of reduced average heterozygosity at neutral loci (Allendorf 2014). In summary, *P. aeneum*'s life history traits and vulnerability to overexploitation make it a species of concern, as it is at risk of genetic changes associated with population decline and reduced recovery potential.

The blue breem is a species that demonstrates age-related sex change from female to male (Buxton & Clarke 1986), making it an excellent study species for examining genetic stock structuring. This is due to the presence of hermaphroditism in the species, which is thought to be a key factor in shaping genetic structuring (Chopelet 2009).

The movement patterns of this species are not well understood (Fennessy 2013), which poses a challenge in determining the stock structure of the *P. aeneum*. Movement patterns are often reliable indicators of demographic, spatial and genetic population structure (Pittman & McAlpine 2001). Highly-migratory species typically display little genetic structuring (Avisé 1998, Nesbø 2000, Waples 1987, Zardoya 2004, Gonzalez 2008). However, as *P. aeneum* belongs to the Sparidae family, which exhibits extreme residency during their lives, it is expected to be genetically structured (Buxton & Allen 1989, Cowley 2002, Kerwath 2007a).

On the other hand, seasonal catch data from southern KZN and northern EC suggest that *P. aeneum* likely undertakes northward and southward spawning migrations (Garratt 1998).

Therefore, when the species utilizes a specific site for spawning, it is likely that ocean currents will transport the pelagic larvae, which should be the main driver of genetic connectivity (Choat 2012, Colin 2012a, Coney 2014).

This suggests that, despite the potential for genetic structuring within *P. aeneum* due to the typical extreme residency exhibited by the species belonging to the Sparidae family, the *P. aeneum* population may exist as a single mixed panmictic stock, a result of seasonal spawning migrations undertaken by *P. aeneum* individuals.

Pachymetopon aeneum displays conflicting indications in its life history, demographic and reproductive biology, raising the possibility that it may exist either as a multi-stock species with distinct populations within a larger metapopulation or as a single panmictic stock.

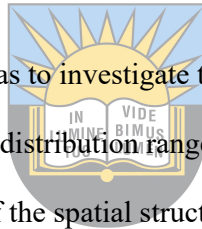
Therefore, a genetic stock assessment was conducted in order to aid in stock delimitation and the implementation of effective management measures (NOAA 2009).

1.5 Aims and objectives of this study

The overarching aim of this project was to investigate the stock structure of *P. aeneum* from different ecological regions within its distribution range, using microsatellite markers, in order to improve our understanding of the spatial structuring, genetic diversity, and the underlying factors affecting the genetic structuring of the *P. aeneum* population in South Africa.

The objectives of this project were to:

- Assess the extent and distribution of genetic diversity of the *P. aeneum* across its range by utilizing nuclear microsatellite genotypic data.
- Identify levels of differentiation within the *P. aeneum*, and identify any potential discrete stocks.



University of Fort Hare
Together in Excellence

Chapter 2

Materials and methods

2.1 Sampling

Permits [under the Department of Agriculture, Forestry, and Fisheries (DAFF) mandate] and ethical clearance (GOU011SYOK01) were obtained before the start of the project. A total of 173 fin clips of *P. aeneum* were collected between March 2009 and October 2015. Sampling occurred during the annual research surveys of DAFF. Other fin clip samples were obtained from observers on fishing vessels, at fishing or angling competitions, and from the Oceanographic Research Institute (ORI) tagging programme. The collected fin clips samples of *P. aeneum* were preserved in 99% ethanol and stored in individually-labelled vials. For each sample, the location, date of capture, fish length (total length and/or fork length), weight, sex and reproductive stage were recorded.

The samples were collected from 27 locations, spanning the full distribution range of *P. aeneum* from northern KZN to the Agulhas Offshore region. In order to increase sample size and improve representativeness, neighbouring sampling locations were grouped based on proximity and biogeographic zones, resulting in seven geographic regions used in analyses: Agulhas Inshore (AI), Agulhas Offshore (AO), Port Alfred (PA), East London (EL), Transkei (TKI), South KZN (S KZN) and North KZN (N KZN) (Figure 2.1). This sampling strategy ensured that a minimum of 19 samples were collected per region, providing a robust dataset for analysis.

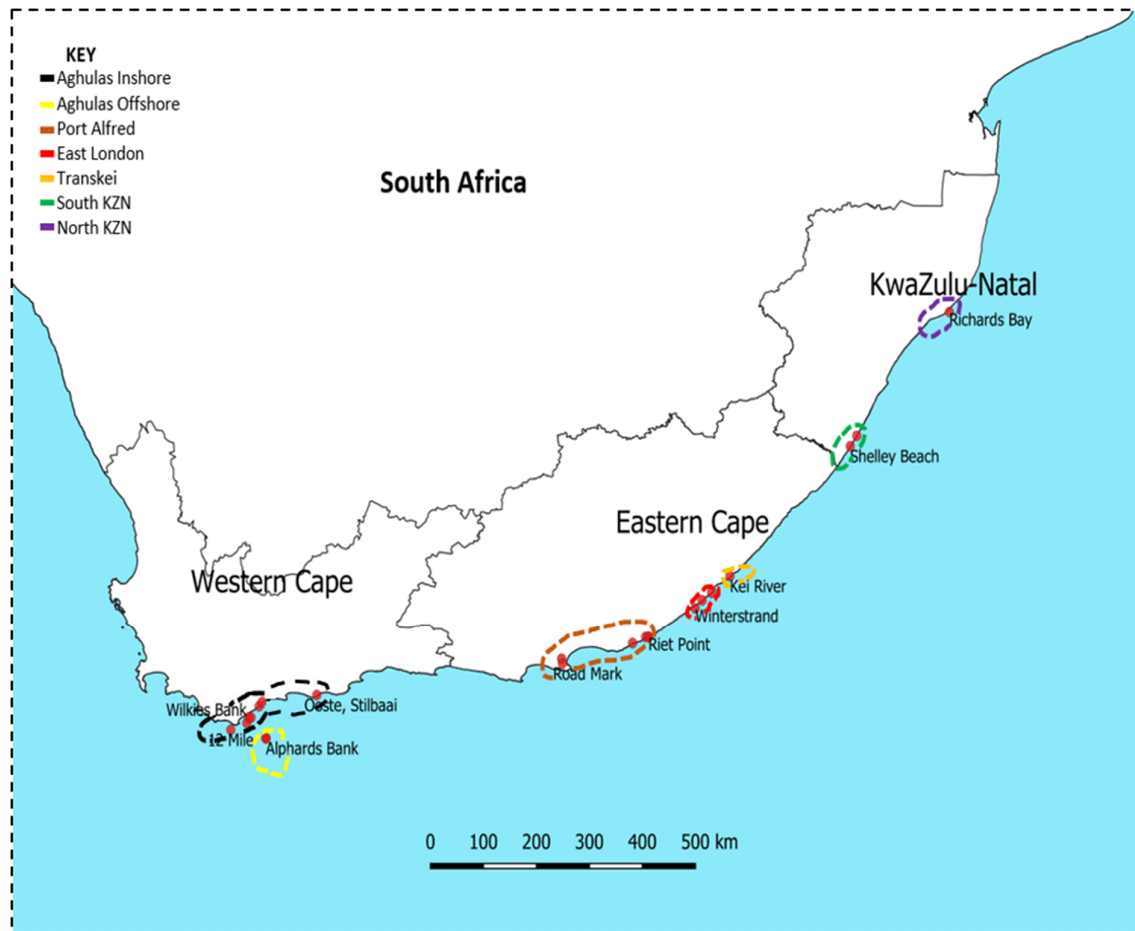


Figure 2.1: Map of South Africa depicting the locations where *Pachymetopon aeneum* samples were collected. The sampled locations were grouped together based on proximity and biogeographic zones, and these seven sampling regions are denoted by broken lines and represented in the key

2.2 DNA extraction

Genomic DNA was extracted from each sample, using a small piece of the fin clip (approximately 5 mm in length). DNA extraction was carried out using the “salting out” method (Sunnucks & Hales 1996). The extracted DNA was quantified and its quality was assessed using a Nanodrop 2000c spectrophotometer. The extracted DNA was then stored at -4°C for future use in generating microsatellite data.

2.3 Microsatellites data processing

2.3.1 PCR amplification and genotyping

Reid et al. (2012) developed microsatellite markers for two Sparidae species (*Pachymetopon blochii* and *Lithognathus lithognathus*) and then tested the transferability of these markers to other sparid species. Eight tetranucleotide markers (with their respective primer sequences), were chosen for the current study (Table 2.1) and incorporated into two multiplex PCR reactions, taking into account the fluorescent dye labels and allelic size ranges. The selection of markers was based on the results of cross-species amplification and polymorphism reported by Reid et al. (2012) for these markers in other sparid species.

The microsatellite loci were amplified for each sample by PCR, using fluorescently-labelled forward and unlabelled reverse primers. The labelled fluorescent dyes aid in the separation of the microsatellite markers, as the alleles of each marker are distinguished by their designated fluorescent dye (colour) and size ranges (Table 2.1). Furthermore, the use of fluorescently-dyed primers allows the amplification of the many microsatellite markers in a single PCR multiplex reaction, significantly saving time and expense (Neff et al. 2000).



University of Fort Hare
Together in Excellence

Table 2.1: Microsatellite markers developed by Reid et al. (2012) and used in the current study of genetic diversity and structure in the blue bream (*Pachymetopon auneum*) along the South African coast, with selected microsatellite markers, along with their respective dye labels and allele size ranges.

Multiplex	Marker	Fluorescent Dyes	Allele size ranges	Primer Sequence
1	PBt007	6FAM (Blue)	135 – 380	F: 5'-CGCAGGATTCATACGTAAAGAAAT-3', R: 5'-TTCCATTCCATGCCCTGCC-3'
1	PBt018	VIC (Green)	120 – 285	F: 5'-GTCAAGCCTGTCAAACCTGGG-3', R: 5'-TGGAAAGAAACACGAGCCG-3'
1	PBt-OVI-D106	VIC (Green)	320 – 455	F: 5'-ATGAGCGATGAAGATGGAGA-3', R: 5'-GCCCTCAGCTACAGTTGAGA-3'
2	PBt011	NED (Yellow)	117 – 320	F: 5'-CACAGACGGGAATGCACAC-3', R: 5'-AGGGTGTTCAGGACAGAACG-3'
2	PBt001	PET (Red)	290 – 390	F: 5'-AGGTTCCCCACAGAAGGTC-3', R: 5'-AGTACCTGGGAAACAGCCC-3'
2	PBt003	PET (Red)	115 – 280	F: 5'-TTGGGGAAAAGGAGAGGCG-3', R: 5'-CCGACTCGCTCTGTATGTTG-3'
2	Pbt014	VIC (Green)	125 – 285	F: 5'-GTTGGGTGATGGAGCCAAG-3', R: 5'-CAGAGGTGACTAGAAGCCCC-3'
2	LLt006	VIC (Green)	291 – 440	F: 5'-TCATCTGACCACAGACCAGC-3', R: 5'-CACGCCATTTGTGACAGTAAG-3'

In this study, the PCR reactions were performed in a 96-well plate format, with each reaction having a volume of 25 μ L. To ensure the specificity of the amplification, 2 μ L of high-quality (concentration between 50ng/ μ L and 100ng/ μ L, purity ratios $\sim$ 1.8 for 260/280 and between 2.0 -2.2 for 260/230) DNA template was added to each reaction; along with a 0.4 μ M multiplex primer master mix. MyTaq HS (BioLine) DNA polymerase was used as the enzyme for the PCR reactions, and deionized water was used to make up the remainder of the reaction volume. A negative control PCR reaction was performed as well, along with the runs containing a template.

The PCR reactions were performed in an Applied Biosystems Veriti 96 well Thermal Cycler, which provides precise temperature control and consistent results. The PCR cycling conditions included an initial denaturing step at 94° C for 2 min, followed by 30 cycles of denaturing at 95° C for 1 min, annealing at 52° C for 45 s, and extension at 72° C for 1 min. The final extension step was at 72° C for 7 min. These conditions were carefully optimized to ensure the specificity and efficiency of the amplification.



University of Fort Hare
Together in Excellence

After the PCR amplification, the amplified fragments were combined with ABI Hi-Di Formamide and an internal lane standard (GS600LIZ) to generate electrophoretic profiles on an ABI-Hitachi Genetic Analyser 3500 at the NRF-SAIAB Aquatic Genomics Research Platform.

2.3.2 Allele calling and genotyping

GeneMapper software version 5.0 was used to score the microsatellite data. The software was used to size the DNA fragments, by comparing the fragment to a set of known size standards. In this study, an initial automated binning step was done, then followed by manual verification and adjustment of bins that were outside the expected motif, to ensure accuracy and consistency in allele calling. Calling alleles refers to the process of identifying the

specific variant (a DNA fragment of a particular size) of a genetic marker present in a sample, while genotyping refers to the process of determining the combination of alleles present at a locus in an individual sample.

MICRO-CHECKER v2.2.3 (Van Oosterhout et al. 2004) was used to detect potential genotyping errors, allele dropout, and allele stuttering, which can arise during PCR amplification. Allele dropout occurs when one allele in a heterozygous pair is not detected, and allele stuttering refers to an amplification artifact where additional alleles appear at smaller sizes than the true allele, leading to overestimation of genetic variation (Van Oosterhout et al. 2004). While MICRO-CHECKER was employed to identify these genotyping errors, no corrections or calculations of null allele frequencies were performed or reported in this study.



2.3.3 Microsatellite Analysis

2.3.4 Hardy Weinberg Equilibrium (HWE)

University of Fort Hare
Together in Excellence

Hardy-Weinberg Equilibrium (HWE) is based on the assumptions of random mating, no migration, no mutation and no selection within a population of infinite size (Hardy 1908, Weinberg 1908, Weir 1996). This theorem predicts the expected frequency of homozygous and heterozygous variants based on the allele frequency (Weir 1996). If the observed frequencies deviate from the expected values, it may indicate a violation of one or more of these assumptions (Weir 1996). However, it is important to note that departure from HWE can also be a result of genotyping errors and not necessarily a violation of the assumptions (Weir 1996).

To test for deviations from HWE, an exact test was performed using Arlequin 3.5 (Raymond & Rousset 1994). The exact test is considered advantageous because it is accurate, unbiased,

works well with small sample sizes and is not affected by ploidy level (Raymond & Rousset 1994). The test was performed using a Markov-chain method, which included 100,000 steps and 10,000 dememorization steps (Guo & Thompson 1992). This method is robust and efficient, making it suitable for testing HWE in large datasets (Guo & Thompson 1992).

2.3.5 Linkage disequilibrium

Linkage disequilibrium (LD) is a measure of the non-random association between genetic markers in a population (Hedrick 2005). It tells us whether pairs of genetic markers or loci show a closer association than expected, indicating that these are not segregating independently. The independence or correlation of expected allele frequencies at different loci is crucial for interpreting genetic variation within and between populations (Weir & Hill 2002). In population genetics, LD is an important concept as it can also provide insight into various population processes, such as population bottlenecks, evolutionary events and demographic processes that have shaped a population (Slatkin 2008).

Tests of genotypic linkage disequilibrium were carried out in Arlequin 3.5 to determine if there is a non-random association of alleles at all eight loci (Slatkin 2008). The test, conducted between pairs of loci across all the regions, used 10 000 steps in the Markov chain. Expectation algorithm (EM) initial conditions were set to start at 5, this algorithm searches for the maximum likelihood through a repetitive process of refinement from a set number of random initial conditions. Because this analysis applied multiple statistical tests, significance levels were corrected using standard Bonferroni corrections (Rice 1989).

2.3.6 Genetic diversity

Genetic diversity is a crucial aspect of population genetics because it illuminates a population's potential for evolution and potential risks from overexploitation (Hughes et al.

2008). Genetic diversity can be assessed using a number of indices, such the number of alleles, allelic richness and heterozygosity. In this study, these indices were used to gain a comprehensive understanding of the genetic diversity within and between populations.

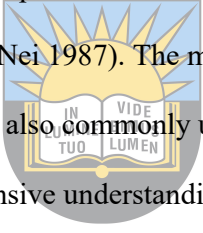
Allelic richness is the number of alleles corrected for sample size (Hughes et al. 2008) and is an important measure of genetic diversity. The correction ensures unbiased comparisons of genetic diversity between populations that are not represented by the same number of samples. Allelic richness has been demonstrated to be an important evolutionary indicator (Allendorf 1986, Allendorf et al. 2012, Caballero & Garcia-Dorado 2013), shedding light on a population's capacity to survive or adapt to environmental changes (Fisher 1930), and can help guide population management and conservation strategies (Petit et al. 1998; Barker 2001, Simianer 2005). Allelic richness (AR) values vary depending on the dataset and the number of alleles observed per locus within the studied population, calculated as the total number of different alleles across all loci divided by the number of loci examined.

Heterozygosity is another important measure of genetic diversity. It measures the proportion of individuals that possess different alleles at a given locus. Expected heterozygosity (H_E) and observed heterozygosity (H_O) are commonly used measures of heterozygosity. Observed heterozygosity has been shown to be correlated to evolutionary fitness (Reed & Frankham 2003, Szulkin et al. 2010), making heterozygosity an important measure for ecological inferences. H_E is considered a better estimator to assess genetic variability in a population (Nei 1987, Toro et al. 2009) because it is more sensitive to allele frequency than to the number of alleles (Nei 1987).

Observed heterozygosity (H_O) and expected heterozygosity (H_E) value range from 0 to 1. Higher H_O values indicate a higher proportion of heterozygous individuals within the population, suggesting greater genetic diversity. Higher values of H_E indicate greater genetic

diversity within the population, as there is a higher probability of randomly sampling different alleles from the gene pool. While lower values indicate the opposite.

Wright's inbreeding coefficient (F_{IS}), is also an important measure of genetic diversity. This quantifies the deviation of observed heterozygosity from expected heterozygosity under HWE (Weir & Hill 2002, Reed & Frankham 2003). The inbreeding coefficient (F_{IS}) ranges from -1 to 1. Negative values indicate heterozygote excess, while positive values of F_{IS} indicate heterozygote deficit, which can occur due to inbreeding, resulting in fewer heterozygotes than expected. This deficit reflects reduced genetic diversity within the population and can have negative effects on fitness (Reed & Frankham 2003).

The number of alleles per locus is an important measure of genetic diversity, reflecting the genetic variation within a population (Nei 1987). The mean number of alleles, which is the average number of alleles per locus, is also commonly used in conjunction with the number of alleles to provide a more comprehensive understanding of the population's genetic diversity (Weir & Hill 2002).  Polymorphism information content (PIC) (Kalinowski et al. 2007).

In this study, Arlequin 3.5 (Excoffier & Lischer 2010, 2011) was used to calculate all indices of genetic diversity, except for mean allelic richness, which was calculated in Fstat v2.9.3 (Goudet 1995, 2002). The data were analysed for each region and for the entire population.

2.3.7 Population differentiation

Pairwise population comparisons were also conducted using Slatkin's (1995) R_{ST} , which is analogous to F_{ST} but assumes a stepwise mutation model that is more appropriate for microsatellite data. R_{ST} considers the stepwise accumulation of mutations, making it useful for detecting differentiation influenced by genetic drift and mutation. This approach provides

insight into the genetic structure of populations by accounting for the mutational processes that shape microsatellite variation. Analyses were conducted in Arlequin version 3.5.

2.3.8 Analysis of Molecular Variance (AMOVA)

To effectively assess the population structure of a species, Analysis of Molecular Variance (AMOVA) serves as a valuable tool. This statistical method, introduced by Excoffier et al. (1992), allows for the evaluation of genetic variability at different hierarchical levels, ranging from among groups, within groups, within populations, and within individuals.

Among-group variance reflects the genetic differentiation between predefined groups of populations. In this study, samples were organized into three distinct groups based on their biogeographical provinces, which are influenced by climate, ecological differences and oceanographic conditions. Each biogeographical province consists of multiple sampling localities that were grouped based on their geographic proximity and shared environmental conditions. The KZN group includes both North KZN and South KZN, representing tropical and subtropical regions, respectively, while the EC group (TKI, EL, and PA) falls within a warm temperate region. The WC group (AI and AO) is situated in a cool temperate region. These distinct ecological zones, characterized by varying climate conditions such as temperature and humidity, are likely to influence genetic differentiation. These environmental variation is expected to shape genetic diversity within populations, with each region potentially harbouring unique genetic traits adapted to its specific ecological conditions. The within-groups consideration of variance (F_{SC}) examines genetic variation among populations within each predefined group, revealing the diversity present among populations making up the same group. The among-individuals component of variance assesses genetic differences among individuals within populations, providing insight into patterns and levels of diversity among loci. Finally, the within-individuals component of variance represents the genetic

variation within individual organisms, helping to quantify the genetic diversity of individual organisms. AMOVA examines these components as percentages of the total molecular variance, this can help explain how genetic variation is partitioned at different hierarchical levels within the population, integrating all the F-statistics into a cohesive understanding of population genetic structure. Arlequin was used to assess AMOVA. To test the significance of the variance components and fixation indices, 10,000 permutations were conducted. The permutation test is a robust method for evaluating the statistical significance of the observed results, as it allows for the estimation of the probability of observing the same or more extreme results by chance.

2.3.9 Principal Coordinate Analysis (PCoA).

Principal Coordinate Analysis (PCoA) is a multivariate statistical method used in biology and ecology to reduce the dimensionality of complex data sets (Peakall & Smouse 2006). PCoA is frequently employed to visualize patterns and relationships, identify clustering or grouping of samples, and capture the majority of the variance, thus making the data more interpretable (Gower 1966, Peres-Neto et al. 2006). For population geneticists, PCoA is valuable as it helps interpret and visualize complex genetic data, providing insights into the evolutionary history and relationships of populations (Excoffier et al. 2009).

PCoA considers pairwise genetic distances, such as genetic dissimilarities or F_{ST} values, calculated from genetic markers like microsatellites. The resulting plot displays the spatial arrangement of individuals or populations in reduced-dimensional space, with each point representing an individual. Clustering patterns in the PCoA plot indicate genetic similarity or divergence: tight clusters suggest high genetic similarity, while dispersed points indicate genetic divergence. Outliers may reveal individuals or populations with unique genetic

compositions, potential admixture, or genetic isolation. In this study, PCoA was carried out using Nei's distance in GenAlEx (Peakall & Smouse 2006).

2.3.10 Structure

Inferring population genetic structure holds importance in ecology, evolution, and conservation biology. Among the range of tools available, STRUCTURE (Pritchard et al. 2000) stands out as the most widely employed and referenced software for detecting population genetic structure. In order to identify population structure from the observed multilocus microsatellite genotypes in the global sample, a Bayesian clustering method was employed in the software STRUCTURE (Pritchard et al. 2000). The algorithm uses a Markov chain Monte Carlo (MCMC) method to assign individuals showing similar genotypes to one or more hypothetical discrete putative populations (K) that are defined by the user. The method optimises the assignment by minimising deviations from HWE within each group, thus identifying the most likely number of distinct populations. For this study, the number of putative populations (K) was tested from $K = 1$, representing a single well-mixed population, to $K = 8$, as literature review suggested potential substructure or genetic differentiation within the sampled populations beyond what is immediately apparent from the number of sampled sites.

For each value of K, 20 iterations were conducted, from which the mean log probability that the observed set of genotypes in the global sample was drawn from the K populations was estimated. STRUCTURE analyses were conducted using the admixture model with correlated allele frequencies, which is appropriate for detecting subtle population structure and allowing for gene flow between populations. LOCPRIOR was set to 1 to use geographic or environmental information to assign individuals to populations based on their location. Each iteration consisted of 50 000 burn-in steps and an additional 50 000 steps in the MCMC.

Once the optimal K value was determined, the results from each of the 20 iterations for the chosen K were synthesized to produce the final population assignments for each individual. A bar plot was generated to visualise the population assignments for each individual and the proportion of each individual's genetic makeup from each putative population. These results were then used to further infer population structure of the blue breem.

STRUCTURE works well with even sample sizes but struggles with uneven samples or hierarchical structures (Puechmaille 2016). The ΔK method by Evanno et al. (2005) often underestimates true structure by focusing on the highest hierarchical level. Therefore, to address this, this study used StructureSelector (Li & Liu 2018) with four more accurate estimators (MedMedK, MedMeaK, MaxMedK, and MaxMeaK) that count clusters in at least one subpopulation (Puechmaille, 2016). These methods work better with both even and uneven samples. Along with the new estimators, the traditional $\ln \Pr(X|K)$ and ΔK were calculated. The most probable number of populations (K) was selected based on the value of K that maximizes the log probability (Falush et al. 2003) and accounts for the most variability in the data (Pritchard et al. 2010); this was done using StructureSelector.

2.3.11 Isolation-by-distance

Isolation-by-distance (IBD) is a fundamental concept in population genetics, describing how genetic differentiation among populations increases with greater geographic distance between them (Wright 1943). This phenomenon arises due to the limited dispersal distance of individuals, leading to reduced gene flow between distant populations (Waples 1998). In marine species with wide distributions and high levels of gene flow, population structure can be formed due to limited dispersal distance rather than physical barriers (Wright et al. 2015).

IBD was investigated by examining how genetic differentiation among populations correlate with the geographic distances separating them. Geographic distances were calculated using

Euclidean distance between the central points of each pooled region, where each region comprises multiple localities. Arlequin 3.5 was used to perform a Mantel (1967) test to determine if there was a correlation between a matrix of pairwise genetic distances among the sampling regions and the matrix of geographic distances between them. A matrix of pairwise F_{ST} values was used, then adjusted to Slatkin's linear genetic distance ($F_{ST}/(1-F_{ST})$), and the negative F_{ST} values were changed to zero (Slatkin 1995). IBD was performed in GenAlEx, to test for a significant correlation with the matrix values used in 9 999 permutations. This approach allowed us to determine if any observed genetic differentiation among populations was due to IBD or other factors such as natural selection or genetic drift.

2.3.12 Relatedness

The estimation of relatedness among individuals within a population is integral to population genetics research (Wright 1921, 1922). This approach is applied across various research areas, including population structure and gene flow, and holds a particular relevance for a species like *P. aeneum*, sampled from different locations (Weir et al. 2006). By genotyping individuals from these locations, relatedness analysis allows for the assessment of genetic similarity between individuals from the same or different locations (Weir et al. 2006).

Relatedness is rooted in the principle that genetically similar individuals share more similar genotypes at a locus, indicative of a higher probability of identity by descent (Lynch and Walsh 1998). This shared genetic ancestry leads to a common genealogical history among related individuals (Wang 2017). Estimating relatedness enables researchers to discern the extent of genetic similarity among *P. aeneum* individuals and its relationship with their spatial distribution.

To quantify relatedness, researchers utilize measures such as pairwise genetic distance or the coefficient of relatedness, which gauge the degree of genetic similarity between individuals.

Higher relatedness values signify greater genetic similarity, implying closer genetic relationships. Significance testing helps in determining if observed relatedness values significantly deviate from random expectations, suggesting that genetic similarity is not merely due to chance.

To perform this analysis, the software GenAlEx version 6.5 (Peakall & Smouse 2006) was employed. This software uses three estimator methods, those developed by Queller and Goodnight (1989), Ritland (1996), and Lynch and Ritland (1999).

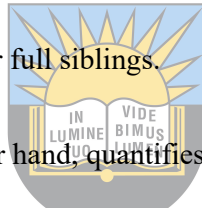
The Queller and Goodnight (1989) index measures relatedness by assessing the proportion of alleles shared identical by descent. It ranges from -1 to 1, where -1 signifies individuals are completely unrelated, 0 indicates no more relatedness than expected by chance, and 1 implies individuals are genetically identical or full siblings.

The Ritland (1996) index, on the other hand, quantifies genetic relationships based on allele sharing probabilities. It ranges from 0 to 1, with 0 indicating no genetic relationship (unrelated) and 1 denoting individuals are genetically identical or full siblings.

Lastly, the Lynch and Ritland (1999) index is a modification of the Ritland index, specifically designed for diploid organisms with incomplete pedigree information. Similar to the Ritland index, it also ranges from 0 to 1, with interpretations akin to the Ritland index.

These estimator methods enable the calculation of pairwise comparisons among all individuals from the global sample, providing insights into the relatedness and genetic relationships within the studied populations.

The default max setting for all three methods is 0.5 (Peakall & Smouse 2006). However, as per Lynch and Ritland (1999), the check x2 option must be selected to give a max of 1 (Peakall & Smouse 2006). These estimators have been shown to be unbiased, accurate and

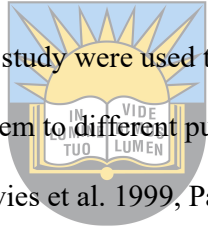


University of Fort Hare
Together in Excellence

consistent under the assumption that there are no known errors in allele frequency in a suitably-defined population (Lynch & Ritland 1999, Van de Castele et al. 2001, Wang 2002, 2011b). The mean relatedness estimates were calculated within each region and were statistically compared to overall relatedness within the global sample, using 9,999 permutations (Peakall & Smouse 2006).

2.3.13 Population assignment

In this study, population assignment techniques were used to analyse current dispersal patterns. According to Berry et al. (2004), these techniques are frequently used in population genetics studies because they offer a more accurate and efficient method of estimating dispersal patterns than do conventional direct observation techniques. The population assignment methods employed in this study were used to identify the source population of each sampled individual and assign them to different putative populations simultaneously (Waser 1998, Cornuet et al. 1999, Davies et al. 1999, Patkeau et al. 2004, Field et al. 2017).



GenAIEx was used to carry out the population assignment analysis. The analysis aimed to detect any isolated populations within the study population. Two different approaches were employed: a “leave one out” approach and an “as is” approach. The “leave one out” approach was used to correct for potential biases (Waser & Strobeck 1998) by excluding the individual’s genotype from the allele frequency estimation of its original sampled location and then estimating allele frequency for the calculation of the expected frequency for that specific genotype (Peakall & Smouse 2009). The “as is” approach, on the other hand, used a simpler method of calculating allele frequency by using the original observed allele frequency of the individuals.

Chapter 3

Results

A total of 173 *P. aeneum* individuals were sampled and successfully genotyped. The samples were collected from 27 locations, covering the full distribution range of the species from N KZN to the AO region. To enhance sample size and improve representativeness, sampling localities in close proximity, typically within the same town or municipality, were grouped together. As a result, the seven geographic regions included the following sample sizes: N KZN (29), S KZN (26), TKI (25), EL (19), PA (25), AI (24), and AO (25).

Eight microsatellite loci were genotyped, namely PBt007, PBt018, PBt-OVI-D106, LLt006, PBt001, PBt003, PBt011, and PBt014. Additionally, the polymorphic information content (PIC) values, which provide a relative estimate of marker informativeness within the studied population, ranged from 0.427 for PBt001, indicating lower informativeness, to 0.878 for PBt018, suggesting high informativeness. Overall, the PIC results demonstrate varying levels of informativeness across loci within this population. The proportion of missing data varied across these loci, with PBt018 having the lowest frequency of missing data at 1.73%, while PBt001 exhibited the highest at 38.73%. The overall average proportion of missing data across all loci was approximately 17.00%.

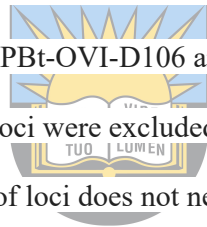
3.1 Microchecker

Eight microsatellite loci were chosen for genotyping for this study. However, during the quality control analysis using Micro-Checker, it was discovered that one of the selected loci, PBt001, had an excessive amount of missing data in three sampling regions (AO, N KZN, and S KZN). It is possible that this is because the genetic samples from these areas did not successfully amplify for this particular microsatellite locus (Van Oosterhout et al. 2004).

Despite this, PBt001 was still included in the subsequent genetic analyses due to the overall high quality of the genotyping results. The genotyping results of the remaining loci were in line with the expectations of Micro-Checker, indicating that there was no indication of large allele dropout, null alleles or stuttering. This implies that the genotyping data for these loci was of high quality and appropriate for further population genetic analyses (Van Oosterhout et al. 2004).

3.2 Linkage disequilibrium

The results of the global sample analysis (Table 3.1) showed that four loci were in linkage disequilibrium (LD), implying a higher-than-expected association between these loci (Ramakrishnan 2013). After applying the Bonferroni correction to adjust the p-values ($p = 0.01$), only one pairwise comparison (PBt-OVI-D106 and PBt007) remained significant. Loci were considered independent and no loci were excluded from any subsequent analysis. Significant LD between a single pair of loci does not necessarily indicate a systematic issue, as this is not a consistent feature, it may result from factors such as population structure, genetic drift, or sampling effects rather than true physical linkage, and therefore included these loci in the downstream analyses. Retaining all loci ensures comprehensive data utility while allowing for robust comparisons to determine if LD patterns are region-specific or widespread.



Together in Excellence

Table 3.1: Pairwise linkage disequilibrium tests among eight microsatellite loci in *Pachymetopon aeneum* across all seven localities sampled along South African distribution. Significant values (at adjusted threshold of $p = 0.01$) are highlighted in bold.

Locus	PBt007	PBt018	PBt-OVI-D106	LLt006	PBt001	PBt003	PBt011	PBt014
PBt007								
PBt018	0.16							
PBt-OVI-D106	0.00	0.08						
LLt006	0.03	0.16	0.05					
PBt001	0.03	0.24	0.06	0.02				
PBt003	0.13	0.12	0.18	0.11	0.23			
PBt011	0.12	0.35	0.19	0.06	0.15	0.18		
PBt014	0.15	0.27	0.22	0.24	0.26	0.13	0.20	



3.3 Genetic diversity

University of Fort Hare
Together in Excellence

Across the eight microsatellite loci, the average number of alleles (N_A) per locus ranged from 12.0 ± 3.32 SD (PBt014) to 17.0 ± 4.90 SD (PBt007) (Table 3.1). The mean allelic richness (A_R) varied across the eight loci, with values ranging from 2.67 (PBt001) to 5.03 (PBt018) (Table 3.2). The observed heterozygosity (H_O) values (ranged from 0.81 ± 0.04 to 0.98 ± 0.08) were found to be close to the expected heterozygosity (H_E) values (0.87 ± 0.06 to 0.93 ± 0.03 range) for all loci.

Furthermore, the inbreeding coefficient (F_{IS}) values were mostly negative or close to zero, ranging from -0.09 (PBt001 and PBt018) to 0.08 (PBt014), indicating little to no inbreeding and a potential excess of heterozygotes at several loci.

Table 3.2: Genetic diversity indices for the blue breem (*P. aeneum*) across seven South African coastal regions. The table includes the number of alleles (N), mean ($\pm$ SD) number of alleles (N_A) ($\pm$ Standard Deviation), mean allelic richness (A_R), observed heterozygosity (H_o) ($\pm$ SD), expected heterozygosity (H_E) ($\pm$ SD), inbreeding coefficient (F_{IS}), Hardy-Weinberg exact test (HWE) ($p = 0.01$), polymorphic information content (PIC) and missing data frequencies (%) for each of the eight microsatellite loci across the seven sampling regions.

Locus	N	N_A	A_R	H_o	H_E	F_{IS}	HWE	PIC	Missing Data Freq
LLt006	104	14.9 (± 4.38)	4.03	0.89 (± 0.03)	0.91 (± 0.04)	0.01	0.09	0.848	24.86%
PBt001	110	15.7 (± 5.59)	2.67	0.98 (± 0.04)	0.91 (± 0.06)	-0.09	0.19	0.427	38.73%
PBt003	93	13.3 (± 2.14)	3.69	0.91 (± 0.04)	0.91 (± 0.03)	0	0.33	0.784	20.81%
PBt007	119	17.0 (± 4.90)	3.05	0.96 (± 0.08)	0.92 (± 0.04)	-0.05	0.1	0.735	2.31%
PBt011	104	14.9 (± 2.97)	3.55	0.96 (± 0.06)	0.93 (± 0.02)	-0.04	0.27	0.784	4.62%
PBt014	84	12.0 (± 3.32)	3.7	0.81 (± 0.04)	0.87 (± 0.06)	0.08	0.33	0.844	16.76%
PBt018	90	12.9 (± 2.91)	5.03	0.98 (± 0.08)	0.91 (± 0.03)	-0.09	0.21	0.878	1.73%
PBt-OVI-D106	117	16.7 (± 3.68)	3.06	0.96 (± 0.08)	0.93 (± 0.03)	-0.04	0.18	0.783	16.19%

The average number of alleles per locus varied among the seven South African coastal regions sampled, ranging from 9.50 in the S KZN region to 18.00 in both the PA and AI regions (Table 3.3). Mean allelic richness, a measure of the number of different alleles in a population corrected for sample size, revealed values ranging from 2.56 in the PA region to 5.01 in the AO region (Table 3.2), indicating that the AO region possessed the highest level of genetic diversity on average. The inbreeding coefficients (F_{IS}) for the regions also

indicated low levels of inbreeding, and randomly mating population, values ranging from slightly negative (-0.13 for S KZN) to slightly positive (0.06 for AO) (Table 3.3).

Table 3.3: Genetic diversity statistics for *Pachymetopon aeneum* across seven South African coastal regions. The table includes the sample size per locality (N), the number of alleles per locus (N_A) with standard deviation (SD), allelic richness (A_R), the number of polymorphic loci (#PL), observed heterozygosity (H_o) with SD, expected heterozygosity (H_E) with SD and inbreeding coefficient (F_{IS}).

Locality	N	N_A	A_R	#PL	H_o	H_E	F_{IS}
N KZN	29	10.25 (± 1.75)	4.54	8	0.92 (± 0.14)	0.87 (± 0.03)	-0.06
S KZN	26	9.50 (± 1.20)	4.80	8	0.96 (± 0.10)	0.85 (± 0.04)	-0.13
TKI	25	17.25 (± 2.31)	2.62	8	0.90 (± 0.09)	0.94 (± 0.02)	0.04
EL	19	15.75 (± 2.61)	2.78	8	0.92 (± 0.06)	0.93 (± 0.01)	0.01
PA	25	18.00 (± 3.16)	2.56	8	0.95 (± 0.03)	0.94 (± 0.02)	0.00
AI	24	18.00 (± 2.67)	2.87	8	0.95 (± 0.06)	0.95 (± 0.01)	-0.10
AO	25	13.88 (± 2.23)	5.01	8	0.95 (± 0.11)	0.90 (± 0.02)	0.06

3.4 Population comparison

The population comparison analysis of *P. aeneum* across seven South African regions revealed moderate genetic differentiation among the populations. The pairwise R_{ST} values indicated strong differentiation between the N KZN and S KZN populations compared to others such as TKI, EL, PA, and AI (Table 3.6). However, the results also showed a lack of differentiation between AO and the KZN sites, suggesting a possible genetic connection or

ongoing gene flow between these locations. Additionally, TKI was found to be significantly differentiated from EL and PA, further highlighting the complex genetic structuring within the species across its range.

Table 3.4: Comparison of pairwise R_{ST} values (above diagonal) with associated p-values (below diagonal) in *Pachymetopon aeneum* populations sampled from seven regions from the South African distribution and based on eight microsatellite loci. Significance values at $p = 0.05$ are indicated in bold.

	N	KZN	S KZN	TKI	EL	PA	AI	AO
N KZN			0.042	0.480	0.491	0.439	0.342	-0.001
S KZN	0.135			0.480	0.476	0.417	0.325	0.024
TKI	0.000	0.000			0.038	0.056	0.019	0.407
EL	0.000	0.000		0.039		-0.023	0.009	0.412
PA	0.000	0.000		0.014	0.983		0.010	0.366
AI	0.000	0.000		0.101	0.261	0.215		0.262
AO	0.593	0.245		0.000	0.000	0.000	0.000	

3.5 Bayesian clustering/assignment

The results of a Bayesian clustering analysis, based on genotype frequencies, suggest that the sampled blue bream individuals likely belong to two distinct subpopulations ($K = 2$, Figure 3.1). The highest Delta K value is at $K=2$ (1051.81425), indicating that the optimal number of clusters for the given population structure is two. This means that the data best fits a model with two distinct genetic clusters.

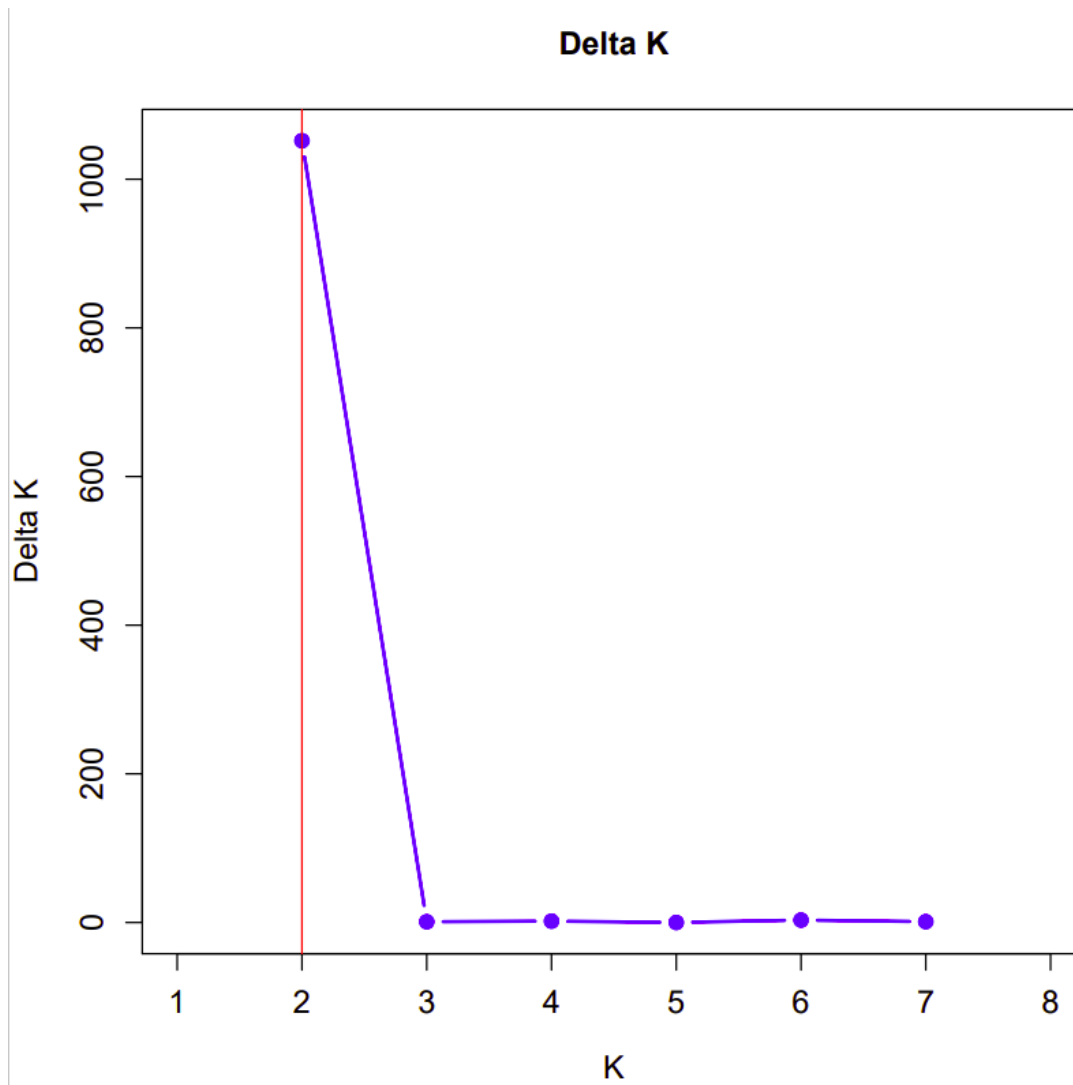


Figure 3.1: STRUCTURE results for blue breem samples drawn from seven localities representing the species' distribution range, with estimation of the most probable K using the Delta K method by Evanno et al. (2005). The analysis suggests that the optimal number of clusters is two.

Figure 3.1 presents the probability of individuals from each of the seven regions belonging to each of the identified genetic clusters (K) based on their genotypes. The analyses revealed that for K=2, 3, 4, 6, 7, and 8, the AO population consistently aligns with the KZN regions, suggesting a close genetic relationship and shared ancestry. The Bayesian clustering analysis indicates that the blue breem populations are optimally divided into two main genetic

clusters, corresponding to the AO and KZN regions, and the other localities.

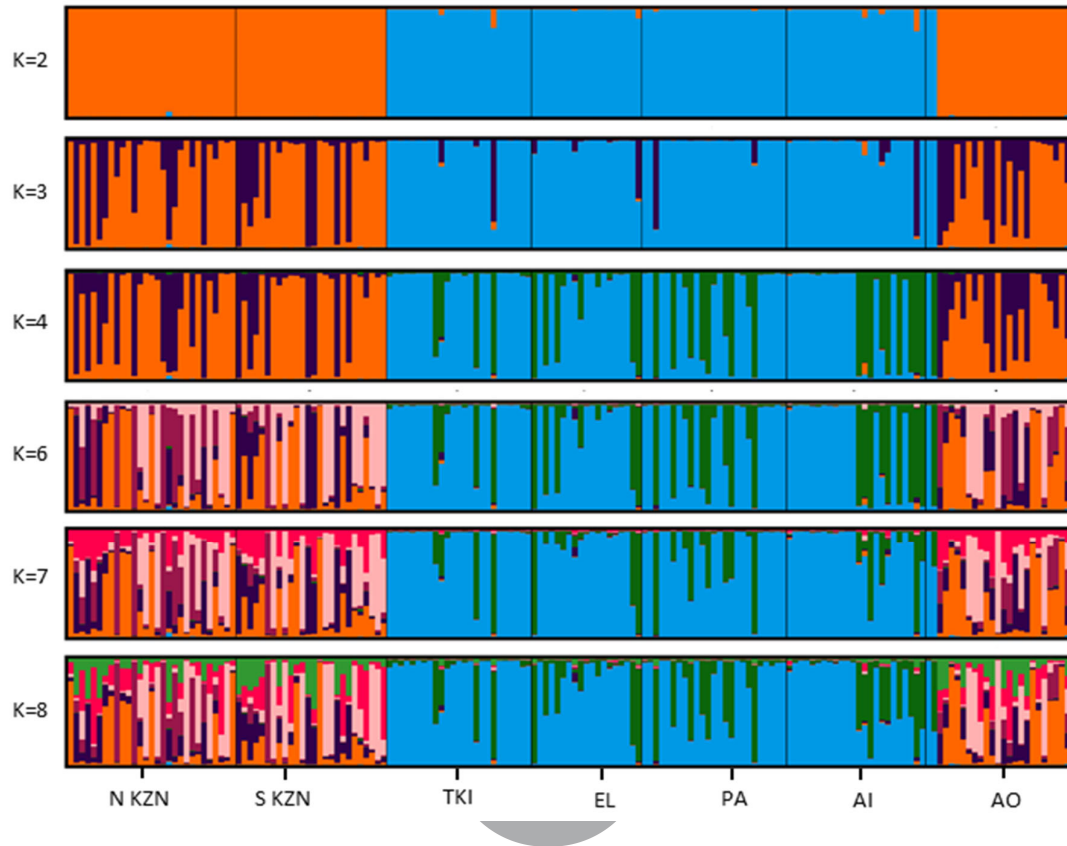


Figure 3.2: Genetic assignment and admixture analysis of blue bream (*P. aeneum*) from seven South African coastal regions, using 8 microsatellite loci and Bayesian clustering. Each vertical bar represents the inferred ancestry in each genetic group, and bar plots depict genetic structure at various K values (2, 3, 4, 6, 7, 8).

The Structure Selector analysis of *P. aeneum* revealed a genetically structured population, delineating two genetic clusters within the sampled individuals. This was strongly supported by a consensus among all estimators (MedMedK, MedMeaK, MaxMedK, and MaxMeaK), uniformly suggesting that the model with two clusters (K=2) best explains the genotypic data (Figure3.3).

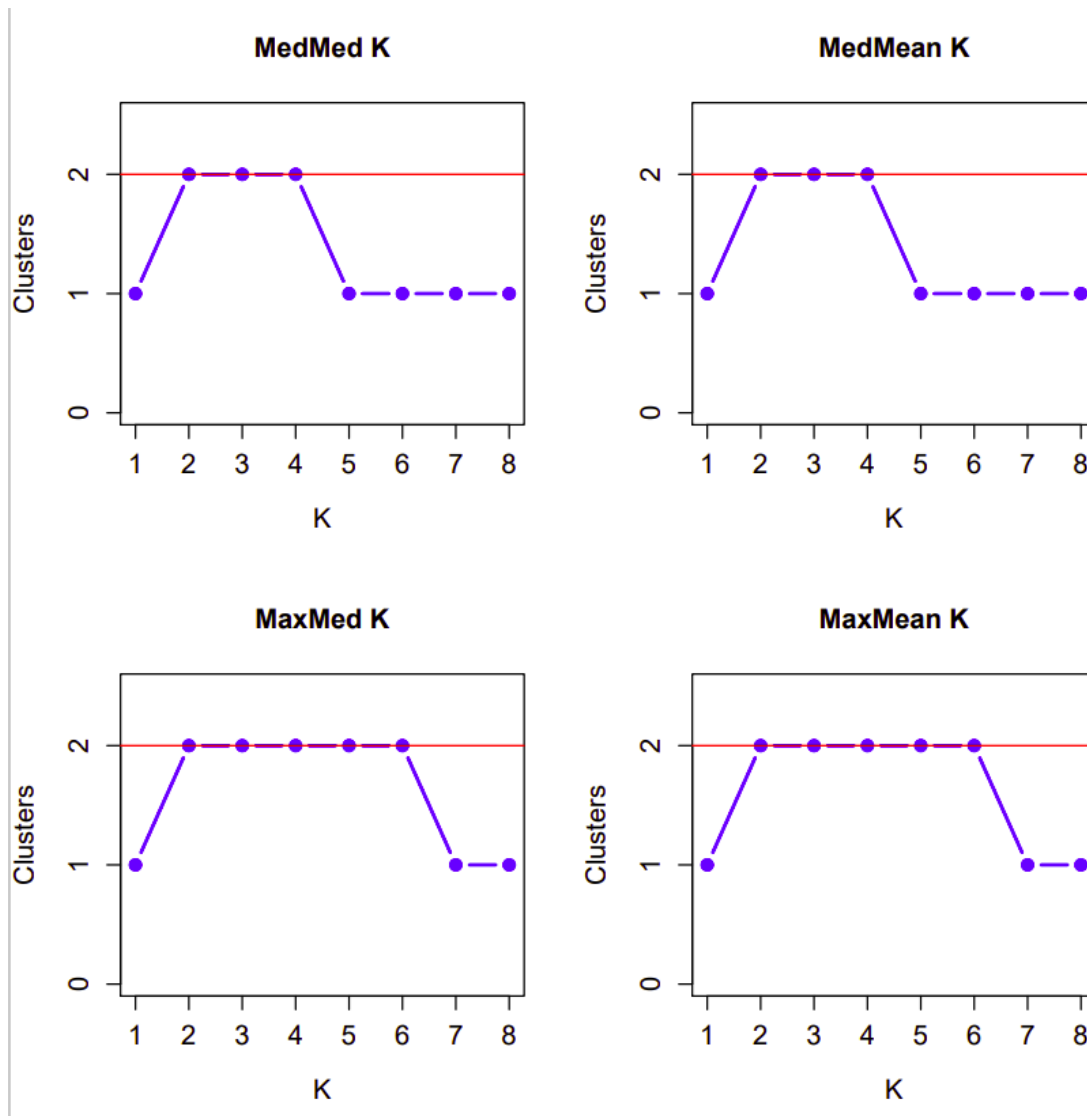


Figure 3.3: Cluster analysis results, generated by Structure Selector, revealing potential clusters within the *Pachymetopon aeneum* population across its distribution in South America. The analysis employed four different estimators: MedMedK, MedMeanK, MaxMedK, and MaxMeanK.

3.6 Principal Coordinate Analysis (PCoA)

The PCoA results revealed two apparently separate clusters of genetically similar individuals. Cluster one consisted of individuals from three sampling localities (N KZN S KZN, and AO),

while cluster two consisted of individuals from four sampling localities (TKI, EL, PA and AI) (Figure 3.4).

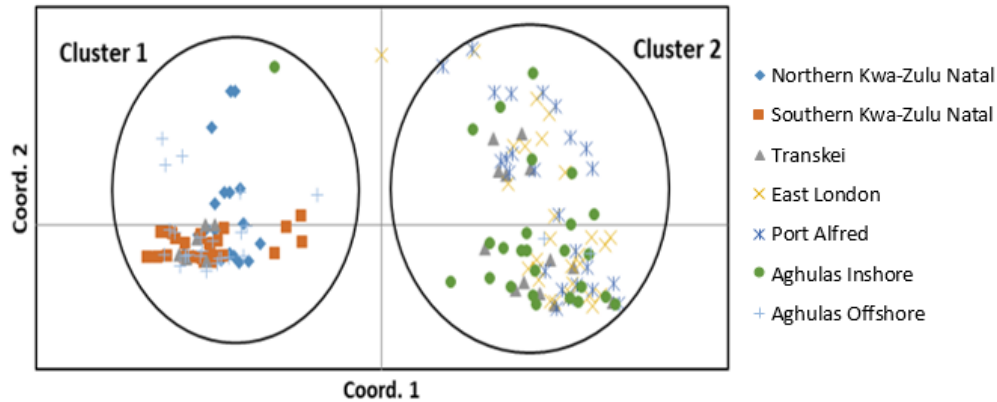


Figure 3.4: Principal Coordinates Analysis (PCoA) of microsatellite genotypic data from eight loci for *Pachymetopon aeneum* individuals from seven sampling localities across South Africa. The plot shows all sampled individuals, colour-coded by their sampling site. The axes represent the first and second principal coordinates (Coord1 and Coord2), which explain the major patterns of genetic variation.

3.7 Analysis of molecular variance (AMOVA) Fort Hare Together in Excellence

The Analysis of Molecular Variance (AMOVA) was performed on seven regions of the South African coast for the species *P. aeneum*. The genetic variance among groups was 28.48% of the total genetic variation, with a fixation index (F_{CT}) of 0.285, indicating a significant but low levels of differentiation among groups ($p = 0.000$). Among populations within groups, the genetic variance contributed 15.01% to the total variation, with a fixation index (F_{SC}) of 0.210, suggesting a significant but low differentiation among populations within groups ($p = 0.000$) (Table 3.5). Among individuals within populations, the genetic variance was 21.58%, with a fixation index (F_{IS}) of 0.382, pointing to significant but low levels of inbreeding or genetic structuring within populations ($p = 0.000$). The largest proportion of genetic variance was seen within individuals, at 34.93%, with an overall

fixation index (F_{IT}) of 0.651, indicating a high level of variation and diversity among the individual loci within the individuals ($p = 0.000$).

Table 3.5: Analysis of Molecular Variance (AMOVA) results for the partitioning of microsatellite genetic variation among hierarchical groups of *Pachymetopon aeneum* populations from seven South African regions. The variation was analysed among groups (KZN, EC, and Agulhas), among populations within groups, among individuals within populations and within individuals. The table includes estimates of the Sum of squares, percentage of variation for each hierarchical level, associated fixation indices and p-values.

Source of variation	Sum of squares	Var. Components	Percentage variation	Fixation index	p value
Among groups	1164465.757	4720.430	28.479	F_{CT} 0.285	0.000
Among pop within groups	462700.589	2487.602	15.008	F_{SC} 0.210	0.000
Among Individuals within populations	1874477.462	3577.119	21.581	F_{IS} 0.382	0.000
Within individuals	941992.500	5790.289	34.931	F_{IT} 0.651	0.000
Total	4443636.308	16575.440			

3.9 Isolation-by-distance

The Mantel test to assess isolation by distance (IBD) revealed a weak relationship between the geographic distance and genetic distance matrices in the studied population ($R^2 = 0.1268$), with a lack of statistical significance ($p = 0.070$) (Figure 3.5). This suggests that other factors, such as migration and gene flow, may be more responsible for the genetic structure of the population, rather than geographic distance alone.

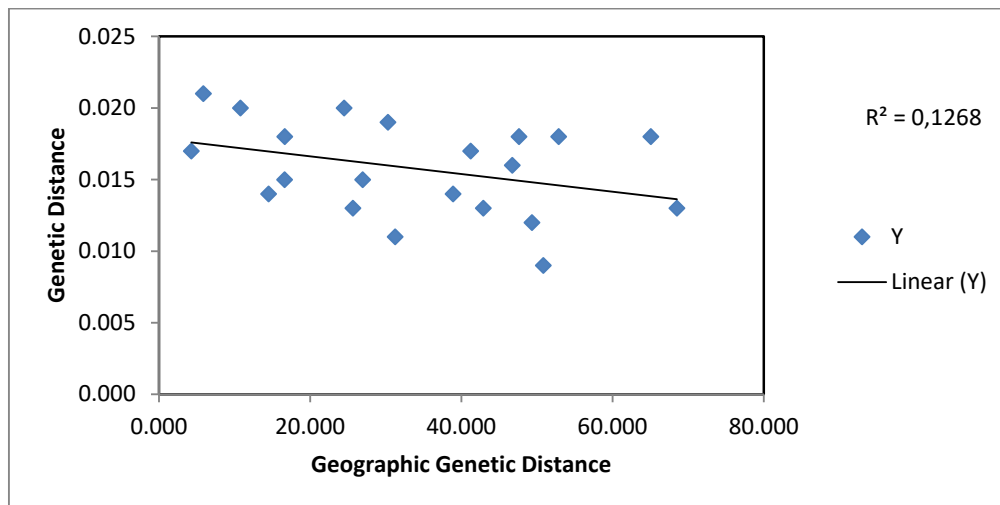


Figure 3.5: Isolation-by-distance scatterplot for blue bream. The plot shows the relationship between geographic and genetic distances among *Pachymetopon aeneum* samples from seven localities along the South African coast. Geographic distances were calculated as the shortest oceanic distances between sampling sites, with a pairwise geographic distance matrix generated. Genetic distances were calculated using Slatkin's linearized F_{ST} approach, with pairwise F_{ST} values representing genetic differentiation across eight microsatellite loci.

3.10 Relatedness

The Queller and Goodnight index has shown that the mean relatedness values range from 0.109 to 0.369 for all localities, with the highest mean relatedness observed in the PA (0.369). The Ritland index showed mean relatedness values spanning from 0.028 to 0.055, with the N KZN exhibiting the highest mean relatedness (0.055). The Lynch and Ritland index revealed the mean relatedness values ranging from 0.014 to 0.118, with the TKI exhibiting the highest mean relatedness (0.118). Despite the variations, the data from all three indices consistently showed low mean relatedness values within each locality. Furthermore, the negative mean relatedness values observed in the total population suggest that, on average, individuals across all localities are less related than expected by chance, possibly due to high genetic diversity among populations.

Table 3.6: Estimates of relatedness based on analyses of eight microsatellite loci for the total sample of *Pachymetopon aeneum* from its South African distribution and for each of the seven sampled regions. The sample sizes (N), the number of pairwise comparisons conducted within the population (NC), and the means of each relatedness index, as determined by Queller and Goodnight (1989), Ritland (1996), and Lynch and Ritland (1999) methods, are listed in the table.

			Queller and Goodnight (1989)	Ritland (1996)	Lynch and Ritland (1999)
Locality	N	NC	Mean	Mean	Mean
N KZN	29	406	0.128	0.055	0.042
S KZN	26	325	0.138	0.046	0.032
TKI	25	300	0.356	0.043	0.118
EL	19	171	0.331	0.040	0.075
PA	25	300	0.369	0.035	0.106
AI	24	276	0.284	0.038	0.089
AO	25	300	0.109	0.028	0.014
Total Population	173	14878	-0.011	-0.006	-0.004

3.11.1 Population assignment (“As is” method)

Population assignment, using the “As is” population assignment method, successfully assigned 69% of the individuals back to their respective populations (Figure 3.6). The N KZN and S KZN populations had the highest successful population reassignment proportions of 93% and 92%, respectively. The two KZN regions, which had high levels of reassignment, likely do not receive as many individuals (or gene flow) from the other regions. The PA population was shown to have the lowest successful reassignment of 48% (Figure 3.6). However, this does not mean that they (N KZN and S KZN) don’t contribute to the other

regions, as their individuals were assigned to Agulhas Offshore (not presented). However, no individuals from other regions were found in KZN; both KZN populations were mixed with other individuals from the KZN region (not presented).

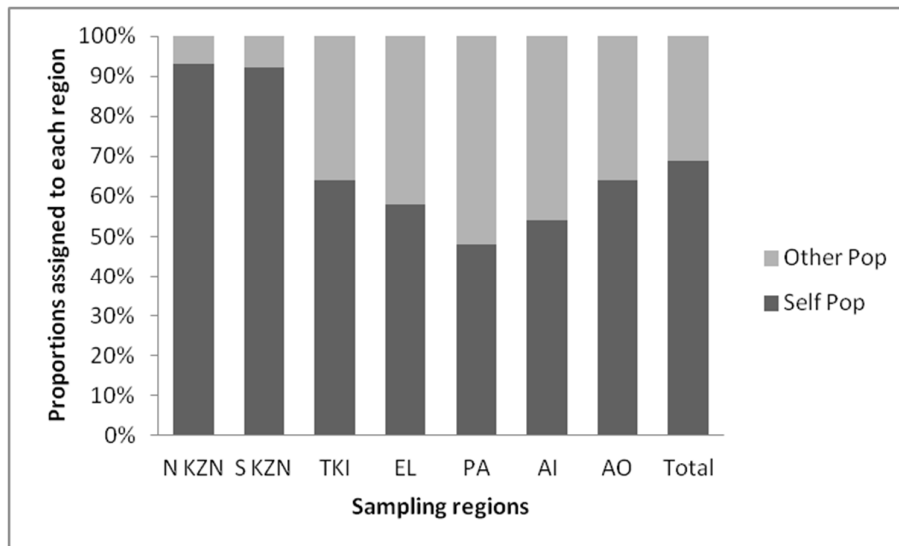


Figure 3.6: The results of a population assignment analysis for *Pachymetopon aeneum*. The "As is" method was used to assign individuals to one of seven populations, each of which represents a different geographic region within the species' distribution range in South Africa. The analysis was based on data from eight microsatellite loci. The figure shows the proportion of individuals assigned to each population, with the self-population represented in dark grey and other populations represented in light grey.

3.11.2 Population Assignment ("Leave-one-out" method)

The "Leave-one-out" population assignment method was employed to assign individuals to the respective populations. The results revealed that 42% of the individuals were successfully assigned back to their original populations, while the majority (58%) were assigned to new populations (Figure 3.7). The N KZN, TKI and S KZN populations exhibited the highest proportion of individuals successfully reassigned to the same/source population at 59%, 52% and 50%, respectively. The PA population, on the other hand, had the lowest rate of individuals successfully reassigned to the same/source population at 24% (Figure 3.7).

The results were consistent with the "As-is" results, revealing a trend where the KZN regions had higher levels of successful population reassignment, followed by the TKI and Agulhas regions, and EL (Figure 3.6). PA continued to exhibit the lowest levels of successful reassignment. Additionally, the “Leave-one-out” results revealed that individuals from the AO population were assigned to the KZN regions and vice versa (not presented). This observation suggests that these regions, which are at opposite ends of the species’ distribution, likely share similar genetic characteristics.

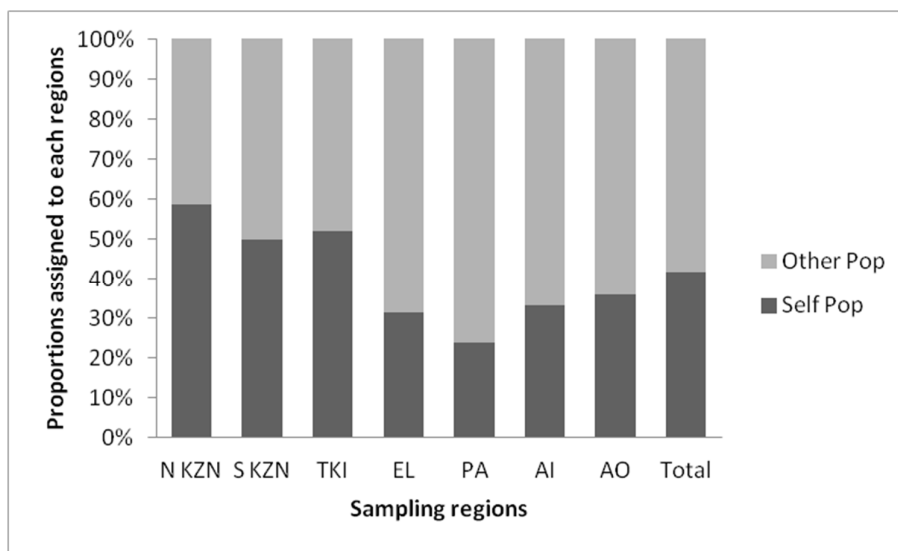
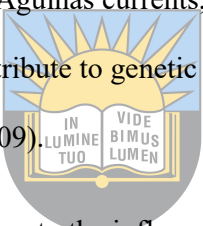


Figure 3.7: Proportion of individuals from *Pachymetopon aeneum* assigned to their respective populations using the "Leave one out" method. The analysis was based on eight microsatellite loci and assigned individuals to one of seven populations, each representing a different geographic region within the species' distribution range in South Africa. Dark grey represents the proportion of individuals assigned to their self-population while light grey represents the proportion of individuals assigned to other populations.

Chapter 4

Discussion

The microsatellite analysis of *P. aeneum* revealed higher genetic diversity with H_0 values of 0.90–0.95 and A_R of 2.87–5.01. These values were more comparable to *C. puniceus* (Duncan et al., 2015), which exhibited slightly higher allelic richness but lower heterozygosity. This indicates substantial genetic variation in *P. aeneum*, supporting its adaptive potential and population resilience. The analysis also provided evidence for significant genetic structuring, with the identification of two genetically distinct subpopulations. These findings support previous studies suggesting that the unique oceanographic conditions along the South African coastline, such as the influence of the Agulhas currents, combined with the biological characteristics of *P. aeneum*, may contribute to genetic differentiation among populations (Teske et al. 2011, von der Heyden 2009).



Genetic differentiation was expected due to the influence of factors such as regional ocean currents, habitat discontinuity, and historical biogeographic barriers, which can limit gene flow and promote local adaptation within populations (Lett et al. 2024). The differentiation observed between the KZN and EC regions suggests that biogeographic features, such as the transition between tropical and temperate zones, play a crucial role in shaping population structure (Lett et al. 2024, Teske et al. 2019).

High levels of gene flow were observed within the population, with a directional pattern from KZN to AO, as indicated by the extensive genetic material shared between these regions, while EC and AI also exhibited high levels of gene exchange. The informativeness of the microsatellite markers, as indicated by their high polymorphic information content, further supports the reliability of these findings, reinforcing the robustness of the genetic data in assessing population structure and connectivity in *P. aeneum*.

The observed genetic structure within *P. aeneum* suggests that, while larval dispersal mechanisms, such as ocean currents, wind, and tides, likely facilitate gene flow, they do not result in complete genetic homogeneity across regions (Lett et al. 2024). Instead, the genetic differentiation observed between the KZN and EC regions highlights the influence of regional oceanographic and biogeographic barriers in shaping population structure. The high genetic similarity between KZN and AO, for instance, may be attributed to prevailing southward-flowing ocean currents, such as the Agulhas Current, which could transport larvae and facilitate gene exchange over long distances (Schumann 1998, Lett et al. 2024). However, the biogeographic transition zones and habitat discontinuities along the coastline likely act as barriers that contribute to genetic structuring (Lett et al. 2024).

The South African coastline is characterized by complex oceanographic features, with the warm Agulhas Current dominating the east coast and moving southward along the KZN region, while the colder Benguela Current influences the west coast, flowing northward along the Western Cape. The interaction of these currents with coastal topography and wind-driven upwelling events can create distinct ecological zones that influence the dispersal and connectivity of marine populations (Lett et al. 2024). These environmental factors, coupled with potential spawning migrations (Garratt, 1998), may explain the observed patterns of genetic differentiation and connectivity in *P. aeneum*.

4.1 Comparison with other species

The findings of this study were inconsistent with some previous studies conducted on the population genetics of other South African sparids. Specifically, studies on *Lithognathus lithognathus*, utilizing both mtDNA and microsatellites markers, as well as studies on *Pachymetopon grande*, *Rhabdosargus holubi* (Oosthuizen 2006), *Petrus rupestris* (Gouws et al. 2020), *Chrysoblephus puniceus* (Duncan et al. 2015), *Chrysoblephus laticeps* (Teske et al.

2010), *Chrysoblephus laticeps* (Teske et al. 2010), *Lichia amia* (Henriques et al. 2012) and *Atractoscion aequidens* (Henriques et al. 2014) have all reported high levels of genetic diversity within populations, yet failed to detect significant genetic structuring (Table 4.1).

Table 4.1: Comparison of studies looking at genetic diversity in commercially-significant South African fish species: The studies either used mitochondrial DNA control region or microsatellites or both to test for genetic structure.

Species	Genetic marker	Genetic Structure	Reference
<i>Pachymetopon aeneum</i>	Msat	Yes	Current study
<i>Petrus rupestris</i>	mtDNA	No	Gouws et al. 2020
<i>Chrysoblephus puniceus</i>	mtDNA	No	Duncan et al. 2015
<i>Atractoscion aequidens</i>	mtDNA	No	Henriques et al. 2014
<i>Lithognathus lithognathus</i>	mtDNA & msat	No	Bennet 2012
<i>Lichia amia</i>	mtDNA	Yes	Henriques et al. 2012
<i>Chrysoblephus laticeps</i>	mtDNA	No	Teske et al. 2010

University of Fort Hare
Together in Excellence

In terms of genetic diversity, *P. aeneum* exhibited high levels based on several metrics of genetic variation. Observed heterozygosity (H_o) for *P. aeneum* ranged from 0.90 to 0.95, and H_e ranged from 0.85 to 0.95. These values indicated strong genetic variability within the species. The number of alleles per locus (N_A) ranged from 9.50 to 18.00, and allelic richness (A_R) ranged from 2.87 to 5.01. These statistics suggested substantial allelic diversity within *P. aeneum* populations.

Compared to other sparid species, such as *C. puniceus* and *L. lithognathus*, some variation in genetic diversity was observed. *Chrysoblephus puniceus* exhibited H_o ranging from 0.75 to

0.79, and H_E from 0.81 to 0.84 (Duncan et al. 2015), which were slightly lower than those observed in *P. aeneum*. Similarly, *L. lithognathus* exhibited H_O values ranging from 0.72 to 0.77, and H_E values from 0.78 to 0.80 (Bennet, 2012), also lower than *P. aeneum*. The number of alleles per locus (N_A) in *C. puniceus* was slightly higher than that of *P. aeneum*, while that of *L. lithognathus* was lower. Although the allelic richness (A_R) in these species had higher ranges than in *P. aeneum*, the latter still demonstrated overall relatively higher genetic diversity. Overall, the genetic diversity observed in *P. aeneum* was consistent with, and in some cases exceeded, that of other fish species. For example, *Sebastes schlegelii* and *Konosirus punctatus*, both marine species, exhibited lower H_O and H_E values (Wang et al. 2020, Liu et al. 2022) compared to *P. aeneum* (Table 4.2). However, it is also important to note that genetic diversity can vary widely depending on the species, the specific loci scored, the number of loci scored, and the sample size (Fritsch et al. 2007).

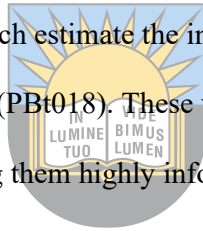
Table 4.2: Comparison of studies looking at genetic diversity in marine fish species: All the studies used microsatellites to test for genetic structure.

Species	N_A	A_R	H_O	H_E	# of loci	Reference
<i>P. aeneum</i>	9.50 - 18.00	2.87 - 5.01	0.90 - 0.95	0.85 - 0.95	8	Current study
<i>C. puniceus</i>	15.8 - 19.4	15.3 - 16.6	0.75 - 0.79	0.81 - 0.84	10	Duncan et al., 2015
<i>L. lithognathus</i>	8.33 - 10.00	7.62 - 8.15	0.72 - 0.77	0.78 - 0.80	12	Bennet PhD 2012
<i>S. schlegelii</i>	-	-	0.67 - 0.74	0.70 - 0.76	38	Wang et al., 2020
<i>K. punctatus</i>	8.35 - 11.75	8.35 - 9.73	0.60 - 0.75	0.72 - 0.81	20	Liu et al., 2022
<i>R. holubi</i>	9-19	1-21	0.47 - 1.00	0.82 - 0.96	2	Ooisthuizen MSc 2006

4.2 Biological interpretation

Fundamental data suitability

In population genetics, assessing the fundamental data suitability of genetic markers is crucial for ensuring the validity and reliability of the results. Most loci were in Hardy-Weinberg Equilibrium, with p-values exceeding the threshold of 0.01, indicating random mating within populations and no significant effects of selection, migration, or genetic drift disrupting equilibrium. The low levels of significant LD and the general adherence to HWE across loci affirm the suitability of the selected microsatellite markers for assessing the genetic diversity and structure of *P. aeneum*. These markers accurately reflect the natural population dynamics without the bias of physical linkage or non-random mating. Additionally, the polymorphic information content (PIC) values, which estimate the informativeness of genetic markers, ranged from 0.427 (PBt001) to 0.878 (PBt018). These values indicate that the loci exhibit a high degree of polymorphism, making them highly informative for population genetic analyses.



University of Fort Hare
Together in Excellence

Genetic diversity

Allelic richness (A_R), which is a measure of the number of alleles per locus, provides an insight into the genetic diversity within the blue bream population (Hughes et al., 2008). Although this is relative rather than absolute measure, the KZN regions and the AO population have a greater variety of alleles per locus compared to the EC regions and AI. This could be due to the historical gene flow, larger effective population sizes, or more stable population histories and higher potential to adapt to future environmental changes in KZN and AO populations (Greenbaum 2014).

High H_O and H_E values (~ 0.9) were observed throughout the blue bream sampled populations. These high values indicate a high level of genetic variability within all the

studied populations. Because these values are close to the maximum, they suggest extensive genetic diversity and a lack of recent inbreeding events (Reed & Frankham 2003). This is beneficial as it suggests that the populations have maintained a healthy level of genetic diversity, which is important for adaptability and resilience (Reed & Frankham 2003). Furthermore, when H_o (the proportion of heterozygous individuals observed) is close to H_E (the proportion expected under Hardy-Weinberg equilibrium), it suggests that the populations are randomly mating and not experiencing significant inbreeding or selection pressures (Reed & Frankham 2003, Hedrick 2016).

For inbreeding, negative F_{IS} in KZN and AI populations could suggest an excess of heterozygotes (Hedrick 2016) compared to what is expected under Hardy-Weinberg equilibrium. This could indicate outbreeding, which might be due to migration or gene flow from other populations. In general, low F_{IS} was observed throughout the blue bream populations. Low but positive F_{IS} values suggest a slight deficiency of heterozygotes, which could be due to inbreeding or non-random mating (Reed & Frankham 2003, Hedrick 2016). However, low values indicate that this effect is minimal and not a major concern for these populations.

The overall high genetic diversity indicated by high A_R , N_A , H_o , and H_E suggests that the blue bream populations are genetically healthy, which is favourable for their long-term survival and adaptability. This trend of high gene flow and genetic differentiation is consistent with patterns observed in other marine species that exhibit high levels of adult dispersal (Teske et al. 2003, von der Heyden et al. 2007), while the opposite is true for species with low levels of adult dispersal. The *P. aeneum* populations may exhibit adult dispersal in the form of a suggested spawning migration (Garratt 1998), which could contribute to the observed patterns of reduced genetic differentiation. This is consistent with

the expectation that marine species, due to their commonly larger population sizes and the absence of physical barriers in the marine environment, exhibit higher levels of gene flow compared to their terrestrial or freshwater counterparts (Grant & Bowen 1998, Billington 2003). The high genetic diversity indicated by the results suggests that blue bream populations are genetically healthy, which is favourable for their long-term survival and adaptability (Amos & Hardwood 1998).

Overall patterns of population structure

When comparing populations across seven South African regions, genetic differentiation was observed. Pairwise R_{ST} values indicated strong differentiation, particularly between the N KZN and S KZN populations compared to others such as TKI, EL, PA, and AI. This genetic differentiation between KZN populations and TKI, EL, PA, and AI populations can be attributed to the boundaries between different biogeographic regions as identified by Teske et al. (2011). The observed genetic differentiation among regional populations of *P. aeneum* could be influenced by several ecological and environmental factors (Teske et al. 2011). Oceanographic features, such as the transition between the warm Agulhas Current in KwaZulu-Natal and the cooler upwelling-influenced waters of the Eastern Cape may create distinct environmental conditions that limit larval dispersal and gene flow. Additionally, habitat preferences and differences in coastal topography, including estuarine versus rocky reef environments, could contribute to localized adaptations and restricted movement between regions (Teske et al. 2011, Lett 2024)

The significant among-group variance ($F_{CT} = 0.285$, $p = 0.000$) suggests that these biogeographical groupings indeed capture meaningful genetic divergence, likely due to the regional ecological differences that influence local adaptation and genetic differentiation. Moderate genetic differentiation was observed among populations within the same group

($F_{SC} = 0.210$, $p = 0.000$), indicating some degree of sub-structuring within these regions. This suggests that, while populations within each group are genetically related, local adaptations or micro-environmental factors (such as variations in habitat, salinity, or ocean currents) may contribute to subtle genetic differences between populations. These patterns of moderate differentiation are likely reflective of environmental variability within each biogeographical province, which can create unique selective pressures.

Finally, the highest genetic variation (34.9%) was found within individuals ($F_{IT} = 0.651$, $p = 0.000$), suggesting that a large portion of genetic diversity exists at the within-population level. This is consistent with the expectation that gene flow within populations is high, which could result in the maintenance of genetic diversity within local populations despite regional differentiation. Therefore, AMOVA results highlight the importance of ecological and environmental factors in shaping genetic differentiation at both regional and local scales. While the significant differentiation among groups suggests the influence of biogeographical provinces, the moderate sub-structuring within groups and high variation within populations reflect the complex interplay of local adaptation and gene flow.

High F_{IS} values suggest considerable genetic diversity at the individual level within populations, possibly due to a high rate of outbreeding and gene flow. High individual genetic diversity is beneficial for the adaptive potential and resilience of the populations. The substantial genetic variation within individuals underscores the high level of heterozygosity and overall genetic diversity. This highlights the genetic robustness of the populations, providing a buffer against environmental changes and diseases.

The PCoA further supported the genetic structure existing in the population of the blue bream, by revealing clear genetic differentiation among blue bream populations. Individuals from the KZN regions and the AO population formed a distinct cluster, while populations

from the EC regions and AI showed more overlap. This suggests that KZN and AO populations are genetically distinct, whereas EC and AI populations are more genetically similar. This genetic differentiation could be due to several factors, including historical separation, different environmental conditions, or ecological barriers such as ocean currents and habitat preferences that limit gene flow between these regions. For *P. aeneum*, this differentiation suggests that KZN and AO populations may have unique adaptations to their local environments. It also implies that conservation efforts should consider these genetic differences to maintain the evolutionary potential of each population.

The STRUCTURE analysis results provide robust evidence of significant genetic differentiation among the blue breem populations, with results showing the highest Delta K value at K=2. This suggests that the population of *P. aeneum* is divided into two major clusters, which agrees with the PCoA. This significant division could be due to historical separation, geographic barriers, or ecological differences between the two main clusters. In this case, it suggested that a major genetic distinction is between the KZN populations and the AO population versus other regions. As for all tested values of K (2, 3, 4, 6, 7, and 8), the analysis revealed this consistent genetic division (KZN-AO populations vs others). The differentiation of N KZN, S KZN, and AO from other regions suggests that geographic or environmental barriers are influencing gene flow. These barriers could be physical (e.g., ocean currents) or ecological (e.g., habitat preferences, spawning sites), limiting interbreeding between populations. The close genetic relationship within these clusters indicates shared ancestry, possibly due to historical connectivity or recent gene flow. This can happen through natural dispersal mechanisms, such as ocean currents facilitating the movement of individuals between regions.

The IBD analysis aimed to assess whether genetic differentiation among populations is due to geographic distance. The lack of statistical significance ($p < 0.05$) in this analysis indicates no positive correlation between genetic and geographic distances. This suggests that factors other than geographic distance, such as migration and gene flow, are more responsible for the genetic structure of the population. The absence of a strong IBD pattern means that gene flow occurs relatively freely across all distances, possibly facilitated by ocean currents or the species' own migratory behaviors. This high gene flow can help maintain genetic diversity and reduce the risks associated with inbreeding. *P. aeneum*'s ability to move and migrate over large distances during spawning seasons can contribute to the observed genetic homogeneity across geographically-separated regions. This behavior helps maintain genetic connectivity and diversity across the population.

The analysis of relatedness provided insights into the genetic similarity among individuals within blue breem populations. All mean relatedness values for sampled populations were low and less than zero, indicating low intra-population relatedness and suggesting high levels of gene flow (Lynch & Ritland 1999). This genetic diversity could result from large population sizes, historical admixture, or other factors that increase genetic variability (William & Harwood 1998). However, this is not always the case, because of drift (Prober & Brown 1994, Raijmann et al. 1994). Similar patterns of low intra-population relatedness have been observed in other sparid species (Bennet 2012). However, Bennett et al.'s (2012) study on *L. lithognathus* stock delineation revealed no genetic structuring. *Pachymetopon aeneum* adults are known to move around during the spawning season, which may contribute to the observed levels of genetic homogeneity among adults (Budford & Larson 2006). This provides further evidence for the presence of high levels of gene flow among *P. aeneum* populations and suggests that spawning migrations may also play a role in shaping the gene flow of this species.

The observed low intra-population relatedness and significant genetic homogeneity suggest that *P. aeneum* populations are well-mixed with minimal genetic differentiation. This implies that there are no strong barriers to gene flow within these populations, allowing for random mating and high genetic exchange. The role of ocean currents, habitat connectivity, and other ecological factors in facilitating dispersal is critical. Understanding these factors can help identify key habitats and corridors that are essential for maintaining gene flow.

The Population Assignment analysis showed that most individuals were correctly assigned to their respective populations with high probability. Individuals from EC and AI populations showed mixed ancestry, as they were always assigned to each other, indicating gene flow between these regions. On the other hand, KZN and AO populations were consistently assigned within their respective groups, suggesting limited gene flow from other regions, confirming the distinct genetic clusters identified by STRUCTURE. The high assignment probability of KZN individuals to their putative population suggests strong genetic cohesion within this region. This could mean that KZN populations are relatively isolated, receiving fewer migrants from other regions (from AO only, per population assignment results), possibly due to environmental parameters or localized adaptations. Despite the geographical distance between KZN and AO, there is evidence of gene flow between these regions. This could be facilitated by ocean currents or migratory behavior during spawning seasons, which enables individuals from AO to mix with KZN populations.

4.6 Conclusion

The comprehensive genetic analysis of the blue bream (*Pachymetopon aeneum*) populations, utilizing microsatellite markers, has revealed significant insights into the genetic structuring, diversity, and gene flow within and among populations across various regions. The key findings from genetic indices, population comparisons, LD analysis, relatedness, PCoA,

AMOVA, STRUCTURE analysis, IBD, and population assignment have collectively high levels of genetic diversity and genetic structuring of *P. aeneum* populations. Microsatellite analysis indicated high levels of genetic diversity within populations, as evidenced by low mean relatedness values and low LD, which suggest extensive gene flow. The PCoA revealed distinct genetic clusters, particularly highlighting the genetic differentiation of KZN and AO populations relative to other regions. This pattern is supported by the AMOVA results, which quantified significant genetic variation among predefined biogeographical groups (28.479%) and moderate differentiation among populations within these groups (15.008%), suggesting a degree of genetic structuring at both regional and local scales.

The observed genetic differentiation and high genetic diversity within *P. aeneum* populations could have several implications for management and conservation. Conservation efforts should prioritize maintaining high levels of genetic diversity within populations, which is crucial for blue bream's adaptability and resilience. Given the distinct genetic clusters, particularly in KZN and AO, region-specific conservation strategies should be developed to preserve unique genetic traits and adaptations.

Microsatellites have been widely used in population genetics studies as they have been shown to be highly effective in detecting genetic structuring and subtle differences between populations in different localities. They are also useful in indicating the current gene flow patterns within a population (Hauser & Ward 1998, Gold & Turner 2002). However, for a more comprehensive and in-depth understanding of intraspecific genetic variation, mtDNA analysis should also be incorporated into future studies. This is because mtDNA has been shown to provide valuable insights into the historical demographic and evolutionary history of populations (Wilson et al. 1985, Gold & Turner 2002, Lemaire et al. 2005).

Additionally, it is important to conduct movement studies focusing on the movement patterns of *P. aeneum*. These studies will provide valuable information on the spawning migration of *P. aeneum*, which can help to better understand the presence or form of genetic variation and gene flow observed in this study. Larval behavioral studies can also be conducted as larvae can play an active role in their own dispersal (Tilney et al. 1996, Mora & Sale 2002, Patrick 2008). These studies can provide further information on how environmental factors, such as currents and tides, can influence the dispersal patterns of *P. aeneum* larvae, ultimately affecting the genetic structure and gene flow of populations



University of Fort Hare
Together in Excellence

References

- Allendorf, F.W., 1986. Genetic drift and the loss of alleles versus heterozygosity. *Conservation Biology*, 1(1), pp.19-23.
- Allendorf, F.W., Luikart, G. and Aitken, S.N., 2007. *Conservation and the genetics of populations*. *Mammalia*, 2007(2007), pp.189-197.
- Cowen, R.K. and Sponaugle, S., 2009. Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, 1, pp.443–466.
- Duncan, M., James, N., Fennessy, S.T., Mutombene, R.J. and Mwale, M., 2015. Genetic structure and consequences of stock exploitation of *Chrysoblephus puniceus*, a commercially important sparid in the South West Indian Ocean. *Fisheries Research*, 164, pp.64-72.
- Evanno, G., Regnaut, S. and Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), pp.2611-2620.
- Excoffier, L. and Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), pp.564-567.
- Excoffier, L. and Lischer, H.E., 2011. Arlequin ver 3.5: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, 7, pp.371-373.
- Excoffier, L., Lischer, H.E.L. and Petrov, D.A., 2009. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 9(3), pp.564-567.

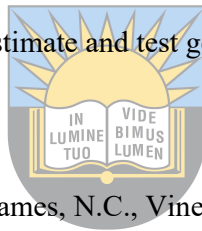
Fay, J.C. and Wu, C.I., 2000. Hitchhiking under positive Darwinian selection. *Genetics*, 155(1), pp.1449-1460.

Gold, J.R. and Turner, T.F., 2002. Population structure of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico, as inferred from variation in nuclear-encoded microsatellites. *Marine Biology*, 140, pp.249-265.

Goldstein, D.B., 1995. Microsatellites: evolution and applications. *Trends in Ecology and Evolution*, 10(4), pp.130-135.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, 86(6), pp.485-486.

Goudet, J., 2002. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2).



Gouws, G., Kerwath, S.E., Potts, W.M., James, N.C., Vine, N.G. and Cowley, P.D., 2020. High genetic diversity and limited spatial structure in an endangered, endemic South African sparid, the red steenbras *Petrus rupestris*. *African Journal of Marine Science*, 42(3), pp.295-306.

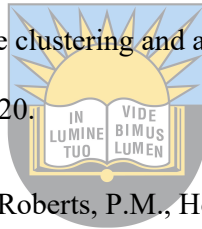
Gower, J.C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3-4), pp.325-338.

Hauser, L. and Ward, R.D., 1998. Population identification in pelagic fish: the limits of molecular markers. In: Carvahlo, G.R. (ed.), *Advances in molecular ecology*. Amsterdam: IOS Press, pp.191-224.

Hedrick, P.W., 2005. Population genetics. The Hardy-Weinberg equilibrium. *Journal of Heredity*, 96(6), pp.773-779.

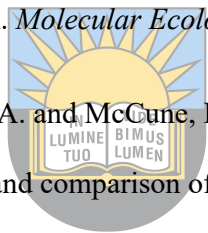
- Henriques, R., Potts, W.M., Santos, C.V., Sauer, W.H.H. and Shaw, P.W., 2014. Population connectivity and phylogeography of a coastal fish, *Atractoscion aequidens* (Sciaenidae), across the Benguela Current region: Evidence of an ancient vicariant event. *PLoS ONE*, 9(2), p.e87907.
- Henriques, R., Potts, W.M., Sauer, W.H.H. and Shaw, P.W., 2012. Evidence of deep genetic divergence between populations of an important recreational fishery species, *Lichia amia* L. 1758, around southern Africa. *African Journal of Marine Science*, 34(4), pp.585-591.
- Hudson, R.R., 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, 18(2), pp.337-338.
- Hughes, J.M., Bell, G. and Kaufman, D.S., 2008. Microsatellites for population genetics. *Nature Reviews Genetics*, 9(2), pp.129-140.
- Jakobsson, M. and Rosenberg, N.A., 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), pp.1801-1806.
- Kalinowski, S.T., Taper, M.L. and Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16(5), pp.1099-1106.
- Kruglyak, L., 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics*, 22(3), pp.139-144.
- Lemaire, C., Versini, J.J. and Bonhomme, F., 2005. Maintenance of genetic differentiation across a transition zone in the sea: Discordance between nuclear and cytoplasmic markers. *Journal of Evolutionary Biology*, 18(1), pp.70-80.

- Li, Y.L. and Liu, J.X., 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, 18(1), pp.176-177.
- Levin, L.A., 2006. Recent progress in understanding larval dispersal: New directions and digressions. *Integrative and Comparative Biology*, 46(3), pp.282-297.
- Lynch, M. and Ritland, K., 1999. Estimation of pairwise relatedness with molecular markers. *Genetics*, 152(4), pp.1753-1766.
- Lynch, M. and Walsh, B., 1998. *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2), pp.209-220.
- Mendenhall, E., Hendrix, C., Nyman, E., Roberts, P.M., Hoopes, J.R., Watson, J.R., Lam, V.W. and Sumaila, U.R., 2020. Climate change increases the risk of fisheries conflict. *Marine Policy*, 117, p.103954.
- Menozzi, P., Piazza, A. and Cavalli-Sforza, L.L., 1978. Synthetic maps of human gene frequencies in Europeans. *Science*, 201(4357), pp.786-792.
- Mora, C. and Sale, P.F., 2002. Are populations of coral reef fish open or closed? *Trends in Ecology and Evolution*, 17(9), pp.422-428.
- Novembre, J. and Stephens, M., 2008. Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*, 40(5), pp.646-649.



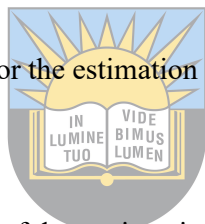
University of Fort Hare
Together in Excellence

- Oosthuizen, C.J., 2006. *Genetic variation within Cape stumpnose, Rhabdosargus holubi Steindachner (Teleostei: Sparidae)*. Doctoral dissertation, University of Pretoria, South Africa.
- Palumbi, S.R., 2003. Population genetics, demographic connectivity and the design of marine reserves. *Ecological Applications*, 13, S146-S158.
- Patrick, P., 2008. Larval fish dynamics in the shallow nearshore of eastern Algoa Bay with particular emphasis on the effects of currents and swimming abilities on dispersal. MSc thesis, Rhodes University, South Africa.
- Peakall, R. and Smouse, P.E., 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), pp.288-295.
- Peres-Neto, P.R., Jackson, D.A., Elliot, J.A. and McCune, B., 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, 87(9), pp.2614-2625.
- Petit, R.J., El Mousadik, A. and Pons, O., 1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12(6), pp.844-855.
- Piredda, R., Mottola, A., Cipriano, G., Carlucci, R., Ciccacese, G. and Di Pinto, A., 2022. Next generation sequencing (NGS) approach applied to species identification in mixed processed seafood products. *Food Control*, 133, p.108590.
- Pontes, A.I., Loeb, M.V., Mai, A.C.G., Mott, T., Batalha-Filho, H. and Jacobina, U.P., 2024. Pleistocene sea level variation and changes in the São Francisco river drove the evolution of Atlantic sabretooth anchovy. *Estuarine, Coastal and Shelf Science*, 298, p.108628.



University of Fort Hare
Together in Excellence

- Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945-959.
- Queller, D.C. and Goodnight, K.F., 1989. Estimating relatedness using genetic markers. *Evolution*, 43, pp.258-275.
- Reed, D.H. and Frankham, R., 2003. Correlation between fitness and genetic diversity. *Conservation Biology*, 17(1), pp.230-237.
- Reich, D.E. and Lander, E.S., 2001. On the allelic spectrum of human disease. *Trends in Genetics*, 17(2), pp.502-510.
- Rice, W.R., 1989. Analyzing tables of statistical tests. *Evolution*, 43(1), pp.223-225.
- Ritland, K., 1996. Extensions of models for the estimation of pairwise relatedness. *Heredity*, 76, pp.261-269.
- Simianer, H., 2005. Genetic conservation of domestic animal populations: from the perspective of population genetics. *Journal of Animal Breeding and Genetics*, 122(1), pp.1-19.
- Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139(1), pp.457-462.
- Slatkin, M., 2008. Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9(4), pp.477-485.
- Szulkin, M., Merilä, J. and Sheldon, B.C., 2010. Heterozygosity and fitness in natural populations of birds and mammals. *Molecular Ecology*, 19(2), pp.397-407.
- Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17(22), pp.6463-6471.



University of Fort Hare
Together in Excellence

- Teske, P.R., Cherry, I.M. and Matthee, C.A., 2003. Population genetics of the endangered Knysna seahorse, *Hippocampus capensis*. *Molecular Ecology*, 12, pp.1703-1715.
- Teske, P.R., Von der Heyden, S., McQuaid, C.D. and Barker, N.P., 2011. A review of marine phylogeography in Southern Africa. *South African Journal of Science*, 107, pp.1-11. (doi:10.4102/sajs.v107i5/6.514).
- Teske, P.R., Forget, F.R.G., Cowley, P.D., Von der Heyden, S. and Beheregaray, L.B., 2010. Connectivity between marine reserves and exploited areas in the philopatric reef fish *Chrysoblephus laticeps* (Teleostei: Sparidae). *Marine Biology*, 157, pp.2029-2042.
- Tilney, R.L., Nelson, G., Radloff, S.E. and Buxton, C.D., 1996. Ichthyoplankton distribution and dispersal in the Tsitsikamma National Park marine reserve, South Africa. *South African Journal of Marine Science*, 17, pp.1-14.
- Toro, M.A., Fernandez-Lopez, J. and Toro, M.A., 2009. Heterozygosity-fitness correlations in natural populations. *Biological Reviews*, 84(4), pp.641-658.
- Van de Castele, T., Palsbøll, P.J. and De Meyer, T., et al., 2001. Estimating relatedness from microsatellites: a comparison of methods and a review of software. *Molecular Ecology*, 10, pp.2469-2485.
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P. and Kelly, D.W., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), pp.535-538.
- von der Heyden, S., Lipinski, M.R. and Matthee, C.A., 2007. Mitochondrial DNA analyses of the Cape hakes reveal an expanding, panmictic population for *Merluccius capensis* and population structuring for mature fish in *Merluccius paradoxus*. *Molecular Phylogenetics and Evolution*, 42, pp.517-527.

- Wang, J., 2002. An estimator for pairwise relatedness using molecular markers. *Genetics*, 160(3), pp.1203-1215.
- Wang, Y., 2017. Recent advances in the study of the genetics of fish. In: *Fish genetics and aquaculture* (pp. 1-27). Springer, Singapore.
- Waples, R.S., 1998. Definition of "species" under the Endangered Species Act: genetic vs. evolutionary considerations. *Conservation Biology*, 12(3), pp.844-848.
- Weir, B.S., 1996. *Genetic data analysis II: methods for discrete population genetic data*. Sinauer Associates.
- Weir, B.S. and Hill, W.G., 2002. Estimating F-statistics for the analysis of population structure. *Evolution*, 56(8), pp.1358-1370.
- Weir, B.S. and Hill, W.G., 2002. *Genetic data analysis II: Methods for discrete population genetic data*. Sinauer Associates.
- Weir, B.S. and Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38, pp.1358-1370.
- Wenne, R., 2023. Single nucleotide polymorphism markers with applications in conservation and exploitation of aquatic natural populations. *Animals*, 13(6), p.1089.
- Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M., 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, 26, pp.375-400.
- Wright, S., 1922. Systems of mating. I, II. *Genetics*, 6(2), pp.111-178; 6(3), pp.259-294.



University of Fort Hare
Together in Excellence

Wright, S., Gaggiotti, O.E. and Hedrick, P.W., 2015. Genetics and conservation of marine fishes.

Marine Ecology, 36(1), pp.9-22.

Wright, S., 1930. The genetical theory of natural selection: a review. *Journal of Heredity*, 21(8),

pp.349-356.

Zardi, G.I., McQuaid, C.D., Teske, P.R. and Barker, N.P., 2007. Unexpected genetic structure of

mussel populations in South Africa: indigenous *Perna perna* and invasive *Mytilus*

galloprovincialis. *Marine Ecology Progress Series*, 337, pp.135-144



University of Fort Hare
Together in Excellence