DNA barcoding of Euphausiids and Chaetognaths off the East Coast of South Africa

By

Pariksha Govender

BSc. (*Hons*) Genetics

Submitted in fulfillment of the academic requirements for the degree of Master of Science in
the Discipline of Genetics
School of Life Sciences

College of Agriculture, Engineering, and
Science University of KwaZulu-Natal
Pietermaritzburg South Africa

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

Supervisor: Prof. S. Willows-Munro

Signature: [redacted]

Date: 29 May 2025

Supervisor: Dr A. Govender

Signature: [redacted]

Date: 29 May 2025

Preface

The research contained in this dissertation was completed by the candidate while based in the discipline of Genetics, School of Life Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa under the supervision of Prof. S. Willows-Munro and Dr A. Govender-Paul.

These studies represent original work by the candidate and have not otherwise been submitted in any form to another University. Where use has been made of the work by other authors it has been duly acknowledged in the text.

As the candidate's Supervisors I agree to the submission of this dissertation.

Supervisor: Prof. S. Willows-Munro

Signature: [redacted]

Date: 29 May 2025

Supervisor: Dr A Govender

Signature: [redacted]

Date: 29 May 2025

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE

DECLARATION 1 - PLAGIARISM

I, Pariksha Govender, declare that

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

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

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

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

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

[redacted]

Signature: Pariksha Govender

Date: 29 May 2025 *Declaration Plagiarism*

Abstract

DNA barcoding is a well-established method that has been used to augment traditional morphology-based methods to identify species. A 650 base pair region of the mitochondrial cytochrome oxidase I gene (COI) is used as the standard DNA barcode across many animal groups. The reliability of species delimitation by DNA barcoding can be biased by factors such as incomplete reference databases which link sequence data to taxonomically validated species; a lack of regional records for widespread species that may have below-species-level genetic structure; and not enough variable characters in the barcode region to confidently separate closely related species. Unfortunately, only a few studies investigate the accuracy of standard barcode markers, such as COI, in delimiting species within their chosen group of study.

This study focuses on the DNA barcoding of Euphausiids and Chaetognaths. These two phyla form an important component of zooplankton. Zooplankton serve as critical links in marine food chains, contribute to nutrient recycling, and their presence and diversity act as bio-indicators of ecosystem health where changes in their populations can signal changes in environmental conditions. Identifying the species that make up zooplankton is crucial in diversity assessments, biomonitoring, and the detection of invasive species. The overall aim of this MSc study was to determine the utility of COI as a DNA barcode marker for Euphausiid and Chaetognath species. This was done by: (i) analysing all publicly available COI sequence data to evaluate whether there is a statistically significant separation between the maximum genetic differences observed within individuals belonging to the same species and the minimum genetic differences observed between different species. This separation is known as the “DNA barcode gap”. If the DNA barcode gap is present in the COI data for Euphausiids and Chaetognaths, then the COI region can be used to accurately delimit species. If there is overlap within-species and between-species genetic differences, then additional genetic markers or other information are needed to confidently separate species; (ii) generating new DNA barcodes for Euphausiids and Chaetognaths specimens collected from five different localities off the East Coast of South Africa in collaboration with an expert taxonomist to strengthen the current reference libraries for the two groups; (iii) evaluating the completeness of DNA reference libraries for Euphausiids and Chaetognaths by using data from recent metabarcoding studies to determine how many Euphausiid and Chaetognath sequences, amplified from zooplankton samples collected off the East Coast of South Africa, can be confidently assigned (>96% sequence similarity) to reference sequences available on BOLD and GenBank. These analyses are important as the exact number of species found in South African waters is still unclear.

Chapter One provides an overall introduction to topics relevant to the dissertation. This chapter highlights the importance of monitoring the species that make up zooplankton. The chapter also

addresses the taxonomic challenges associated with identifying species within Euphausiids and Chaetognaths. It introduces DNA barcoding as a vital tool to enhance the speed and accuracy of species identification within these two important zooplankton groups. The chapter concludes with an outline of the overall aims of the dissertation.

Chapter Two aims to assess if the standard COI barcode region can statistically delimit Euphausiid and Chaetognath species. This was done by testing for the presence of the DNA barcode gap in the COI data currently available for Euphausiid and Chaetognath species. In this study, I extended this concept to also examine if sequence divergence differences consistently occur between different genera and families belonging to Euphausiids and Chaetognaths. The data analyzed in this chapter came primarily from BOLD and GenBank. Although analyses were conducted separately for Euphausiids and Chaetognaths, the results were very similar. In Chaetognaths, a clear DNA barcode gap was observed, indicating that sequences with more than 3% sequence divergence can be regarded as distinct species. Similarly, in Euphausiids, a clear DNA barcode gap was observed, indicating that sequences with more than 3 - 4% sequence divergence can be regarded as distinct species. Species pairs found in the overlap regions have low divergence values and are phenotypically very difficult to differentiate. Further statistical analyses using the Jeffries-Matusita test was conducted, confirming COI to be successful in separating different species and genera belonging to Chaetognaths and Euphausiids.

In **Chapter Three**, the primary goal was to improve the records of Euphausiid and Chaetognath species from South Africa that are accessible in DNA reference libraries. The secondary goal was to test the utility and accuracy of these reference libraries using metabarcoding data generated in a previous zooplankton metabarcoding study. Euphausiid and Chaetognath specimens collected from five localities off the East Coast of South Africa were sorted and identified to the lowest taxonomic level using morphology-based keys under the supervision of a trained taxonomist. Up to five specimens of each species per locality were used for DNA barcoding. DNA barcodes (COI) were generated for 11 Euphausiid species belonging to four genera and one family, as well as for one species of Chaetognath from one genus and one family. These new records were added to BOLD. This study increased the number of barcodes on BOLD by 118 new records making up of two new DNA barcodes for Chaetognaths and 114 new DNA barcodes for Euphausiids, resulting in an increase in data available from South Africa for these taxa. New species records added to BOLD, include *Euphausia recurva, Thysanoessa gregaria,* and *Sagitta enflata*. To evaluate the impact of these new barcode records on species identification and determine the extensiveness of the DNA reference library, new sequences from the metabarcoding results were matched (>96% sequence similarity) in Euphausiids, at a 77% success rate, with 25 species and one genus identified and for Chaetognaths at a 64% success rate, with 12 species and one genus being identified.

In **Chapter Four** I provide an overall summary of the work conducted towards the MSc. This study statistically tests if the COI sequence data can accurately separate species, genera and families of Euphausiids and Chaetognaths. Both Euphausiids and Chaetognaths play an essential role in marine food chains and are used as ocean health indicators, however, despite their importance, very few COI studies with species identification have been conducted. Euphausiids and Chaetognaths display cryptic morphology, and their taxonomy relies heavily on morphology-based methods of species delimitation, therefore molecular-based methods like DNA barcoding, has the ability to provide faster and more accurate species identification. For Chaetognaths, nuclear genes should be barcoded and combined with COI as they have unwonted patterns of mitochondrial evolution and display extraordinary heterogeneity. Hence COI together with nuclear genes will allow for a more valid and represented DNA barcode database for Chaetognaths. Through DNA barcoding, Euphausiids and Chaetognaths can be assessed for as biological indicators of the overall health of South African marine ecosystems.

Acknowledgments

I would like to express my sincerest gratitude to the following individuals for their unwavering support throughout this journey:

To my supervisors, **Prof. Sandi Willows-Munro,** your guidance, patience, and steadfast belief in me have been invaluable. Your encouragement and support over the years have shaped my growth in more ways than I can express. Thank you for your patience, encouragement, and unwavering belief in my potential.

Dr Ashrenee Govender-Paul, I am deeply appreciative for your insightful guidance and the invaluable skillset you have imparted to me. Your support has not only enriched my research but has also ignited a passion that I hope to pursue further. Thank you for pushing me to improve and for always inspiring me to strive for excellence.

To my husband, **Sagren**, words cannot fully capture the depth of my gratitude. Your unconditional love, patience, and encouragement have been a beacon on the harder days. Thank you for always believing in me, for lifting me up in moments of doubt, and for walking this journey by my side with unwavering support.

To my incredible friends—Ramona, Jamelee, Sabeeha, Esmerelda, Anashree, Rehana, Aunty Sylvie, Maiyuran and Duran—each of you have played an indispensable role in my life and in this achievement. Your encouragement, love, and motivational talks have kept me going on days I felt I couldn’t. I am endlessly grateful for your support, words of wisdom, generous gestures and the laughter that has carried me through this journey.

To my beloved family—you are my heart and my purpose. Without you, none of this would hold meaning. **Poorasen** and **Deshyln**, my precious boys, never stop reaching for the stars. No matter where life takes you, know that I will always be there to catch you and cheer you on. *Always*.

Table of contents

List of Figures

List of Tables

Chapter One:

Literature review and aims of the dissertation

Zooplankton, the name derived from the Greek words "zoon" (animal) and "planktos" (wandering), constitutes a diverse array of floating organisms that are crucial to marine ecosystems. The vast majority of species that form zooplankton are microscopic, uni- or multicellular and exhibit a variety of unique morphological features (Goswami *et al*. 2004; for examples of zooplankton diversity see Figure 1-1). Zooplankton are ecologically important and abundant marine animals in pelagic ecosystems and are classified into two groups based on their life cycle: holoplankton and meroplankton (McGinty *et al.* 2023). Holoplankton spend their lives drifting in the water column as plankton, whilst meroplankton are planktonic only during early stages of their life cycle (Teodosio and Barbosa 2021).

Figure 1-1: A selection of zooplankton species showing varied morphological traits and diversity (photos taken from Bucklin *et al.* 2021).

1.1 Zooplankton as sentinels of environmental change

Zooplankton are essential in nutrient recycling in marine environments as they transfer nutrients from phytoplankton (primary producers) to the higher trophic levels in food chains and help sustain populations of larger marine animals (Atkinson *et al.* 2008; D'Alelio *et al*. 2022; Wang *et al*. 2020). Zooplankton occur at all depths and is widespread across the varied biogeographical regions of the world's oceans. Zooplankton are sensitive to environmental variability, including changes in water temperature (Richardson 2008), nutrient levels (Karpowicz *et al.* 2020), and water quality (Bucklin *et al.* 2004). Furthermore, zooplankton contributes to the biological carbon pump, sequestering carbon dioxide and as a result, lessening the extent and pace of climate change (Richardson 2008).

Zooplankton populations exhibit swift responses to environmental stressors (Sun *et al.* 2023). Tracking changes in zooplankton communities are invaluable as indicators of ecosystem health and in detecting early signs of ecosystem disturbance (Curk *et al.* 2020), as climate change and eutrophication significantly affect their abundance and community structure (Wang *et al.* 2020). Monitoring zooplankton communities, in particular variation in species composition over time (Marques *et al.* 2024), can be used to indicate how marine biodiversity is affected by factors like salinity (Sun *et al.* 2023), low oxygen conditions (Karpowicz *et al.* 2020), climate change and temperature shifts (Richardson 2008). For these reasons the monitoring of zooplankton communities has become a widely accepted and important tool for marine scientists and decision makers (Pitois and Yebra 2022).

1.2. Unstable taxonomy

The taxonomy of many zooplankton taxa is complex, and identification can be difficult due to their small body size, high species richness, fragility and cryptic morphology (Machida *et al*. 2009; Questel *et al.* 2019). The scarcity of diagnostic characteristics particularly for egg and larval stages presents a significant challenge in the interpretation of global and local zooplankton patterns of biodiversity and biogeography (Bucklin *et al*. 2016). Morphological identification of zooplankton species often necessitates the skills of well-trained specialists as their morphology varies drastically as seen in Figure 1-1 (Bucklin *et al.* 2011). This dissertation will focus on two important zooplankton groups, namely, the Euphausiidae and Chaetognatha. These two groups are not well studied in South African waters (Singh *et al*. 2021) but are important as they play an instrumental role in marine food chains (Atkinson *et al.* 2008) and serve as indicators of ocean health (Huntley and Sykes 1989).

1.3. Chaetognatha

Arrow worms (Phylum: Chaetognatha) comprise over 133 species assigned to 26 genera and eight families, globally (Perez *et al.* 2021), with all species inhabiting marine environments and exhibiting hermaphroditic reproduction (World Register of Marine Species (WoRMS) (2025) *Chaetognatha*. Available at : https://www.marinespecies.org/aphia.php?p=taxdetails&id=2081 (Accessed: 15 August 2025). Morphologically Chaetognaths are characterized by a simple body plan, but with abundant intraspecific morphological variation, making species delimitation based on morphological characters challenging (Marletaz *et al.* 2017). Despite this phylum having a low number of species (Atkinson *et al.* 2008), Chaetognaths are abundant in many pelagic environments (Bucklin *et al.* 2021). Their transparent bodies, grasping hooks and fine teeth, Chaetognaths leading predators in many marine food chains (Jennings *et al*. 2010). Due to their fragile bodies, Chaetognaths are not easily identified to species-level, as they are often damaged during sampling (Marlétaz and Parco 2008). DNA barcoding could significantly aid with the identification of cryptic Chaetognath species, as studying the COI gene can reveal genetic differences not evident from morphology alone (Bucklin *et al.* 2021). Globally, 121 Chaetognath species had barcode records, but only 28 DNA barcodes are available from taxa collected in South African waters (Singh *et al*. 2021). The 28 barcodes are mostly one species of Chaetognath (*Sagitta enflata*), one of the seven species which are found in South African waters (Duró and Saiz 2000). This indicates a need for further genetic studies to enhance species identification and ecological understanding.

1.4. Euphausiidae

The Euphausiidae (Phylum: Athrophoda: Crustacea), are among the most essential phyla in most zooplankton assemblages (Bucklin *et al. 2*007) as they serve as a primary food source for larger fish and marine mammals, making them integral to marine food chains. More than 20 species of krill have been recorded in South African waters (Shabangu 2016). Euphausiids constitute nominally 50% of the total zooplankton biomass along South Africa's coast (Gibbons 1995), additionally recent surveys indicate that euphausiids remain as a dominant component of zooplankton found in Southern African and surrounding waters, with species richness being highest at deeper depths and lower along the shore (Hewitt *et al.* 2004; Krafft *et al.* 2021). Furthermore, Singh *et al*. (2021) found that Euphausiids represent 2% of South African DNA barcodes for known species, highlighting a significant gap in genetic data that could aid in species identification and conservation efforts.

Traditionally, microscopic analysis of morphology has been the cornerstone of zooplankton species identification (Verheye *et al.* 2016). The taxonomic keys from Boltovskoy (1999) can be used to identify Euphausiids and Chaetognaths, however, morphology-based identification can still be challenging and limiting as taxonomic keys for the region is sorely lacking for many groups and the ones available are outdated (Koch *et al.* 2022). Due to the low rate of species identification using traditional microscopy, advancements in molecular-based methods of_species identification, such as DNA barcoding have been proposed to augment traditional taxonomic approaches and enhance the accuracy of species identification (Odah 2023).

1.5. The use of DNA barcoding in zooplankton

Molecular data can play an important role in addressing fundamental questions regarding species diversity and their distribution (Bucklin *et al*. 2007). DNA barcoding, a species delimitation method, relies on the cytochrome oxidase 1 (COI) gene as the standard genetic marker utilized to identify species in most animal taxa (Hebert *et al.* 2003). This method provides a powerful complementary approach to traditional species identification based on morphology (Jinbo *et al*. 2011; Smith *et al.* 2005), allowing researchers to assign specimens or samples such as tissue fragments or gut contents to species by sequencing a standardized DNA fragment. DNA barcoding works by comparing

COI sequences generated from an unknown taxon against a reference library containing sequence data linked to taxonomically verified species to facilitate species identification (Bucklin *et al*. 2011; Hebert *et al*. 2003; Laakmann *et al*. 2020). These same reference libraries are used in DNA metabarcoding, which allows for the identification of multiple species simultaneously in mixed samples using high-throughput sequencing methods (Taberlet *et al*. 2012). The global online reference libraries for DNA barcode data are contained in the Barcode of Life Data Systems (BOLD, www.boldsystems.org) and GenBank (www.ncbi.nlm.nih.gov/genbank). These web-based platforms also provide a standardized procedure for species identification that is accessible to researchers who are not specialist taxonomists. Another system for storing data on zooplankton is MetaZooGene (Bucklin *et al.* 2021; https://metazoogene.org/MZGdb) which currently holds barcode data for approximately 8 322 zooplankton species in 2021 (Bucklin *et al.* 2021). An inclusive reference sequence database would be an invaluable tool for scientific investigations, of ocean ecosystem monitoring and assessment as well as to further enhance global metabarcoding studies (Bucklin *et al.* 2021).

DNA barcoding is a widely accepted molecular tool for resolving taxonomic ambiguity, strengthening biodiversity inventories and supporting applications that require accurate species identification (Bucklin *et al.* 2011; Govender *et al.* 2019; Herbert *et al.* 2003). It has been successfully utilized to clarify the taxonomy of many groups of organisms, leading to improved conservation strategies and a deeper understanding of ecological relationships (Ratnasingham and Hebert 2007). Despite DNA barcoding being the most widely used molecular method for species identification, there are several challenges that need to be considered such as incomplete reference libraries, and difficulty in distinguishing closely related species using only a single mitochondrial gene marker (Meyer and Paulay 2005; Pinna *et al*. 2021; Rubinoff *et al*. 2006).

DNA barcoding of Euphausiids using COI has successfully resolved relationships among closely related species within genera (Bucklin *et al.* 2016). Similar successes are observed in Chaetognaths, where COI-based barcode analyses has accurately discriminated among established species (Bucklin *et al*. 2010; Hebert *et al*. 2003). Unfortunately, a vast number of species do not have reference records available on publicly available reference libraries. In particular, Singh *et al.* (2021) reported that only an estimated 13% of the 12 000 described marine faunal species in South Africa had DNA barcodes which were publicly available. Although there are approximately 200 known Chaetognath species globally, there have been 28 records of this group being added to BOLD from South Africa with only one of seven known species found in South Africa (Duró and Saiz 2000). Euphausiids have 49 publicly available DNA barcodes from South Africa (Singh *et al*. 2021), with 23 of 86 species found in South African waters (Brinton *et al*. 2000; Sell *et al*. 2024). Unfortunately, incomplete reference libraries also impact on the ability to detect undescribed species.

Species delimitation through DNA barcoding hinges on the assumption that the genetic distance between different species (interspecies genetic distance) is greater than the genetic distance within the same species (intraspecies genetic distance). The aperture in the distribution of these two genetic distance classes is known as the "DNA barcode gap" (Figure 1-2a). If there is a complete overlap in the interspecific and intraspecific genetic distances (Figure 1-2b), then no DNA barcode gap can be established. However, the presence of a barcode gap is seldomly explicitly tested, highlighting the need for rigorous validation in barcoding studies (Collins and Cruickshank 2013).

The DNA barcode gap has been conducted in Chaetognaths (Marletaz *et al.* 2017); Nair *et al.* 2015) and Euphausiids (Bezeng and van der Bank 2019). These studies have tested the barcode gap at species level only. Our study aims to test for the DNA barcode gap at species level using intraspecific and interspecific genetic distances (Figure 1-2. a), at genus level using congeneric and intergeneric genetic distances and at family level using confamilial and interfamilial generic distances. This study will be using the DNA barcode gap if established, to identify the percentage sequence similarity to distinguish between species, genera and family.

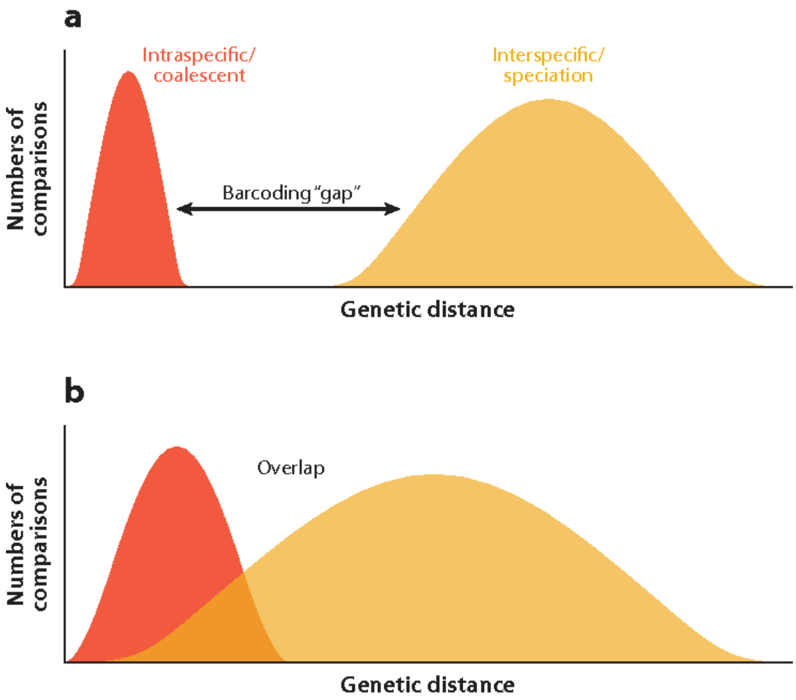

Figure 1-2: DNA "barcode gap" where interspecies genetic distance is greater than intraspecies genetic distance (graphic sourced from Meyer and Paulay 2005)

Lack of explicit testing for the presence of the DNA barcode gap has meant that only a few studies highlight this as a problem. Examples include a study on the *Lycaenidae* (Blue Butterflies), which found that there is an 18% overlap in the intraspecific and interspecific COI sequence genetic distances. This overlap was attributed to low interspecific divergence among closely related species

(Wiemers and Fiedler 2007). Other taxa which display the absence of a DNA barcode gap include species belonging to Mollusca and Annelida *(*Pinna *et al.* 2021). Possible reasons for the lack of resolution may include recent speciation events, the presence of cryptic species diversity, and insufficient sampling across different taxa (Ratnasingham and Hebert 2007; Wiemers and Fiedler 2007). If the COI barcode marker is not able to reliably resolve species, then alternative molecular markers like the mitochondrial 16S rDNA gene have also been utilized for the identification and discrimination of zooplankton species (Marletaz *et al.* 2017). Ali *et al*. (2014), sees barcoding using the COI region as unreliable and prone to errors in identifications due to sequence variability, primer design, amplification bias and incomplete reference libraries. However, more recent data from Bucklin *et al*. (2021) promotes the COI barcode region as useful for species identification and discrimination. While DNA barcoding, particularly through the COI gene, offers a robust framework for species identification and biodiversity studies in many taxa; ongoing efforts are essential to expand reference libraries and validate methodologies to ensure accurate and reliable results (Bucklin *et al.* 2021; Ratnasingham and Hebert 2007).

1.6 Aims of this Masters research

The principal aim of this study was to determine the utility of COI as DNA barcode marker for the identification of Euphausiid and Chaetognath species. This was done by: First, using all available COI sequence data for Chaetognaths and Euphausiids, determining if there is a statistical separation between the maximum within-species, genera or family genetic differences and the minimum between-species, genera or family genetic differences (DNA barcode gap). If a DNA barcode gap is identified, the sequence divergence threshold required to differentiate species, genera, and families for Euphausiids and Chaetognaths will be determined, and if there is an overlap in within- and between- species, genera and family genetic distances then additional genetic markers will be needed to confidently separate species. Second, this study endeavored to generate new DNA barcodes from Euphausiids and Chaetognaths sampled off the East Coast of South Africa in collaboration with a taxonomist. The objective was to add new barcode records to public reference libraries namely, Barcode of Life Database (BOLD) and NCBI GenBank. Third, this study aimed to determine the completeness of DNA reference libraries for Euphausiids and Chaetognaths by using data from recent metabarcoding studies in South Africa to determine how many Euphausiid and Chaetognath sequences amplified from zooplankton samples collected off the East Coast of South Africa can be confidently assigned (>96% sequence similarity) to reference sequences available on BOLD and GenBank.

By understanding the marine biodiversity of South Africa, we will be better equipped to protect and preserve the endemic and other species found in our coastal waters. By learning more about biota and their environment, we can ensure the continuation of species and understanding of marine ecosystems, their relevance and sensitivity to change.

Each chapter of this dissertation has been written in manuscript format. Although some repetition has been included, this has been kept to a minimum.

1.7. References

Ali M.A., Gyulai G., Hidvegi N., Kerti B., Al Hemaid F.M., Pandey A.K. & Lee J. (2014) The changing epitome of species identification–DNA barcoding. *Saudi Journal of Biological Sciences* **21**, 204-231.

Atkinson A., Siegel V., Pakhomov E., Rothery P. & Loeb V. (2008) A re-appraisal of the total biomass and annual production of Antarctic krill. Deep Sea Research Part I: *Oceanographic Research Papers* **55**, 362-373.

Bezeng B.S. & van der Bank H.F. (2019) DNA barcoding of southern African crustaceans reveals a mix of invasive species and potential cryptic diversity. *PLOS ONE* **14**, e0222047.

Boltovskoy D. (1999) South Atlantic Zooplankton Vol. 2. Leiden: Backhuys, 869-1706.

Bucklin A., Nishida S. & Schiel S. (2004) Census of marine zooplankton: A new global survey of marine biodiversity. *Global Change News Letter* **60**, 8-10.

Bucklin A., Wiebe P.H., Smolenack S.B., Copley N.J., Beaudet J.G., Bonner K.G., Färber-Lorda J. & Pierson J.J. (2007) DNA barcodes for species identification of euphausiids (Euphausiacea Crustacea). *Journal of Plankton Research* **29**, 483-493.

Bucklin, A. Jennings, R.M., and Pierrot-Bults, A.C. (2010) Barcoding of Arrow Worms (Phylum Chaetognatha) from Three Oceans: Genetic Diversity and Evolution within an Enigmatic Phylum, *PLoS ONE*, 5(4), p. e9949. doi: 10.1371/journal.pone.0009949

Bucklin A., Steinke D. & Blanco-Bercial L. (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science* **3**, 471-508.

Bucklin A., Lindeque P.K., Rodriguez-Ezpeleta N., Albaina A. & Lehtiniemi M. (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *Journal of Plankton Research* **38**, 393-400.

Bucklin A., Peijnenburg K.T., Kosobokova K.N., O'Brien T.D., Blanco-Bercial L., Cornils A., Falkenhaug T., Hopcroft R.R., Hosia A., Laakmann S. & Li C. (2021) Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology* **168**, 1-26.

Brinton E., Ohman M.D., Townsend A.W., Knight M.D. & Bridgeman A.L. (2000) Euphausids of the World Ocean. ETI. CD-ROM.

Collins R.A. & Cruickshank R.H. (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources* **13**, 969-975.

Curk M., Glavan M. & Pintar M. (2020) Analysis of nitrate pollution pathways on a vulnerable agricultural plain in Slovenia: Taking the local approach to balance ecosystem services of food and water. *Water* **12**, 707.

D'Alelio D., Russo L., Del Gaizo G. & Caputi L. (2022) Plankton under pressure: How water conditions alter the phytoplankton–zooplankton link in coastal lagoons. *Water* **14**, 974.

Duró A. & Saiz E. (2000) Distribution and trophic ecology of chaetognaths in the western Mediterranean in relation to an inshore–offshore gradient. *Journal of Plankton Research* **22**, 339–361.

Gibbons M.J. (1995) Observations on euphausiid assemblages of the south coast of South Africa. *South African Journal of Marine Science* **16**, 141-148.

Goswami S.C., Dhargalkar V.K. & Verlecar X.N. (2004) Zooplankton Methodology Collection & Identification – A Field Manual. Dona Paula Goa: Bioinformatics Centre National Institute of Oceanography.

Govender A, Groeneveld J, Singh S, Willows-Munro S (2019) The design and testing of mini-barcode markers in marine lobsters. PLoS ONE 14(1): e0210492.

Griffiths C.L., Robinson T.B., Lange L. & Mead A. (2010) Marine biodiversity in South Africa: an evaluation of current states of knowledge. *PLOS ONE* **5**, e12008.

Hebert P.D., Cywinska A., Ball S.L. & DeWaard J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B Biological Sciences* **270**, 313-321.

Hewitt R.P., Watkins J.L., Naganobu M., Tshernyshkov P., Brierley A.S., Demed D.A., Kasatkina S., Takao Y., Goss C., Malyshko A., Kawaguchi S. & Yoshida T. (2004) Setting a precautionary catch limit for Antarctic krill. *Oceanography* **15**, 26–33.

Huntley M. & Sykes P. (1989) A study of Chaetognath abundance, body size and feeding in the Atlantic and Antarctic oceans. *Marine Biology* **103**, 153-162.

Jennings R.M., Bucklin A. & Pierrot-Bults A. (2010) Barcoding of arrow worms (Phylum Chaetognatha) from three oceans: genetic diversity and evolution within an enigmatic phylum. *PLOS ONE* **5**, e9949.

Jinbo U., Kato T. & Ito M. (2011) Current progress in DNA barcoding and future implications for entomology. *Entomological Science* **14**, 107-124.

Karpowicz M., Ejsmont-Karabin J., Kozłowska J., Feniova I. & Dzialowski A.R. (2020) Zooplankton community responses to oxygen stress. *Water* **12**, 706.

Koch, W., Elven, H. and Finstad, A.G. (2022) Clavis: An open and versatile identification key format, *PLOS ONE*, 17(12), p. e0277752. doi: 10.1371/journal.pone.0277752.

Krafft B.A., Fielding S., Brierley A.S., Tarling G.A., Hill S.L., Reiss C.S., Kawaguchi S., Murphy E.J., Atkinson A., Nicol S. & Hosie G.W. (2021) Distribution and biomass estimation of Antarctic krill (Euphausia superba) in the South Orkney Islands region from 2011 to 2020. ICES *Journal of Marine Science* **80**, 1472–1488.

Laakmann S., Blanco-Bercial L. & Cornils A. (2020) The crossover from microscopy to genes in marine diversity: from species to assemblages in marine pelagic copepods. *Philosophical Transactions of the Royal Society B* **375**, 20190446.

Machida R.J., Hashiguchi Y., Nishida M. & Nishida S. (2009) Zooplankton diversity analysis through single-gene sequencing of a community sample. BMC *Genomics* **10**, 1-7.

Marlétaz F. & Parco Y. (2008) Careful with understudied phyla: The case of chaetognath. BMC *Evolutionary Biology* **8**, 251.

Marletaz F., Le Parco Y., Liu S. & Peijnenburg K.T. (2017) Extreme mitogenomic variation in natural populations of Chaetognaths. *Genome Biology and Evolution* **9**, 1374-1384.

Marques R., Otto S.A., Di Pane J., Boersma M., Meunier C.L., Wiltshire K.H., Möllmann C. & Renz J. (2024) Response of the meso- and macro-zooplankton community to long-term environmental changes in the southern North Sea. ICES *Journal of Marine Science* **81**, 526–539.

McGinty N., Irwin A.J., Finkel Z.V. & Dutkiewicz S. (2023) Using ecological partitions to assess zooplankton biogeography and seasonality. *Frontiers in Marine Science* **10**, 989770.

Meyer C.P. & Paulay G. (2005) DNA barcoding: Error rates based on comprehensive sampling. PLOS *Biology* **3**, e422.

Nair V.R., Kidangan F.X., Prabhu R.G., Bucklin A. & Nair S. (2015) DNA barcode of Chaetognatha from Indian Waters. *Indian Journal of Geo-Marine Sciences* **44**, 1366–1376.

Odah M.A.A. (2023) Unlocking the genetic code: Exploring the potential of DNA barcoding for biodiversity assessment. AIMS *Molecular Science* **10**, 263–294.

Perez Y., Müller C.H.G. & Harzsch S. (2021) Chapter 15: Chaetognatha. In: Schierwater B. & DeSalle R. (eds.) Invertebrate Zoology: A Tree of Life Approach. CRC Press, 231.

Pinna M., Saccomanno B., Marini G., Zangaro F., Kabayeva A., Khalaj M., Shaimardan L., D'Attis S., Tzafesta E. & Specchia V. (2021) Testing the influence of incomplete DNA barcode libraries on ecological status assessment of Mediterranean transitional waters. *Biology* (Basel) **10**, 1092.

Pitois S. & Yebra L. (2022) Contribution of marine zooplankton time series to the United Nations Decade of Ocean Science for Sustainable Development. ICES *Journal of Marine Science* **79**, 722-726.

Questel J.M., Hopcroft R.R., DeHart H.M., Smoot C.A., Kosobokova K.N. & Bucklin A. (2021) Metabarcoding of zooplankton diversity within the Chukchi Borderland Arctic Ocean: improved resolution from multi-gene markers and region-specific DNA databases. *Marine Biodiversity* **51**, 1-19.

Ratnasingham S. & Hebert P.D.N. (2007) BOLD: The Barcode of Life Data System. *Molecular Ecology Notes* **7**, 355-364.

Richardson A.J. (2008) In hot water: zooplankton and climate change. ICES *Journal of Marine*

Science **65**, 279-295.

Rubinoff D., Cameron S. & Will K. (2006) A genomic perspective on the shortcomings of mitochondrial DNA for barcoding identification. *Journal of Heredity* **97**, 581-594.

Sell A.F., von Maltitz G.P., Auel H., Biastoch A., Bode-Dalby M., Brandt P., Duncan S.E., Ekau W., Fock H.O., Hagen W., Huggett J.A., Koppelmann R., Körner M., Lahajnar N., Martin B., Midgley G.F., Rixen T., van der Lingen C.D., Verheye H.M. & Wilhelm M.R. (2024) Unique Southern African Terrestrial and Oceanic Biomes and Their Relation to Steep Environmental Gradients. *Sustainability of Southern African Ecosystems in Ecological Studies* **248**.

Shabangu F., Somhlaba S. & Leslie R. (2016) Should South Africa harvest krill? *Fishing Industry News Southern Africa* **17,** 18-20.

Singh S.P., Groeneveld J.C., Huggett J., Naidoo D., Cedras R. & Willows-Munro S. (2021) Metabarcoding of marine zooplankton in South Africa. *African Journal of Marine Science* **1**,1-13.

Smith M.A., Fisher B.L. & Hebert P.D.N. (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society B: *Biological Sciences* **360**, 1825-1834.

Sun X., Zhang H., Wang Z., Huang T., Tian W. & Huang H. (2023) Responses of Zooplankton Community Pattern to Environmental Factors along the Salinity Gradient in a Seagoing River in Tianjin, China. *Microorganisms* **11**, 1638.

Taberlet P., Coissac E., Pompanon F., Brochmann C. & Willerslev E. (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* **21**, 2045-2050.

Teodósio M.A. & Barbosa A.M.B. (eds.) (2021) Zooplankton ecology. CRC Press Taylor & Francis Group.

Verheye H.M., Lamont T., Huggett J.A., Kreiner A. & Hampton I. (2016) Plankton productivity of the Benguela current large marine ecosystem (BCLME). *Environmental Development* **17**, 75-92.

Wang R., Liu X., Wu J., Wai T.-C., Shen P. & Lam P.K.S. (2020) Long-term variations of phytoplankton community in relation to environmental factors in Deep Bay China from 1994 to 2016. *Marine Pollution Bulletin* **153**, 111010.

Wiemers M. & Fiedler K. (2007) Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera Lycaenidae). *Frontiers in Zoology* **4**, 8.

World Register of Marine Species (WoRMS) (2025) *Chaetognatha*. Available at : https://www.marinespecies.org/aphia.php?p=taxdetails&id=2081 (Accessed: 21 August 2025).

Chapter Two

Examining the DNA barcode gap in two abundant zooplankton taxa: Euphausiids and Chaetognaths

Abstract

DNA barcoding is an effective tool for swift species identification. While it has been successfully used to identify many species, the overlap of interspecific and intraspecific genetic distances in some taxa can lead to unreliable species identification. Few studies explicitly assess for the presence of a distinct "barcode gap" separating interspecific and intraspecific genetic distances in target taxa. This study examines the utility of the mitochondrial cytochrome oxidase I (COI) barcode marker in distinguishing species, genera and family boundaries in Euphausiids and Chaetognaths by statistically testing for a DNA barcode gap. As key bio-indicators with significant roles in marine ecosystems, Euphausiids and Chaetognaths are ecologically important. Despite their ecological significance and global distribution, species identification in these taxa has been challenging due to morphological similarities and unreliable diagnostic characteristics. A statistically significant DNA barcode gap was observed for Euphausiid species (2-3% sequence divergence; Jeffries-Matusita value =1.88) and genera (5-6% sequence divergence; J-M= 1.99). At higher taxonomic levels, confamilial and interfamilial genetic distances slightly overlapped but remained statistically separable (8-10% sequence divergence J-M value = 1.99). For Chaetognaths, overlaps were present at all taxonomic levels, though a statistical barcode gap was evident at the species (4-5% divergence; J-M = 1.80) and genus levels (7-8% divergence; J-M = 1.80). No clear barcode gap was found at the family level. These results confirm that the COI gene can effectively separate species and genera in Euphausiids and Chaetognaths, supporting its application in rapid, automated species identification.

2.1. Introduction

Zooplankton play a pivotal role in marine ecosystems as primary consumers and also facilitate energy transfer for higher trophic levels (Atkinson *et al.* 2008). Zooplankton serve as sentinel taxa, with species turnover reflecting changes in the environmental such as nutrient availability, salinity, and temperature (Beaugrand *et al.* 2002; Halsband and Hirche 2001; Miller *et al.* 2004). For example, variation in water temperature affect metabolic rates and reproductive cycles of some species of zooplankton, leading to changes in population dynamics and community composition (Beaugrand *et al.* 2002). Similarly, fluctuations in salinity influence osmoregulatory processes, impacting the disperal and abundance of species leading to changes in zooplankton communities (Halsband and Hirche 2001). Nutrient availability shapes zooplankton species richness by directly affecting primary food sources also leading to shifts in community composition (Miller *et al.* 2004). Climate change is reshaping ocean ecosystems, making it crucial to monitor zooplankton communities as indicators of biodiversity shifts (Richardson 2008; Ratnarajah *et al.* 2023). DNA barcoding offers a rapid means of species identification, that can be used to detect changes in zooplankton communities enhancing biodiversity monitoring and conservation efforts (Bozelli *et al.* 2021; Duarte *et al.* 2023; Hebert *et al.* 2003; Imtiaz *et al*. 2017).

2.1.1. DNA barcoding as a species identification tool

DNA barcoding entails the sequencing of a short, standardized region of DNA from a sample and comparing this DNA sequence to a reference library containing sequences linked to taxonomically verified species (Hebert *et al.* 2003; Imtiaz *et al*. 2017). A portion of the mitochondrial cytochrome c oxidase subunit 1 (COI), gene which is approximately 650 base pairs (bp) in length, is regarded as the standard DNA barcode region in most animal taxa (Hebert *et al.* 2003).

The use of COI has many advantages as the COI barcode region displays appropriate levels of interspecific genetic variation, allowing for effective species-level identification even among closely related species (Andújar *et al.* 2018; Deagle *et al.* 2014; Imtiaz *et al*. 2017). The COI gene has a higher amplification rate compared to other molecular markers as the region exhibits lower intragenomic variability and is less likely to vary within a species (Cheng *et al.* 2013; Girard *et al.* 2022; Lv *et al.* 2014). The high amplification rate of the COI gene together with well curated reference libraries, makes COI the ideal marker for DNA barcoding and species identification, despite limitations associated with the use of a single mitochondrial marker (Deagle *et al.* 2014). DNA barcoding can be affected by many different factors. Primers may bind preferentially to certain sequences, resulting in skewed species richness assessments (Anderson *et al.* 2023; van der Loos and Nijland 2021). The analysis of sequence data can also be biased by misinterpretation of biodiversity (Alon *et al*. 2011). Furthermore, sequencing errors or critical information missing from databases, may lead to the misidentification of species (Porter and Hajibabaei 2020).

Reference libraries contain sequence data linked to morphological voucher specimens verified by expert taxonomists (Bucklin *et al*. 2007; Ceasri *et al.* 2013; Tempestini *et al*. 2018; Tzafesta *et al.* 2022). These libraries are essential to obtain accurate results from DNA barcoding analyses. Cryptic species, taxonomic synonymy, and genetically distinct populations can be better understood and resolved through DNA barcoding if comprehensive regional reference libraries are established (Bucklin *et al.* 2007). A recent study by Bucklin *et al.* (2021) aimed to create a COI reference library for global marine zooplankton (MetaZooGene) by compiling data from BOLD and Genbank. In their findings, they successfully reported the proportion of species represented by the COI barcodes available in MetaZooGene. 66 out of 86 (76%) Euphausiid species and only 29 out of 200 (15%) specimens for Chaetognaths (Bucklin *et al*. 2021), this suggests that Chaetognaths are underrepresented in barcode libraries compared to Euphausiids. Bucklin *et al.* 2021 further states that other factors such as recent speciation events, cryptic species diversity and inadequate sampling contribute to gaps in barcode databases.

The limited availability of COI sequences for Chaetognaths in publicly available reference libraries can pose challenges in species identification using DNA barcoding (DeHart *et al.* 2020). Thus, greater sampling and sequencing efforts are needed for this taxon. While DNA barcoding has garnered recognition as a valuable tool in the field of biology, it is essential to note that relatively few studies have explored its application in Euphausiids (Bucklin *et al*. 2007; Deagle *et al.* 2019; Ortman *et al.* 2010) and Chaetognaths (Bucklin *et al.* 2007; Jennings *et al.* 2010; Kearse *et al.* 2012; Nair *et al.* 2015). Currently, there are only 114 species records for Euphausiids on BOLD; 49 with DNA barcodes (accessed on 27 May 2025) with 23 of 86 species found in South African waters (Brinton *et al*. 2000; Gibbons1995). For Chaetognaths, 23 species records are on BOLD (accessed on 27 May 2025), with one of which 7 species that occur in South Africa (Duró and Saiz 2000).

2.1.2. The DNA barcode gap

The concept of the DNA barcode gap, defined as the difference between intraspecific and interspecific genetic distances within a group of organisms (Barett and Hebert 2005; Meyer and Paulay 2005), plays a pivotal role in DNA barcoding. The presence of a DNA barcode gap allows sequences to be accurately assigned to species using a predetermined sequence similarity threshold. In many studies this is automatically set at 2–3% sequence divergence, as this threshold has been found to distinguish between species within most taxa (Barrett and Hebert, 2005; Hebert *et al.,* 2003; Chapple and Ritchie, 2013; Meyer and Paulay 2005; Radulovici *et al.*, 2010). In some species there is an overlap between intraspecific and interspecific genetic distances i.e. the barcode gap is not present in the data, examples include Diptera (Arthropoda: Insecta) and Asteropyrum (Kingdom Plantae) (Cheng *et al.* 2021). Other examples of groups that don’t have a DNA barcode gap include blue butterflies (Wiemers and Fiedler 2007), certain rodents e.g. *Rattus tanezumi, Mus musculus* (Ojeda *et al.* 2022) 2020) and selected bat species e.g. *Minipoterus natalensis* (Taylor *et al.* 2016) and *Miniopterus fraterculus* (Miller-Butterworth *et al.* 2005). This emphasizes the complexity of using DNA barcoding in species delimitation, as some species may be separated by very shallow genetic distances (Hebert *et al.* 2003), while being morphologically, geologically and ecologically distinct (Wiens and Penkrot 2002). The lack of a DNA barcode gap could also be due to the mixing of genetic material through introgression (Moraes *et al.* 2024). Recently diverged species that exhibit minimal genetic differentiation also makes them difficult to separate via barcoding alone (Kekkonen *et al.* 2015). Hence, the lack of a barcode gap requires the use of additional genetic markers or integrative taxonomic approaches to ensure accurate species identification.

2.1.3. DNA barcoding of Euphausiid and Chaetognath species

Both Euphausiids and Chaetognaths are marine invertebrates characterized by their distinctive body structures. Euphausiids are characterized by elongated, segmented bodies, with transparent exoskeletons and a pair of large compound eyes, while Chaetognaths are characterized by their long transparent slender bodies (Javidpour *et al*. 2009; Jennings *et al*. 2010). Both taxa play an integral role in food webs and serve as indicators of ocean health, by responding sensitively to changes in temperature and nutrient availability (Atkinson *et al.* 2008; Huntley and Sykes 1989; Nicol 2006). Chaetognaths have weak osmoregulatory functions, as such each species has developed different environmental tolerances and preferences, occupying narrow salinity ranges (Ulloa *et al*. 2000).

The taxonomy of Euphausiids and Chaetognaths relies heavily on morphology-based method of species delimitation (Bucklin *et al*. 2007; Bucklin *et al.* 2021; Gasmi *et al*. 2014). This can be time-consuming and challenging for non-expert taxonomists (Bucklin *et al.* 2007; Raupach and Radulovici 2015). Due to simple variations in body structure fragile body parts are often damaged and degraded when stored (Bucklin *et al.* 2021; Gasmi *et al*. 2014; Johnson and Allen 2012; Stone and Stone 2023). Additionally, many different life stages make species identification based on morphology alone extremely difficult for these two taxa (Johnson and Allen 2012), as they display varied morphology due to alterations in environmental conditions (Nair *et al* .2015)

Previous studies have used COI DNA barcoding to successfully distinguish between different Euphausiid (Bucklin *et al*. 2007; Deagle *et al.* 2019: Ortman *et al.* 2010) and Chaetognath species (Bucklin *et al.* 2007; Jennings *et al.* 2010; Kearse *et al.* 2012; Nair *et al.* 2015). COI has been used to assess genetic diversity within species, resolving cryptic species and exploring geographical patterns for species (Bucklin *et al*. 2007; Gasmi *et al*. 2014; Peijnenburg *et al*. 2004). In a study by Bezeng and van der Bank (2019) the presence of a DNA barcode gap was established using COI data for crustaceans and included only 28 Euphausiid species. There were also five Chaetognath species used in the Bezeng and van der Bank (2019) study. The presence of a DNA barcode gap was established at 95% sequence similarity threshold and species identification was performed using the Kimura 2-parameter (K2P) model. The study by Bezeng and van der Bank (2019) focussed on a few Euphausiid and

Chaetognath specimens, however this could be taken a step further to include a comprehensive list of each taxon as used in this study.

2.1.4. Aims of the study:

This study aims to statistically determine the presence of a DNA barcode gap in Euphausiid and Chaetognath COI data and if a DNA gap is present at the species, genus and family levels. If a DNA barcode gap is identified, the sequence divergence threshold required to differentiate species, genera, and families for Euphausiids and Chaetognaths will be determined. This approach will enhance species identification reliability of these important zooplankton groups using DNA barcoding, as well as contribute to a better understanding of these taxa.

2.2. Materials and methods

2.2.1 Sampling

A total of 130 Euphausiid COI sequences (65 species, 11 genera and 2 families) and 50 Chaetognath COI sequences (27 species, 13 genera and 6 families) were downloaded from GenBank and BOLD (date accessed: 1 December 2023). The Euphausiid alignment includes 76% of all known species and 100% of families and genera were represented, while the Chaetognath alignment accounts for 23% of all known species, 50% of genera, and 75% of families. Each species was represented by two sequences. To accommodate phylogeographic structure within recognised species, where possible, sequences from different geographical regions were selected. Euphausiid and Chaetognath sequences were aligned separately using ClustalW (Larkin *et al*. 2007). Alignments were manually optimized using Bioedit (Hall 1994) to ensure homology. MEGA 5.0 (Tamura *et al*. 2011) was used to estimate the number of variable characters and average nucleotide composition for each alignment.

2.2.2. DNA barcode gap analysis

Using the Kimura 2-Parameter model (K2P; Kimura 1980), genetic distances were calculated in MEGA 5.0 (Tamura *et al*. 2011). This substitution model, used by BOLD, is widely used in the DNA barcoding community (Antil *et al*. 2023; Hebert *et al.* 2003; Meyer and Paulay 2005) and was used in this study for standardization (Collins and Cruickshank. 2013). Genetic distances were grouped into the different categories: intraspecific distances (genetic distance between two individuals that belong to the same species), interspecific distance (genetic distance between individuals belonging to different species), congeneric (species belonging to the same genus), intergeneric (species belonging to different genera), confamilial (species belonging to the same family) and interfamilial (species belonging to different families) for both the Euphausiid and Chaetognath datasets.

To determine if the genetic distances of the different categories were statistically separable (i.e. that a DNA barcode gap is present) the Jeffries-Matusita (J-M) distance function (Trigg and Flasse 2001) was calculated in R-studio (http://www.r-project.org). The J-M test is a good statistical measure to differentiate between different categories, as this test considers the distance between the class mean and the distribution of values from the mean. A J-M distance value of 1.414 or greater indicates the two genetic distance categories are statistically separable (Trigg and Flasse 2001). This statistic has been used previously to determine the presence of a DNA barcode gap in other studies (Govender *et al.* 2019; Phillips *et al.* 2024). The genetic distances inferred from the different categories were plotted in Microsoft Excel (2016). Threshold values were estimated by taking the difference between the genetic distances of the different classes.

2.3. Results

The final Euphausiid alignment contained 639 base pairs (bp) with an average base composition of 66% adenine and thymine, and 34% guanine and cytosine. The Chaetognath alignment contained 627 bp, and an average base composition of 59% adenine and thymine and 41% guanine and cytosine. No stop codons, deletions or insertions were detected for both the Euphausiid and Chaetognath COI alignments.

2.3.1. DNA barcode gap analysis

For Euphausiid species, the intraspecific K2P distance ranged from 0.02 to 0.03, while the interspecific distances ranged from 0.07 to 0.32. No overlap in genetic distances was observed (Figure 2-1A) which was further supported by a J-M distance value of 1.88, confirming statistically separability of these two genetic distance classes. Congeneric K2P distances ranged from 0.05 to 0.06, while intergeneric distances ranged from 0.11 to 0.32. No overlap was observed among genera (Figure 2-1C), with a corresponding J-M value of 1.99. At the family level, confamilial K2P distances ranged from 0.08 to 0.10, while interfamilial distances ranged from 0.18 to 0.32. A slight overlap in genetic distance categories was observed between 0.19 and 0.24 (Figure 2-1E), with a J-M value of 1.99, indicating statistical significance.

For Chaetognath species, intraspecific K2P distances ranged from 0.04 to 0.05, while the interspecific distances ranged from 0.22 to 0.51. A minor overlap between intra- and interspecific genetic distances was observed between 0.01 and 0.03 (Figure 2-1B), with the J-M function returning a value of 1.80, indicating statistically separable distance classes. Congeneric K2P distances ranged from 0.07 to 0.08, while intergeneric ranging from 0.22 to 0.51. A slight overlap between the intra- and intergeneric genetic distances occurred between 0.22 to 0.24 species (Figure 2-1D), with a J-M value of 1.80. At the family level, the K2P distances showed a complete overlap between confamilial and interfamilial distances (Figure 2-1F), with a J-M value of 1.80.

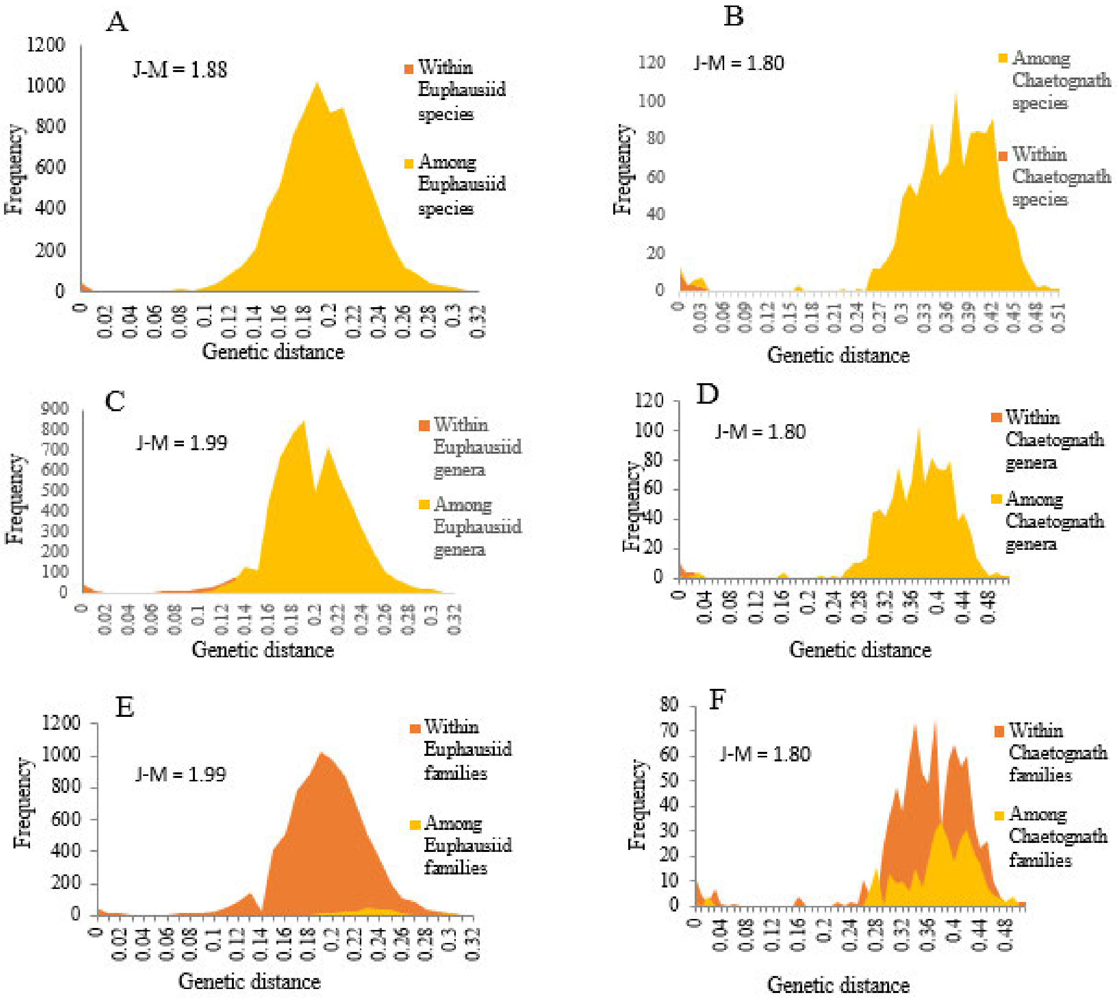

Figure 2-1: Frequency distributions of pairwise K2P genetic distances for Euphausiids and Chaetognaths in COI DNA barcode gap analyses. The first column represents Euphausiids: (A) intra- and interspecific, (C) congeneric and intergeneric, and (E) confamilial and interfamilial distances. The second column represents Chaetognaths: (B) intra- and interspecific, (D) congeneric and intergeneric, and (F) confamilial and interfamilial distances. The distributions were computed using genetic distances for species (A, B), genus (C, D), and family (E, F) in the respective datasets.

2.4. Discussion

DNA barcoding has emerged as a rapid and accurate method for species identification (Antil *et al*. 2023; Hebert and Gregory 2005). However, its success relies on well curated reference libraries available for the expected taxa (Bucklin *et al*. 2021; Hebert *et al* 2003; Kress and Erikson 2007) and the presence of a distinct DNA barcode gap, with minimal overlap between intraspecific and interspecific genetic distances (Meyer and Paulay 2005). This can be particularly useful in cases where using morphological characters alone is inadequate for species identification, such as in this study of species belonging to the ecologically important zooplankton groups of Euphausiids and Chaetognaths being difficult to distinguish morphologically (Bucklin *et al.* 2007). Overlaps between the interspecific and intraspecific genetic distances diminishes the DNA barcode gap, making it more difficult to establish reliable threshold values for species identification (Kelly *et al.* 2019; Phillips *et al.* 2022), which leads to potential misidentification (Deagle *et al.* 2014).

While previous studies have used DNA barcoding to delineate Euphausiid and Chaetognath species (Bucklin *et al*. 2007; Jennings *et al.* 2010; Tempestini *et al*. 2018), no study has experimentally tested for the presence of the DNA barcode gap beyond the species-level. Using all available COI data from GenBank and BOLD, this study provides evidence that the COI gene region is a successful DNA barcode marker for identification in Euphausiids and Chaetognaths. The results obtained from the DNA barcode gap analyses found that the COI gene region is an effective marker for identification of Euphausiid and Chaetognath at the species-level and genus-level as significant DNA barcode gaps was observed in both groups in this study. The J-M distance for species, genus and family levels for Euphausiid and Chaetognath were all greater than 1.414 which suggests the genetic distances at the species, genus and family levels for Euphausiid and Chaetognath are separable.

The DNA barcode gap results at species level from this study are are consistent with previous reports. For example, Jennings *et al.* (2010), tested for the barcode gap at species-level using only their generated Chaetognath COI sequences and found it to be present in their limited data. Bucklin *et al.* (2007) have been successful in species identification of Euphausiids using the COI gene region. Tempestini *et al*. (2018) used the COI gene region for species identification successfully but did not explicitly test for a DNA barcode gap.

Macher *et al*. (2023) suggested general sequence similarity thresholds that can be applicable to a vast range of taxa. These were sequence similarity values ≥ 97% can be used to separate species, ≥ 95% can be used to differentiate different genera, ≥ 90% separate different families and ≥ 85% can be used to separate different orders. From this study more precise values were obtained. In this study different Euphausiid species are separable at sequence similarity > 97-98 %, > 94-95 % for genera, and > 90-92% for families. DNA barcode gap analysis in Chaetognaths, revealed sequence similarity at >

94-95% for species-level identification, and 92-93% for genus-level separation. The latter values are quite different from the general threshold values used and highlights the importance of conducting DNA barcode gap analyses on different taxa.

At family-level for Chaetognaths, DNA barcode analyses revealed an overlap between confamilial and interfamilial genetic distances. This observation suggests that distinguishing between individuals of the same family (confamilial) and those of different families (interfamilial) within these groups may be challenging when relying solely on the COI gene. This may be due to misidentification of species used in our alignment (Wei *et al*. 2018), as Chaetognaths are difficult to identify, due to their delicate body structure. This emphasises the dire need for more extensive taxonomic studies on this group. The overlap in genetic distances may be expected in Chaetognath due to their broad geographic distribution and diverse environmental conditions (Wei *et al*. 2018) and in Euphausiids, is attributed due to high levels of gene flow due to their larval stages, which are free-swimming and can disperse over large distances (Bucklin *et al.* 2021; Wei *et al*. 2018). Specimen pairs found in the overlap regions display low divergence values and are phenotypically very hard to differentiate as both Euphausiid and Chaetognath display significant morphological plasticity influenced by environmental conditions (DeHart *et al.* 2020).

The presence or absence of a DNA barcode gap carries significant implications for the field of taxonomy and species delimitation. In Euphausiids and Chaetognaths, where an overlap in genetic distances has been observed, further research is imperative. More particularly in barcode gap analyses at family level. This overlap reflects the intricacies of species delimitation within these groups, highlighting the need for a multi-disciplinary approach that integrates genetic data with other characters and ecological information.

2.5. Conclusion

The results from this study shows that the COI gene region is effective for accurate species identification of Euphausiids and Chaetognaths with a sequence similarity threshold of > 97%. The standard threshold for using the COI gene region to separate genera and family for Euphausiids are also in line with this study as 95 % sequence similarity for genera and 90 % sequence similarity for family, despite the overlap of genetic distance classes observed. For Chaetognaths, the threshold range using COI gene region to separate genera is 92%, which is below what recent literature cites (Macher *et al.* 2023). The observed overlap in genetic distances particularly at genera and family levels within Chaetognaths necessitates further examination, particularly at family level. Other genes, such as nuclear genes are to be included with the COI gene when DNA barcoding Chaetognaths, as Bucklin *et al.* (2021) has noted that Chaetognaths have atypical mitochondrial evolution and extreme levels of mitochondrial variation. Hence, integrative taxonomic approaches (Phillips *et al.,* 2022 Puillandre *et al.,* 2012; Zhang *et al.,* 2013) such as combining multiple genetic markers, and statistical testing to verify the gaps and thresholds unique to taxa is the way forward for Chaetognaths.

2.6. References

Alon, S., Vigneault, F., Eminaga, S., Christodoulou, D.C., Seidman, J.G., Church, G.M., and Eisenberg, E. (2011) Barcoding bias in high-throughput multiplex sequencing of miRNA, *Genome Research*, 21(9), pp. 1506–1511. doi: 10.1101/gr.121715.111.

Anderson, G., Macdonald, J.I., Potts, J., Feutry, P., Grewe, P.M., Boutigny, M., Davies, C.R., Muir, J.A., Roupsard, F., Sanchez, C., and Nicol, S.J. (2023) Evaluating DNA cross-contamination risk using different tissue sampling procedures on board fishing and research vessels, *ICES Journal of Marine Science*, 80(4), pp. 728–738. doi: 10.1093/icesjms/fsad004.

Andújar, C., Arribas, P., Yu, D.W., Vogler, A.P. and Emerson, B.C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa, *Molecular Ecology*, Available at: https://doi.org/10.1111/mec.14844

Atkinson A., Siegel V., Pakhomov E., Rothery P. & Loeb V. (2008) A re-appraisal of the total biomass and annual production of Antarctic krill. Deep Sea Research Part I: *Oceanographic Research Papers* **55**, 362-373.

Antil, V., Sharma, M., & Kumar, S. (2023). Advances in DNA barcoding and species identification: a focus on substitution models. Gene Reports, 33, 101665

Barrett, R.D.H. and Hebert, P.D.N. (2005) Identifying spiders through DNA barcodes, *Canadian Journal of Zoology*, 83, pp. 481–491.Beaugrand, G., Reid, P.C., Ibanez, F., Lindley, J.A. and Edwards, M., 2002. Reorganization of North Atlantic marine copepod biodiversity and climate. Science, 296(5573), pp.1692-1694.

Beaugrand, G., Reid, P.C., Ibanez, F., Lindley, J.A. and Edwards, M., 2002. Reorganization of North Atlantic marine copepod biodiversity and climate. Science, 296(5573), pp.1692-1694

Bezeng B.S. & van der Bank H.F. (2019) DNA barcoding of southern African crustaceans reveals a mix of invasive species and potential cryptic diversity. *PLOS ONE* **14**, e0222047.

Bozelli, R.L., Costa, M.P., Araujo-Lima, C.A., Nogueira, M.G., Monteiro, L.D., McClain, M.E., Chaves, S.A. and dos Santos, M. (2021) Changes and drivers of zooplankton diversity patterns in the middle reach of Yangtze River floodplain lakes, *Frontiers in Marine Science*, 8, Article 8717274. Available at: https://doi.org/10.3389/fmars.2021.8717274

Brinton E., Ohman M.D., Townsend A.W., Knight M.D. & Bridgeman A.L. (2000) Euphausids of the World Ocean. ETI. CD-ROM.

Bucklin A., Wiebe P.H., Smolenack S.B., Copley N.J., Beaudet J.G., Bonner K.G., Färber-Lorda J. &

Pierson J.J. (2007) DNA barcodes for species identification of euphausiids (Euphausiacea Crustacea). *Journal of Plankton Research* **29**, 483-493.

Bucklin A., Steinke D. & Blanco-Bercial L. (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science* **3**, 471-508.

Bucklin A., Lindeque P.K., Rodriguez-Ezpeleta N., Albaina A. & Lehtiniemi M. (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *Journal of Plankton Research* **38**, 393-400.

Bucklin A., Peijnenburg K.T., Kosobokova K.N., O'Brien T.D., Blanco-Bercial L., Cornils A., Falkenhaug T., Hopcroft R.R., Hosia A., Laakmann S. & Li C. (2021) Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology* **168**, 1-26.

Cesari, M., Guidetti, R., Rebecchi, L., Giovannini, I. & Bertolani, R. (2013) A DNA barcoding approach in the study of tardigrades', *Journal of Limnology*, 72(1), pp. 182–198.

Chapple DG, Ritchie PA (2013) A Retrospective Approach to Testing the DNA Barcoding Method. PLoS ONE 8(11): e77882. doi:10.1371/ journal.pone.0077882

Cheng, F., Wang, M., Sun, S., Li, C. and Zhang, Y., (2013). DNA barcoding of Antarctic marine zooplankton for species identification and recognition. Adv Polar Sci, 24(2), pp.119-127

Cheng, S., Zeng, W., Wang, J., Liu, L., Liang, H., Kou, Y., Wang, H., Fan, D. and Zhang, Z. (2021) Species delimitation of *Asteropyrum* (Ranunculaceae) based on morphological, molecular, and ecological variation, *Frontiers in Plant Science*, 12, 681864. https://doi.org/10.3389/fpls.2021.681864.

Collins R.A. & Cruickshank R.H. (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources* **13**, 969-975.

Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F. and Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match, *Biology Letters*, 10(9), pp. 20140562. doi: 10.1098/rsbl.2014.0562. PMID: 25209199; PMCID: PMC4190964.

Deagle, B.E., Gales, N.J., and Evans, K.L. (2019) DNA barcoding of southern African crustaceans reveals a new species of euphausiid, *PLOS ONE*, 14(9), e0222047. doi:10.1371/journal.pone.0222047.

Décima, M. (2022). Zooplankton trophic structure and ecosystem productivity. *Marine Ecology Progress Series*, 692, 23-42. DOI: 10.3354/meps13920

DeHart, H.M., Blanco-Bercial, L., Passacantando, M., Questel, J.M. and Bucklin, A. (2020). Pathways of pelagic connectivity: *Eukrohnia hamata* (*Chaetognatha*) in the Arctic Ocean, *Frontiers in Marine Science*, 7. doi: 10.3389/fmars.2020.00396.

Duarte, C.M., Weitz, S., Agustí, S., Alcaraz, M., Andrade, S., Badía, M., Bode, A., Calvo, E., Ceballos, S., Daufresne, M., Figueiras, F.G., Gasol, J.M., Gómez, M., Hernández-León, S., Irigoien, X., Juanes, J.A., Martínez, J., Navarro, J., Poulet, S.A., Sarmento, H., Sauter, A., Tassin, B., Teixeira, H., Valente, L. and Zupo, V. (2023) Climate-driven zooplankton shifts cause large-scale declines in food quality for fish, *Nature Climate Change*, 13(3), pp. 232-239. Available at: https://doi.org/10.1038/s41558-023-01630-7

Duró A. & Saiz E. (2000) Distribution and trophic ecology of chaetognaths in the western Mediterranean in relation to an inshore–offshore gradient. *Journal of Plankton Research* **22**, 339–361.

Gasmi, S., Nève, G., Pech, N., Tekaya, S., Gilles, A. and Perez, Y. (2014). Evolutionary history of *Chaetognatha* inferred from molecular and morphological data: a case study for body plan simplification, *Frontiers in Zoology*, 11(84). https://doi.org/10.1186/s12983-014-0084-

Gibbons, M.J., (1995) Observations on euphausiid assemblages of the south coast of South Africa. South African Journal of Marine Science, 16(1), pp.143-155.

Girard, E.B., Langerak, A., Jompa, J., Wangensteen, O.S., Macher, J.-N. and Renema, W. (2022). Mitochondrial Cytochrome Oxidase Subunit 1: A Promising Molecular Marker for Species Identification in Foraminifera, *Frontiers in Marine Science*, 9, Article 809659. Available at: https://doi.org/10.3389/fmars.2022.809659

Govender A, Groeneveld J, Singh S, Willows-Munro S (2019) The design and testing of mini-barcode markers in marine lobsters. PLoS ONE 14(1): e0210492. https://doi.org/10.1371/journal.pone.021049

Hall TA (1994) BioEdit: a user-friendly biological sequence alignment editor and analysis programme for Windows95/98/NT. Nucleic Acids Symposium Series 41:95-98

Halsband, C. and Hirche, H.J., (2001) Reproductive cycles of dominant calanoid copepods in the North Sea. Marine Ecology Progress Series, 209, pp.219-229.

Hebert P.D., Cywinska A., Ball S.L. & DeWaard J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B Biological Sciences* **270**, 313-321.

Hebert, P.D.N. and Gregory, T.R., (2005) The promise of DNA barcoding for taxonomy. Systematic Biology, 54(5), pp.852–859.

Huntley M. & Sykes P. (1989) A study of Chaetognath abundance, body size and feeding in the Atlantic and Antarctic oceans. *Marine Biology* **103**, 153-162.

Imtiaz, A., Mohd Nor, S.A. and Md. Naim, D., (2017) Review: Progress and potential of DNA barcoding for species identification of fish species. Biodiversitas, 18(4), pp.1394-1405

Javidpour, J., Molinero, J. C., Peschutter, J., & Sommer, U. (2009). Seasonal dynamics of zooplankton community structure and trophic processes in the eutrophicated coastal area (Vranjic Basin, eastern Adriatic Sea). Estuarine, Coastal and Shelf Science, 82(2), 311-322.

Johnson, W., & Allen, D. (2012). *Zooplankton of the Atlantic and Gulf Coasts: A guide to their identification and ecology*. Johns Hopkins University Press

Jennings R.M., Bucklin A. & Pierrot-Bults A. (2010) Barcoding of arrow worms (Phylum Chaetognatha) from three oceans: genetic diversity and evolution within an enigmatic phylum. *PLOS ONE* **5**, e9949.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M.Y., Sturrock, S., Thierer, T., Ashton, B. and Meintjes, P. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics*, 28(12), pp. 1647-1649. doi:10.1093/bioinformatics/bts199.

Kekkonen, M., Mutanen, M., Kaila, L., Nieminen, M. & Hebert, P.D. (2015). Delineating species with DNA barcodes: a case of taxon-dependent method performance in moths. *PLoS ONE*, 10(4), e0122481. Available at: https://doi.org/10.1371/journal.pone.0122481

Kelly, L.J., Edwards, R., Morgan-Richards, M., Cordeiro, E.M., Suganuma, N., & Meudt, H.M., (2019) DNA barcoding of native New Zealand plants: A critical assessment of the current status and future prospects. Molecular Ecology Resources, 19(5), pp.1139-1156

Kimura M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16, 111-120

Kress, W. J., & Erickson, D. L. (2007). DNA barcodes: Genes, genomes, and the future of plant biodiversity. *Proceedings of the National Academy of Sciences*, 104(29), 12897-12899

Larkin, M.A. & Blackshields, Gordon & Brown, Nigel & Chenna, Ramaprabu & McGettigan, Paul & Mcwilliam, Hamish & Valentin, Franck & Wallace, I.M.W. & Wilm, Andreas & Lopez, Rodrigo & Thompson, Julie & Gibson, T.J. & Higgins, D.G.. (2007). Clustal W and Clustal X Version 2.0. Bioinformatics (Oxford, England). 23. 2947-8. 10.1093/bioinformatics/btm404.

Lv, J., Wu, S., Zhang, Y., Chen, Y., Feng, C., Yuan, X., Jia, G., Deng, J., Wang, C., Wang, Q., Mei, L. and Lin, X. (2014). Assessment of four DNA fragments (COI, 16S rDNA, ITS2, 12S rDNA) for species identification of the Ixodida (Acari: Ixodida), *Parasites & Vectors*, 7, p. 93. Available at: https://doi.org/10.1186/1756-3305-7-93

Macher T-H, Schütz R, Yildiz A, Beermann AJ, Leese F (2023) Evaluating five primer pairs for environmental DNA metabarcoding of Central European fish species based on mock communities. Metabarcoding and Metagenomics 7: e103856. https://doi.org/10.3897/mbmg.7.103856

Miller, C.B., Frost, B.W., Batchelder, H.P., Clemons, M.J. and Conway, R.E., 2004. Life histories of large, grazing copepods in a subarctic, oceanic gyre: Neocalanus plumchrus, Neocalanus flemingeri and Eucalanus bungii in the Northeast Pacific. Progress in Oceanography, 62(1), pp.41-78

Miller-Butterworth, C.M., Eick, G., Jacobs, D.S., Schoeman, M.C., & Harley, E.H. (2005). Genetic and phenotypic differences between South African long-fingered bats, with a global miniopterine phylogeny. *Journal of Mammalogy*, 86(6), pp.1121–1135. Available at:

https://doi.org/10.1644/05-MAMM-A-021R1.1

Meyer C.P. & Paulay G. (2005) DNA barcoding: Error rates based on comprehensive sampling. PLOS *Biology* **3**, e422.

Moraes Zenker, M., Portella, T.P., Pessoa, F.A.C., Bengtsson-Palme, J. & Galetti, P.M. Jr. (2024). Low coverage of species constrains the use of DNA barcoding to assess mosquito biodiversity. *Scientific Reports*, 14, p.7432. Available at: https://doi.org/10.1038/s41598-024-58071-1Nair V.R., Kidangan F.X., Prabhu R.G., Bucklin A. & Nair S. (2015) DNA barcode of Chaetognatha from Indian Waters. *Indian Journal of Geo-Marine Sciences* **44**, 1366–1376.

Nicol, S. (2006) Krill, currents, and sea ice: Euphausia superba and its changing environment. BioScience, 56(2), 111-120

Ojeda, A.A., Novillo, A., Lanzone, C., Rodríguez, M.D., Cuevas, M.F., Jayat, J.P., Teta, P., Ojeda, R.A. and Borisenko, A., (2022) DNA barcodes highlight genetic diversity patterns in rodents from lowland desert and Andean areas in Argentina. Molecular Ecology Resources, 22(6), pp.2349–2362.

Ortman, B.D., Bucklin, A., Pagès, F. and Youngbluth, M.J. (2010) DNA Barcoding the Medusozoa using mtCOI', *Molecular Phylogenetics and Evolution*, 55(1), pp. 1-10. doi:10.1016/j.ympev.2010.01.002

Peijnenburg, K.T.C.A., Fauvelot, C., Breeuwer, J.A.J. and Menken, S.B.J., (2004) Genetic diversity and population structure of marine species assessed by mitochondrial COI sequences: examples and implications. Molecular Ecology, 13(3), pp. 643-656.

Phillips, J.D., Gillis, D.J. and Hanner, R.H. (2022) Lack of statistical rigor in DNA barcoding likely invalidates the presence of a true species' barcode gap, *Frontiers in Ecology and Evolution*, 10. https://doi.org/10.3389/fevo.2022.859099

Phillips, A.J., Kullander, S.O., & Meyer, A. (2024) New Approaches in DNA Barcoding and Species Delimitation: Insight from South American Cichlids. Molecular Ecology Resources, 24(1), 112-124

Peijnenburg KTCA, Breeuwer JAJ, Pierrot-Bults AC, Menken SBJ. Phylogeographyof the planktonicchaetognath Sagitta setosa reveals isolation in European seas. Evolution. 2004;58:1472–1487

Porter, T.M. and Hajibabaei, M., (2020) Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis. Ecology and Evolution, 10(14), pp. 7141-7150.

Puillandre, N., Lambert, A., Brouillet, S. and Achaz, G., (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology, 21(8), pp. 1864-1877.

Radulovici, A.E., Archambault, P. and Dufresne, F., (2010) DNA barcodes for marine biodiversity: moving fast forward?. Diversity, 2(4), pp.450-472.

Ratnarajah, L., Abu-Alhaija, R., Atkinson, A., Batten, S., Bax, N.J., Bernard, K.S., Canonico, G., Cornils, A., Everett, J.D., Grigoratou, M., Ishak, N.H.A., Johns, D., Lombard, F., Muxagata, E., Ostle, C., Pitois, S., Richardson, A.J., Schmidt, K., Stemmann, L., Swadling, K.M., Yang, G. and Yebra, L. (2023) Monitoring and modelling marine zooplankton in a changing climate, *Nature Communications*, 14, 564. Available at: https://doi.org/10.1038/s41467-023-36241-5

Raupach, Michael & Radulovici, Adriana. (2015). Looking back on a decade of barcoding crustaceans. ZooKeys. 539. 53-81. 10.3897/zookeys.539.6530.

Richardson A.J. (2008) In hot water: zooplankton and climate change. ICES *Journal of Marine Science* **65**, 279-295.

Singh S.P., Groeneveld J.C., Huggett J., Naidoo D., Cedras R. & Willows-Munro S. (2021) Metabarcoding of marine zooplankton in South Africa. *African Journal of Marine Science* **1**, 1-13.

Stone SE, Stone JP. Diversity, seasonal abundance, and environmental drivers of chaetognath populations in North Inlet Estuary, South Carolina, USA(2023) Ecol Evol. 30;13(5):e10151. doi: 10.1002/ece3.10151. PMID: 37261320; PMCID: PMC10227638

Taylor, P. J., Stoffberg, S., Monadjem, A., Schoeman, M. C., Bayliss, J., & Cotterill, F. P. (2016) Genetic diversity and phylogeography of long-fingered bats (Miniopterus spp.) in Africa using mitochondrial DNA barcodes. Acta Chiropterologica, 18(2), 409-424.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M *et al*. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. Mol Biol Evol 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353

Tempestini, A. Rysgaard, S. Dufrense, F. (2018). Species identification and connectivity of marine amphipods in Canadas three oceans.PLos ONE:13(5):e0197174

Trigg S. and Flasse S. (2001) An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. International Journal of Remote Sensing 22, 2641-2647

Tzafesta, E., Saccomanno, B., Zangaro, F., Vadrucci, M.R., Specchia, V. and Pinna, M. (2022) DNA Barcode Gap Analysis for Multiple Marker Genes for Phytoplankton Species Biodiversity in Mediterranean Aquatic Ecosystems, *Biology*, 11(9), p. 1277. Available at: https://doi.org/10.3390/biology11091277

Ulloa, R. , Palma, S. , & Silva, N. (2000). Bathymetric distribution of chaetognaths and their association with water masses off the coast of Valparaíso, Chile. *Deep Sea Research Part I: Oceanographic Research Papers*, 47(11), 2009–2027.

van der Loos, L.M. and Nijland, R. (2021) 'Biases in bulk: DNA metabarcoding of marine communities and the methodology involved', *Molecular Ecology*, 30(13), pp. 3270–3288. doi: 10.1111/mec.15592. Epub 29 August 2020. PMID: 32779312; PMCID: PMC8359149.

Wei, Z., Chen, Q., Zhang, Y., He, Z., (2018) Genetic variation analysis of the cosmopolitan chaetognath Sagitta enflata (Chaetognatha: Sagittidae) in the northern South China Sea based on mitochondrial COI gene sequences. Mitochondrial DNA Part A, 29(5), pp. 705-712.

Wiemers M. & Fiedler K. (2007) Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera Lycaenidae). *Frontiers in Zoology* **4**, 8.

Wiens, J.J. and Penkrot, T.A., (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). Systematic Biology, 51(1), pp.69-9

Zhang, J., Kapli, P., Pavlidis, P. and Stamatakis, A., (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), pp. 2869-2876.

Chapter Three

DNA reference libraries for Euphausiids and Chaetognaths found in South African waters: Implications for future zooplankton metabarcoding research

Abstract

Zooplankton are a primary food source for many larger marine animals and act as important marine ecosystem health indicators. Identification of many zooplankton taxa using only morphology-based keys to species level is challenging. DNA barcoding can aid in the rapid delimitation of species, but the reliability of DNA barcoding is dependent on the availability of comprehensive reference libraries. The completeness of reference libraries is difficult to gauge in under-studied groups and in the biodiversity-rich regions of the world where many undescribed species may exist. In this study the current status of the reference libraries for Euphausiids and Chaetognaths in South African waters were reviewed and expanded. To add to the reference library zooplankton samples were collected from five localities off the East Coast of South Africa: iSimangaliso Wetland Park, Richards Bay, uThukela Marine Protected Area (MPA), Durban and Aliwal Shoal MPA. Each zooplankton sample was divided into two subsamples. From one subsample Euphausiid and Chaetognath specimens were taxonomically identified using morphological keys, Sanger sequenced and added onto the Barcode of Life Data System (BOLD) as new records. The second subsample was used for metabarcoding (simultaneous identification of multiple species using high-throughput sequencing). The results from the metabarcoding analyses were used to measure the completeness of the reference library for Euphausiids and Chaetognaths from South Africa and predict how many, as yet undescribed, species may exist in the region. The DNA barcoding used as a tool in this study resulted in a total of 147 new specimen records on BOLD (from 220 morphologically identified specimens): six new Chaetognath specimen records, with one species identified and;141 Euphausiid specimen records with 127 species records and 13 records identified to genus-level and one record identified to family-level. The metabarcoding results identified 114 Euphausiid species identified by 99 amplicon sequence variants (ASVs) (>97% similarity); this included 23 of the 86 described Euphausiid species found alongside the East Coast of South Africa. Similarly, 15 Chaetognath species were identified using six ASVs (>97% similarity). The efforts of this study have contributed towards 79% of all zooplankton species having DNA barcodes which are publicly available.

3.1. Introduction

Zooplankton play a vital role in marine and freshwater ecosystems as nutrient recyclers and contributes significantly to the functioning and balance of aquatic environments (Brierley 2017). Changes in zooplankton communities can be used to track the ecological condition of aquatic environments, with changes in species richness or community structure signaling changes in environmental conditions (Richardson 2008). To identify and describe species accurately requires experience, sufficient published literature, patience and a specialist taxonomist (Bucklin *et al.* 2007; Goswami *et al*. 2004). The correct identification of species underpins all biological research and is particularly important when species richness is used as a biomonitoring tool (Bucklin *et al.* 2011; Chapple and Ritchie 2013).

3.1.1. Euphausiid and Chaetognath species richness in South African waters

Globally, Copepods make up the majority of biomass in zooplankton communities, though Euphausiids and Chaetognaths also contribute significantly towards species richness (Atkinson *et al.* 2018). Euphausiids play an important role in marine food chains as a primary food source for numerous marine organisms, including larger vertebrate taxa such as whales, fish, seabirds, and seals (Lowe *et al*. 2018). Large Euphausiid populations are important in supporting commercial fisheries in some regions of the world, making them economically significant (Gibbons 1995; Gibbons *et al*. 1995). Globally, Euphausiids consist of 86 described species 23 of which are found in South African waters (Brinton *et al*. 2000; Magnus *et al*. 2017; Pillar *et al.* 1992; Sell *et al*. 2024). Gibbons *et al*. (1995) conducted a study of Euphausiid abundance along the East and West Coasts of South Africa, their findings indicating a significantly higher species abundance on the East Coast with 29 species identified, compared to only ten species found along the West Coast of South Africa (Gibbons and Hutchings 1996).

Chaetognaths play an integral role in regulating the ecological balance of plankton communities, as they primarily feed on other zooplankton, including copepods and small crustaceans (Bucklin *et al*. 2021; Gasmi *et al.* 2014). Chaetognaths have a conserved body plan and display few diagnostic morphological traits, leading to many taxonomic ambiguities as highlighted by previous taxonomic assessments (Haddock *et al*. 2004; Harris and Martin 2014; Lee *et al*. 2005). Chaetognaths include approximately 200 described species (Perez *et al.* 2021), seven of which are found in South African waters (Duró and Saiz 2000).

3.1.2. Use of molecular methods for species identification

Zooplankton communities rapidly respond to changes in their environment and also to pollution and other stressors in the environment and hence are important for climate change monitoring (Ratnarajah *et al.* 2023). By monitoring zooplankton communities over time, a greater understanding of how stressors affect higher trophic levels can be established (Bucklin *et al*. 2004; Muxagata *et al*. 2019; Richardson 2008). DNA barcoding and metabarcoding could be utilized as biomonitoring tools by allowing for more comprehensive diversity assessments, compared to traditional taxonomic methods that are morphology-based (Bucklin *et al*. 2021).

DNA barcoding involves the sequencing of standardized regions of the genome to identify species based on molecular data and has become a popular method to facilitate and aid traditional morphology-based taxonomy (Guo *et al.* 2022; Hebert *et al.*2003; Pires and Marinoni 2010). In most animal taxa, the mitochondrial cytochrome oxidase subunit I (COI) gene is the most commonly used region for DNA barcoding (Hebert *et al*. 2003). The COI gene region exhibits genetic variation for effective species-level identification and discrimination (Andújar *et al.* 2018; Deagle *et al.* 2014).

DNA barcoding has revolutionized species identification of zooplankton, as there are ~5700 described holoplankton species that are linked to COI sequences (Bucklin *et al*. 2021). However, an estimated ~1600 species are yet to be described, and barcoded (Bucklin *et al*. 2021). The proportion of globally described zooplankton species with DNA barcodes has increased from less than 10% in 2011 to 23% in 2019 (Singh *et al.* 2021) and 60% in 2021 (Grant *et al*. 2021). Currently it is estimated that 79% of zooplankton species have DNA barcodes publicly available (https://boldsystems.org/; accessed June 2024).

Metabarcoding merges the fields of DNA barcoding and high-throughput sequencing technologies, for the simultaneous identification of multiple species from a single sample (Govender *et al*. 2022a; Govender *et al*. 2022b; Guo *et al*. 2022). Metabarcoding has emerged as a fundamental methodology in the investigation of biological communities (Macher *et al*. 2024). DNA barcoding and metabarcoding is a powerful analytical tool for accurate species identification, however it is reliant on the availability of reference libraries that contain sequence data linked to taxonomically verified species (Bucklin *et al*. 2016). Marine zooplankton assemblages are taxonomically complex, making tools for accurate species discrimination critically in demand.

3.1.3. Reference libraries for DNA barcoding and metabarcoding

In a recent study by Singh *et al.* (2021) an assessment of DNA barcode reference databases for marine zooplankton in South African waters, such as BOLD and Genbank, was reviewed. The authors found that 90 Euphausiid species barcodes were available in global reference libraries with 49 of these species having sequences from individuals collected in South African waters (Singh *et al.* 2021). Similarly, 121 Chaetognath species had barcode records globally, but only 28 species have been sequenced from taxa collected in South African waters (Singh *et al*. 2021). As such, the current reference libraries for Euphausiids and Chaetognaths in South African waters do not capture the diversity of these taxa, as many species remain undescribed and DNA barcode reference records are not available for most species (Gibbons *et al.*1995; Ratnasingham and Hebert 2013). The taxonomic ambiguities and incompleteness of reference libraries for these two taxa have also been highlighted in other regions (Johnson and Tarling 2017; Wright, 2014). It is important to include representatives of species collected in South African waters as this can resolve the identification of cryptic intraspecific morphological differences linked to environmental heterogeneity. For example, Euphausiid individuals belonging to the same species can differ morphologically due to variations in environmental factors such as temperature, salinity and depth (Boden 1954; Bucklin *et al*. 2004; Richardson 2008; Muxagata *et al*. 2019). A DNA-based system used to identify species together with vital information on species diversity and geological data from various regions of the world will aid in a greater understanding of the roles played by these species ecologically and highlights their importance within marine ecosystems (Atkinson *et al*. 2008).

3.1.4. Aims of the study

Using paired samples collected from five sites off the East Coast of South Africa this study aimed to: (i) use available morphological keys and the assistance of a taxonomic expert to identify species of Euphausiids and Chaetognaths and compare this to the current species list available for the region; (ii) to add to the current South African DNA barcode reference library for Euphausiids and Chaetognaths using Sanger sequencing and; (iii) to use metabarcoding data generated from the same area to assess the efficacy of the current barcode reference library to identify Euphausiids and Chaetognaths to species level.

3.2. Materials and methods

3.2.1. Sampling: South African East Coast

The South Africa coastline is approximately 3,650 km with the oceanographic regime being dominated by two major current systems: the cold Benguela Current along the Atlantic coast to the West and the warm Agulhas Current along the Indian Ocean coast to the East (Griffiths *et al.* 2010). The Agulhas current predominates the East Coast of South Africa, where samples were collected from the following locations: iSimangaliso Wetland Park, Richards Bay, uThukela Marine Protected Area (MPA), Durban and Aliwal Shoal MPA(Figure 3-1).

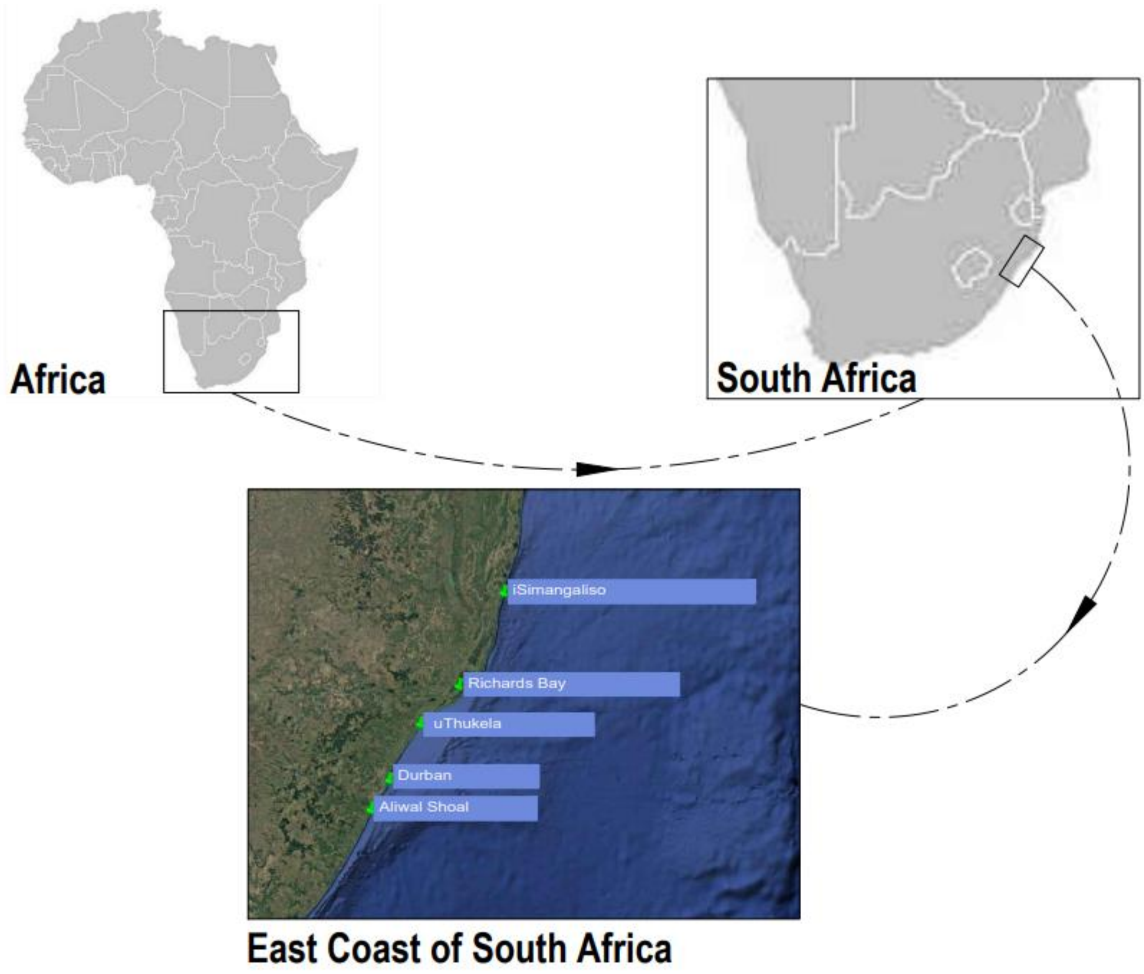

Figure 3-1: Map indicating the five sampling sites where zooplankton samples were collected along KwaZulu-Natal, the East Coast of South Africa.

Zooplankton samples were collected from five stations (Figure 3-1) with longitude and latitude provided in Appendix 1, at 100m and 200m depths, between June 2018 and November 2019. Three net types were used: a ring net with a diameter of 0.8 m and 500 µm mesh; a Manta net with a rectangular frame of 0.5 × 0.15 m and 500 µm mesh; and a WP2 plankton net with a 0.55 m diameter and 200 µm mesh. Three replicate ring net tows were performed at each sampling, resulting in 12 from ring nets (four sets of three replicate tows) and four samples each from the Manta and WP2 nets. The ring- and Manta nets were towed horizontally just below the surface for 5 minutes at a time with a ground speed of 2–3 knots. The WP2 net was lowered to 10 m above the seafloor at each sampling station and then hauled vertically. All tows were undertaken at night, between sunset and sunrise.

Zooplankton samples were stored in 96% ethanol directly after collection and stored at –20°C at the Oceanographic Research Institute (ORI), in Durban, KwaZulu-Natal, South Africa. To ensure long term storage and prevent degradation, ethanol was replaced in sample jars 24 hours after collection. These samples were then split into two subsamples; the first was processed for morphological identification and Sanger sequencing and the second sample was used for metabarcoding which was conducted by Govender *et al.* (2024) and reported in this dissertation as comparison.

3.2.2. Morphological identification and expansion of the DNA barcode reference library

The taxonomic identification of Chaetognatha and Euphausiidae was conducted with the guidance and assistance of Dr. Riaan Cedras (Ecologist, iSimangaliso Wetland Park Authority, St Lucia, South Africa) and using a ZEISS Stemi 2000-C dissecting microscope (Magnification x 50-Carl Zeiss, Oberkochen, Germany) the two taxa: Euphausiids and Chaetognaths were initially separated from collected whole zooplankton subsample. Where possible specimens were categorized to their lowest taxonomic level using available taxonomic keys for Euphausiids (Baker *et al*. 1990), and Chaetognaths (Baker 1990). A minimum of five specimens per species per sampling site (iSimangaliso Wetland Park, Richards Bay, uThukela Marine Protected Area (MPA), Durban and Aliwal Shoal MPA(Figure 3-1)were then chosen for Sanger sequencing. Every specimen was photographed using the ZEISS AxioCam, which was attached to the ZEISS dissecting microscope.

Total genomic DNA was extracted from individual specimens using the ZR Tissue and Insect MicroPrepKit (Zymo Research, Irvine, CA, U.S.A.), following the manufacturer's instructions which was modified to include an initial incubation step at 55˚C overnight. The concentrations and quality of the extracted genomic DNA were quantified with the Nanodrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, U.S.A.). The barcode region of the mitochondrial COI gene was amplified using the universal primers LCOI1490 (5' –GGTCAACAAATCATAAAGATATTGG-3') and HC02198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') originally developed by Folmer *et al*. (1994). The PCRs were carried out in a 25 µl reaction volume made up of: genomic DNA (10-15 ng/µl on average), 12.5 µl OneTaq Quick-Load Master Mix (1X, BioLabs, New England), 0.50 µl forward and reverse primers (10 µM each), 2.5 µl sterile nuclease-free water, 2 µl additional $MgCl_2$ (25 µM) and 2 µl Bovine Serum Albumin (BSA) (1 $mg.m^{-1}$). All PCR reactions were run with a negative control. The thermal cycling program included initial denaturation at 94˚C for 2 minutes, followed by 35 cycles of denaturation at 94˚C for 30 seconds, annealing at 50˚C for 30 seconds, and extension at 68˚C for 1 minute. The final extension step was carried out at 68˚C for 5 minutes.

The PCR products were visualized on a 0.8 % (w/v) TBE agarose gel. A 100-bp molecular weight ladder was used to estimate the size of PCR products. The PCR products were sized using a standard curve and only the PCR products which were of the expected size of 658-bp, were sent for Sanger sequencing. PCR products were bi-directionally cycle sequenced using BigDye Chemistry, v3.1 (Thermo Fisher Scientific, Waltham, MA, U.S.A.), at the Central Analytical Facility at Stellenbosch University, South Africa. Sanger sequencing products were analyzed on an Applied Biosystems 3730xl Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, U.S.A.). The electropherogram of each

PCR amplicon was examined for quality using the BioEdit 7.0.9.0 sequence alignment editor (Hall 1999).

Each COI sequence was queried against the Barcode of Life Data System (BOLD) and NCBI GenBank using the BLASTn feature (www. ncbi.nlm.nih.gov/genbank) and taxa with sequence similarity values > 95 % were assigned to species, sequences with similarity 90%-95% were only identified to genus and sequences with similarity 85%-90% were only assigned to families. To evaluate the impact this study had, and the improvement made in the South African representation in the reference library, species which were already available on BOLD before the addition of sequences from this study were noted. In particular, the total number of Euphausiid and Chaetognath species barcodes on BOLD in 2021 were compared to the total number of (Euphausiid and Chaetognath) barcodes on BOLD in 2024. Barcode-compliant sequences (> 500 bp in length, no stop codons, misidentifications, or contaminants) and associated photographs and collection data generated in this study were uploaded onto BOLD (see Appendix 1) under the project code ORIA (FBIP - DNA barcoding of South African Marine Euphausiacea, Chaetognatha, and Amphipoda).

3.2.3. Metabarcoding of zooplankton

Metabarcoding data previously produced by Govender *et al*. (2024) from the second subsample were included in the present analysis for comparison purposes. The protocol originally applied by Govender *et al.* (2024) is described briefly below. The metabarcoding samples were homogenized in the 97% ethanol solution for 45 seconds using a consumer blender (Defy PB7354X, 350W and 22 000 rpm). To prevent contamination between samples, the blender was washed with a soap solution to remove residual material, and then thoroughly rinsed with a 10% bleach solution and rinsed again with 70% ethanol to remove any residual DNA. Laboratory blanks were also used throughout the DNA extraction and amplification steps to test for any contamination. Three replicates (10 ml of zooplankton slurry) per homogenate were taken and centrifuged at 1200× g for 1 min, repeated to remove excess ethanol. DNA extractions were performed using the Qiagen DNeasy Blood and Tissue Kit (Qiagen GmbH, QIAGEN Str. 1, D-40724).

In Govender *et al.* (2024) a wide range of primers (six primer sets) were utilized to target mainly Malacostraca and Actinopterygii together with one universal primer set which targeted metazoan diversity. These primers also amplified species belonging to Euphausiacea and Chaetognatha. Polymerase chain reactions (PCRs) were performed in triplicate to minimize bias and amplification errors. All PCRs included a no-template negative control. PCR products were visualised on a 1% (w/v) TBE agarose gel containing 0.02% ethidium bromide (EtBr). Amplicon size was determined using a 100-bp molecular weight marker (Solis Biodyne). The triplicate PCR products for each of the six primer sets were pooled and quantified using a Qubit 2.0 Fluorometer (Life Technologies, California, USA). These products were then sequenced on the Illumina MiSeq platform (Illumina) at the KZN Research and Innovation Platform (KRISP, South Africa) using a MiSeq Nano Reagent Kit v.2 for 500 cycles (Govender *et al*. 2024).

Govender *et al* (2024), implemented the dada2 algorithm (Callahan *et al.* 2016) in Qiime2 v. 2019.10 (Bolyen *et al.* 2019) which was used for quality control checks, chimera removal and merging of paired end reads into amplicon sequence variants (ASVs), which were queried on BOLD v4. and NCBI Genbank in January 2024. Taxonomic assignment to species level was implemented using a threshold of 97% sequence divergence. The species detected by metabarcoding were further cross referenced with occurrence records available online and in cited literature.

The records generated in this taxonomic study of individual specimens, were compared with the zooplankton samples identified in the metabarcoding project conducted on East Coast of South Africa by Govender *et al.* (2024). The comparison was based on the following criteria: (i) the number of Euphausiid and Chaetognath species identified; (ii) improvements in sequence similarity; (iii) the addition of new species, previously unrepresented on the BOLD database; (iv) any taxonomic revisions, genus and species name changes; and (v) the number of barcode records for each species.

3.3. Results

3.3.1. Morphological identification of Euphausiid and Chaetognath species

A total of 141 Euphausiid specimens belonging to four genera (*Euphausia*, *Hansarsia*, *Stylocheiron* and *Thysanoessa*) and one family (Euphausiidae) were identified using morphological keys. Of these, 127 could be confidently assigned to species-level, while 13 were identified to genus-level and one to family-level. Chaetognath specimens were difficult to identify morphologically due to their fragile bodies being possibly damaged, or not well preserved by ethanol (Casanova, 1999; Goetze & Jungbluth, 2013). This made key diagnostic features challenging to identify. Additionally, many diagnostic features are only present in mature individuals (Casanova, 1999), compounding the morphological identification challenges experienced. A total of six Chaetognath specimens belonging to two genera (*Flaccisagitta* and *Sagitta*) and one family (Sagittidae) were identified using morphological keys. Only one specimen was identified to genus-level (*Sagitta*) and one specimen to species-level (*Sagitta enflata*).

3.3.2. DNA barcoding

A total of 220 specimens were chosen to be COI barcoded using Sanger sequencing and from these 118 full-length COI Euphausiid and Chaetognath DNA barcodes were generated.

Table 3-1: Euphausiids (Talbot 1974), and Chaetognaths (Stone 1965) species expected along the East and West Coasts of South Africa. Column three represents DNA barcodes generated from species identified from the morphological study. The last column represents DNA barcodes of species identified from the metabarcoding study conducted by Govender *et al*., (2024).

Species in South Africa	**East Coast**	**West Coast**	**Morphology**	**Metabarcode**
Euphausiid				
Euphausia brevis	x	x		
Euphausia diomediae	x	x	x	x
Euphausia hemigibba	x	x		
Euphausia lucens		x		x
Euphausia mutica	x	x		x
Euphausia paragibba	x	x		
Euphausia recurva	x	x	x	x
Euphausia similis				x
Euphausia spinifera	x			
Euphausia tenera	x	x		
Hansaria microps				x
Nematobrachion flexipes	x			
Nematoscelis microps	x	x		
Nematoscelis megalops	x	x		
Nematoscelis tenella	x	x		
Nyctiphanes capensis		x		
Stylocherion affine	x	x	x	
Stylocherion abbreviatum	x	x		
Stylocherion carinatum	x	x	x	x
Stylocherion elongatum	x	x		
Stylocherion longicorne	x	x		
Stylocherion microphthalma	*x*	x		
Stylocherion maximum	*x*	x		
Stylocherion suhmil	x	x		
Thysanopoda acutifrons	x	x		
Thysanopoda aequalis	x	x		
Thysanopoda monacantha	x	x		

Table 3-1. Cont. **Species in South Africa**	**East Coast**	**West Coast**	**Morphology**	**Metabarcode**
Thysonessa gregaria	x	x	x	x
Thysanopoda orientalis	x	x		
Thysanopoda subaequalis	x			
Thysanopoda tricuspidata	x	x		
Chaetognath				
Pterosagitta draco	x			
Sagitta bedoti	x			
Sagitta bipunctata	x			
Sagitta decipiens		x		
Sagitta enflata	x		x	x
Sagitta hexaptera	x			
Sagitta friderici		x		
Sagitta lyra		x		
Sagitta minima		x		
Sagitta regularis	x			
Sagitta robasta	x			
Sagitta serratodentata	x			
Sagitta subtilis		x		

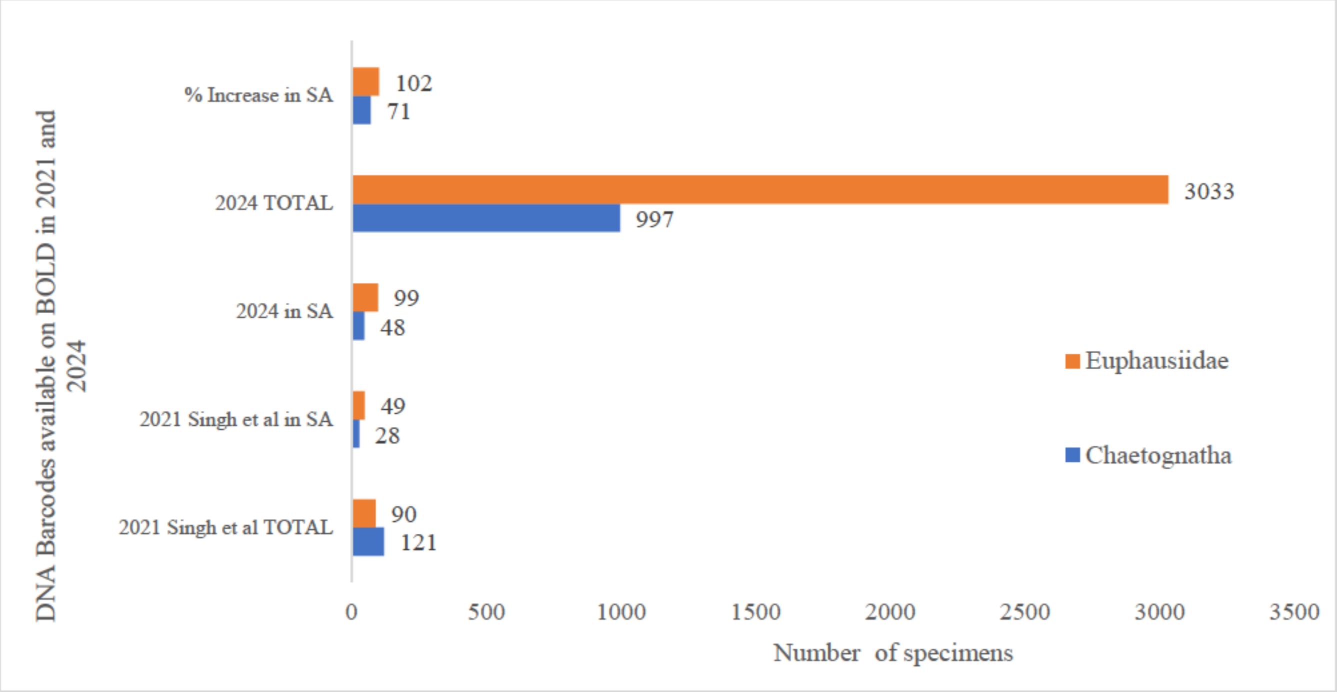

Figure 3-2: Number of Euphausiid and Chaetognath species with barcodes available on BOLD (accessed on 27 May 2024), in total and for South Africa compared to the numbers of each taxa in the year 2021.

In Singh *et al.* (2021), the total number of Euphausiid DNA barcodes available globally were 90 Euphausiid species and 121 Chaetognath species (Figure 3-2). The results of this present study have increased the number of barcodes on BOLD by a total of 118 new specimen records with two new Chaetognath specimen records. Euphausiids had 114 specimen records: 125 species and 12 to individuals only identified to genus-level, using morphological keys. In South Africa, the number of Chaetognath species barcodes were 28 in 2021 (Singh *et al.* 2021) and have increased to 48 species through data added by this study, Euphausiids had 49 species records in 2021 (Singh *et al.* 2021), and this has increased to 99 species with barcodes on BOLD currently (Figure 3-2).

3.3.3. Metabarcoding

As per Govender *et al.* (2024), a final total of 1 047 merged amplicon sequence variants (ASVs) were utilized for analysis. Of the 1 047 ASVs, 589 ASVs could not be confidently assigned to species level (< 97% sequence similarity). Across all sampling sites, metabarcoding detected zooplankton taxa belonging to 27 orders, 89 families, 160 genera and 224 species. From the 1 047 ASVs, 114 Euphausiid species were identified (>97% similarity), including 23 of the 86 described Euphausiid species found on the East Coast of South Africa: Among the seven Chaetognatha species found in Southern African waters, 15 species (>97% similarity) were identified among the 200 species known globally.

3.3.4. Improvement in species identification of taxa: Euphausiids and Chaetognaths

The morphological study has contributed 84 new sequences for *Euphausia recurve* and 11 for *Thysanoessa gregaria;* all of which are added as new sequences on BOLD. Additionally, two sequences have been added for *Stylocheiron carinatum, Euphausia mutica,* and *Euphausia lucens,* five new sequences for *Hansarsia microps* and *Stylocheiron spp* each, and 15 new sequences for *Euphausia diomedeae. H. microps* and *S. carinatum* have ≥ 95% sequence similarity matches on BOLD. *E. brevis* shows an increase in percentage identification accuracy and is new on BOLD. *E. hanseni, E. recurva, H. atlantica,* and *H. megalops* are all new on BOLD and have increased the number of species that can be identified using metabarcoding. There are also changes in species identities due to the new addition of DNA barcodes on BOLD, these being *E. diomedeae, E. hanseni, E. recurva, H. microps,* and *Nyctiphanes capensis* due to possible misidentifications in previous studies (see Appendix 1).

Table 3-2: Increased detection and accuracy of identification for Euphausiid sequences by metabarcoding of bulk tow-net samples, following improvements in the DNA barcode reference database (BOLD) over 5 years (2019-2024). There are multiple changes in species identifications, new barcodes, and increased accuracy for specimen identification, displaying a 43% percentage change in the number of species identified in 2019 compared to 2024.

Criteria		**Rationale**	**Source of change**
Metagenetic species identified in 2019	14		
Metagenetic species identified in 2024	20		
New species identified in 2024	*Euphausia hanseni*	New on BOLD	This study
	Euphausia recurva	New on BOLD	This study
	Hansarsia megalops	New on BOLD	This study
	Hansarsia atlantica	New on BOLD	This study
	Euphausia brevis	New on BOLD	This study
Increased accuracy	*Euphausia brevis*	Increased accuracy	This study
Change in spp identity	*Nyctiphanes capensis*	Possible misidentification in previous study	This study
	Euphausia recurva	Possible misidentification in previous study	This study
	Euphausia hanseni	Possible misidentification in previous study	This study

Table 3-2 Cont **Criteria**		**Rationale**	**Source of change**
	Hansarsia microps	Possible misidentification in previous study	This study
	Euphausia diomedeae	Possible misidentification in previous study	This study

Table 3-3: Increased detection and accuracy of Chaetognath sequences by metabarcoding of bulk tow-net samples, following improvements in the DNA barcode reference database (BOLD) over a 5-year period (2019-2024).

Criteria		**Rationale**	**Source of change**
Metagenetic species identified in 2019	10		
Metagenetic species identified in 2024	10		
New species identified in 2024	*Sagitta enflata*	New on BOLD	This study

Sagitta enflata has two new DNA barcodes uploaded onto BOLD, which will increase identification for these specific species using metabarcoding.

3.4. Discussion

Given the taxonomic challenges identifying zooplankton species, the number of undescribed species in biodiverse countries such as South Africa is expected to be high (Bucklin *et al*. 2021). By comparing the number of species recovered using morphological/Sanger sequencing identification versus metabarcoding, an estimation of the completeness and utility of the current reference library for these two important zooplankton groups was performed. Understanding zooplankton biogeography and documenting undescribed species will aid in better understanding of zooplankton and their role in marine ecosystems (Bieri and Kormanol 2020). An extensive taxonomically verified COI reference sequence database for marine zooplankton is instrumental for DNA barcoding and metabarcoding applications and is pivotal for ocean ecosystem monitoring, health and assessment (Bucklin *et al*. 2021; Song *et al*. 2021). This study contributes towards this global initiative.

The identification of zooplankton to species or genus level using taxonomic keys is a specialist job, is time-consuming, and requires someone highly experienced and trained to do so (Bucklin *et al*. 2021). This was evident as sorting specimens of Euphausiids and Chaetognaths from whole zooplankton samples was challenging. To inexperienced researchers, this is especially hard to execute, and the traits used in taxonomic keys (Baker *et al*. 1990; Baker 1990) are often hard to identify with an inexperienced eye. The need for the combination of morphological identification with DNA barcoding (Macher *et al.* 2024) is in demand and is the way forward to enhance species-level molecular studies, which will benefit metabarcoding efficacy.

3.4.1. DNA barcoding and metabarcoding efficacy

The present study contributed sequences to BOLD, increasing the regional species records for both the Euphausiids (49-99) and Chaetognatha (28-48). DNA barcoding was successful in amplifying *Euphausia diomedia, E.recurva, T. gregaria, S. carinatum, S.affine,* and *Sagitta enflata*. These species were within their known range of distribution along the East and West Coast of South Africa (Duró and Saiz 2000; Magnus *et al*. 2017; Pillar *et al.* 1992; Sell *et al*. 2024; Table 3-1). Species detected with metabarcoding, but which were not encountered during the morphological part of the study, included *E. similis*, *H. microps* and *E.mutica*. Prior studies (Sell *et al.* 2024; Talbot 1974) suggest that *E.lucens* only occur along the west Coast of South Africa, however, the findings of this study suggest the presence of *E.lucens* along the east Coast of Africa.

E. recurva, *E.mutica* and *H.microps* are new records as these species have not been recorded in South African waters before. The species *E. recurva* and *T.gregaria* had no records with sequences on BOLD prior to this study. The addition of 84 *E. recurva* and 11 *T.gregaria* sequence records have increased the likelihood of the identification of these species by 100% (Table 3-1). The number of species identified by metabarcoding in 2019 were 14 Euphausiids and together with the efforts of this study, increased this value to 20 in 2024. This study has also resulted in increased accuracy in species identification for *T. gregaria* and *E.recurva,* which led to the correct species identification with 100% sequence similarity. Some of the changes in species were *E.brevis* (formerly identified as *E. diomedea* with a 96% sequence similarity), *E.lucens* (formerly identified as *N.capensis with* a 95% sequence similarity), and *E.haseni* (formerly identified as *E.similis* with a 98% sequence similarity).

The identification of *E.similis* and *H.microps* in the metabarcoding results demonstrates how important and powerful metabarcoding can be to uncover potential cryptic species and undiscovered diversity. Both of these species (*E.similis* and *H.microps)* are not previously identified in the east or west Coasts of Southern Africa (Talbot 1974); were not identified morphologically, however they were identified in metabarcoding (Table 3-1).Additionally, the detection of *E.lucens,* which has not previously been recorded with traditional methods of observation in the East Coast of South African waters (Stuart and Pillar 1988), was identified through metabarcoding. A reason for this could be due to passive ocean currents or changes in environmental conditions due to season changes (Gibbons 1995). However, this identifies the need for further investigation regarding ocean currents and their effect on zooplankton distribution as well as how environmental factors are impacting these crucial oceanic health indicators. This study has improved the representation of Euphausiid species representative barcodes from South Africa (Figure 3-2., Table 3-1). Despite these advancements, many taxa remain understudied and necessitate future research to expand the taxonomic and geographic scope of DNA reference libraries, particularly for species lacking universal COI gene barcodes and to improve metabarcoding efficacy.

There were no changes in Chaetognath species identified using metabarcoding from 2019 to 2024 as shown in Table 3-3. However, *Sagitta enflata* was the only species identified to species level with morphological techniques and two new DNA barcodes were added onto BOLD. This is due to Chaetognaths being difficult to identify taxonomically as a result of their varied morphology, which changes due to changes in environmental conditions (Nair *et al.* 2015), as well as the damage experienced during the sampling process. Overall, this study has improved the reference DNA barcodes from South Africa for these taxa. However, there is a need for further study into the DNA barcoding of Chaetognaths in South Africa due to difficulties experienced (fragility of Chaetognaths) in this study. In light of the challenges experienced, Casanova (1999) emphasizes that minimizing physic damage through careful sampling and handling is crucial to preserve diagnostic characters for effective taxonomic identifications in Chaetognaths.

The metabarcoding analysis by Govender *et al*. (2024) detected zooplankton belonging to 27 orders, 89 families, 160 genera and 224 species with the dominant taxa being Malacostraca (110), Copepoda (48), Actinopterygii (46) and Gastropoda (20). The metabarcoding analysis also identified a total of 224 / 195 euphausiid and chaetognath species, 37 of which were not previously recorded in South Africa by Govender *et al.* 2024 but are known for its prevalence in the West Indian Ocean (WIO). This could be crucial for species identification and is indicative of how vast and important metabarcoding is for identification and conservation.

3.5. Conclusion

Future research should target taxa that have poor representation in the DNA reference libraries, to enhance their taxonomic and geographic scope. In addition to using the universal COI gene (Hebert *et al*. 2003), other genes should also be barcoded to optimize species identification and improve metabarcoding prolificacy. Notably, from the metabarcoding data, species such as *E.recurva*, and *T.gregaria* were newly recorded from South African waters. *E.similis,* and *H.microps* were species only identified using metabarcoding and not identified in morphological sampling, emphasizing the importance of metabarcoding for cryptic species and undiscovered biodiversity. *E.lucens* was also identified in metabarcoding, a species not previously found on the East Coast of South Africa, hence there is a need for further research on environmental factors and how ocean currents are affecting zooplankton species.

Further research is necessary to investigate the genetic diversity of understudied Chaetognath species in South African waters. One such investigation could be regarding sample collection for Chaetognaths specifically, developing a protocol where there is no damage to the specimen in the collection process and ensuring intact samples or easier morphological identification.

3.6. References:

Andújar C., Arribas P., Yu D.W., Vogler A.P. & Emerson B.C. (2018) Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*.

Atkinson A., Siegel V., Pakhomov E.A., Rothery P., Loeb V.J. & Hewitt R. (2008) A reevaluation of the role of krill in the Southern Ocean food web. *Antarctic Science* **20**, 145–157.

Baker A., de C Boden B. & Brinton E. (1990) A practical guide to the euphausiids of the world. London: Natural History Museum Publications.

Baker A.N. (1990) Chaetognatha of the Southern Ocean and Adjacent Waters. New Zealand *Oceanographic Institute Memoir* **97**, 1–128.

Bieri R. & Kormanol R. (2020) A new species of chaetognath from the deep-sea habitats of the South Pacific Ocean. *Journal of the Marine Biological Association of the United Kingdom* **100**, 1013–1021.

Boden B.P. (1954) The euphausiid crustaceans of southern African waters. *Transactions of the Royal Society of South Africa* **34**, 181–234.

Bolyen E., Rideout J.R., Dillon M.R., Bokulich N.A., Abnet C.C., Al-Ghalith G.A., Alexander H., Alm E.J., Arumugam M., Asnicar F., Bai Y., Bisanz J.E., Bittinger K., Brejnrod A., Brislawn C.J., Brown C.T., Callahan B.J., Caraballo-Rodríguez A.M., Chase J., Cope E.K., Da Silva R., Diener C., Dorrestein P.C., Douglas G.M., Durall D.M., Duvallet C., Edwardson C.F., Ernst M., Estaki M., Fouquier J., Gauglitz J.M., Gibbons S.M., Gibson D.L., Gonzalez A., Gorlick K., Guo J., Hillmann B., Holmes S., Holste H., Huttenhower C., Huttley G.A., Janssen S., Jarmusch A.K., Jiang L., Kaehler B.D., Kang K.B., Keefe C.R., Keim P., Kelley S.T., Knights D., Koester I., Kosciolek T., Kreps J., Langille M.G.I., Lee J., Ley R., Liu Y-X., Loftfield E., Lozupone C., Maher M., Marotz C., Martin B.D., McDonald D., McIver L.J., Melnik A.V., Metcalf J.L., Morgan S.C., Morton J.T., Naimey A.T., Navas-Molina J.A., Nothias L.F., Orchanian S.B., Pearson T., Peoples S.L., Petras D., Preuss M.L., Pruesse E., Rasmussen L.B., Rivers A., Robeson M.S. II, Rosenthal P., Segata N., Shaffer M., Shiffer A., Sinha R., Song S.J., Spear J.R., Swafford A.D., Thompson L.R., Torres P.J., Trinh P., Tripathi A., Turnbaugh P.J., Ul-Hasan S., van der Hooft J.J.J., Vargas F., Vázquez-Baeza Y., Vogtmann E., von Hippel M., Walters W., Wan Y., Wang M., Warren J., Weber K.C., Williamson C.H.D., Willis A.D., Xu Z.Z., Zaneveld J.R., Zhang Y., Zhu Q., Knight R. & Caporaso J.G. (2019) Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**, 852–857.

Brierley A.S. (2017) Plankton. *Current Biology* **27**, 478–483.

Brinton E., Ohman M.D., Townsend A.W., Knight M.D. & Bridgeman A.L. (2000) Euphausids of the World Ocean. ETI. CD-ROM.

Bucklin A., Nishida S. & Schiel S. (2004) Census of marine zooplankton: A new global survey of marine biodiversity. *Global Change News Letter* **60,** 8–10.

Bucklin A., Wiebe P.H., Smolenack S.B., Copley N.J., Beaudet J.G., Bonner K.G., Färber-Lorda J. & Pierson J.J. (2007) DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). *Journal of Plankton Research* **29**, 483–493.

Bucklin A., Steinke D. & Blanco-Bercial L. (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science* **3**, 471–508.

Bucklin A., Lindeque P.K., Rodriguez-Ezpeleta N., Albaina A. & Lehtiniemi M. (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *Journal of Plankton Research* **38**, 393–400.

Bucklin A., Peijnenburg K.T.C.A., Kosobokova K.N., O'Brien T.D., Blanco-Bercial L., Cornils A., Falkenhaug T., Hopcroft R.R., Hosia A., Laakmann S., Li C., Martell L., Questel J.M., Wall-Palmer D., Wang M., Wiebe P.H. & Weydmann-Zwolicka A. (2021) Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology* **168**, 78.

Callahan B.J., McMurdie P.J., Rosen M.J., Han A.W., Johnson A.J.A. & Holmes S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581–583.

Casanova, J.-P. (1999) Chaetognatha, in Boltovskoy, D. (ed.) *South Atlantic Zooplankton*. Leiden: Backhuys Publishers, pp. 1353–1374.

Chapple D.G. & Ritchie P.A. (2013) A retrospective approach to testing the DNA barcoding method. *PLoS ONE* **8**, e77882.

Deagle B.E., Jarman S.N., Coissac E., Pompanon F. & Taberlet P. (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters* **10**, 20140562.

Duró A. & Saiz E. (2000) Distribution and trophic ecology of chaetognaths in the western Mediterranean in relation to an inshore–offshore gradient. *Journal of Plankton Research* **22**, 339–361.

Folmer O., Black M., Hoeh W., Lutz R. & Vrijenhoek R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit 1 from diverse metazoan invertebrates.

Molecular Marine Biology and Biotechnology **3**, 294–299.

Gasmi S., Nve G., Pech N., Tekaya S., Gilles A. & Perez Y. (2014) Evolutionary history of Chaetognatha inferred from molecular and morphological data: a case study for body plan simplification. *Frontiers in Zoology* **11**, 84.

Gibbons M.J., Barange M. & Hutchings L. (1995) Zoogeography and diversity of euphausiids around southern Africa. *Marine Biology* **123**, 257–268.

Gibbons M.J. (1995) Observations on euphausiid assemblages of the south coast of South Africa. *South African Journal of Marine Science* **16**, 141–148.

Gibbons M.J. & Hutchings L. (1996) Zooplankton diversity and community structure around southern Africa, with special attention to the Benguela upwelling system. *South African Journal of Science* **92**.

Goetze, E., & Jungbluth, M. J. (2013). Acetone preservation for zooplankton molecular studies. *Journal of Plankton Research*, 35(5), 972–981. https://doi.org/10.1093/plankt/fbt035

Goswami S.C., Dhargalkar V.K. & Verlecar X.N. (2004) Zooplankton Methodology, Collection & Identification – a Field Manual. Goa: Bioinformatics Centre, National Institute of Oceanography.

Govender A., Groeneveld J., Singh S. & Willows-Munro S. (2022a) Metabarcoding of zooplankton confirms southwards dispersal of decapod crustacean species in the Western Indian Ocean. *African Journal of Marine Science* **44**, 279–289.

Govender A., Singh S., Groeneveld J., Pillay S. & Willows-Munro S. (2022b) Metabarcoding analysis of marine zooplankton confirms the ecological role of a sheltered bight along an exposed continental shelf. *Molecular Ecology* **32**, 6210–6222.

Govender A., Willows-Munro S., Singh S.P. & Groeneveld J.C. (2024) Net type, tow duration and day/night sampling effects on the composition of marine zooplankton derived from metabarcoding. *Metabarcoding and Metagenomics* **8**, 69–85.

Grant D.M., Brodnicke O.B., Evankow A.M., Ferreira A.O., Fontes J.T., Hansen A.K., Jensen M.R., Kalaycı T.E., Leeper A. & Patil S.K. (2021) The Future of DNA Barcoding: Reflections from Early Career Researchers. *Diversity* **13**, 313.

Guo M., Yuan C., Tao L., Cai Y. & Zhang W. (2022) Life barcoded by DNA barcodes. *Conservation Genetics Resources* **14**, 351–365.

Haddock S.M., Schuchert P. & Chapman J.C.P. (2004) Chaetognaths: Morphological and molecular perspectives. *Invertebrate Biology* **123**, 55–69.

Hall T.A. (1994) BioEdit: a user-friendly biological sequence alignment editor and analysis programme for Windows95/98/NT. *Nucleic Acids Symposium Series* **41**, 95–98.

Harris K.N.S. & Martin J.E.P. (2014) Genetic tools in taxonomy: applications for chaetognaths. *Molecular Phylogenetics and Evolution* **75**, 1–15.

Hebert P.D., Cywinska A., Ball S.L. & deWaard J.R. (2003) Biological identifications through DNA barcodes. Proceedings. *Biological Sciences* **270**, 313–318.

Johnson M.L. & Tarling G.A. (2017) Collecting and processing euphausiids. *Journal of Crustacean Biology* **37**, 228–231.

Kürzel K., Kaiser S., Lörz A., Rossel S., Paulus E., Peters J., Schwentner M., Martinez A.P., Coleman C.O., Svavarsson J. & Brix S. (2022) Correct Species Identification and Its Implications for Conservation Using Haploniscidae (Crustacea, Isopoda) in Icelandic Waters as a Proxy. *Frontiers in Marine Science* **8**.

Larkin M.A., Blackshields G., Brown N., Chenna R., McGettigan P., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J., Gibson T.J. & Higgins D.G. (2007) Clustal W and Clustal X Version 2.0. *Bioinformatics* **23**, 2947–2948.

Lee T.H.S.M., Li T.K.Y. & Lee H.S.M. (2005) Molecular systematics of chaetognatha: a review. *Journal of Plankton Research* **27**, 665–676.

Lowe M.R., Lawson G.L. & Fogarty M.J. (2018) Drivers of euphausiid distribution and abundance in the Northeast U.S. Shelf Large Marine Ecosystem. *ICES Journal of Marine Science* **75**, 1280–1295.

Macher J.K., Martinez A., Cakir S., Cholley P.-E., Christoforou E., Galletti M.C., van Galen L., Garcia-Cobo M., Jondelius U., de Jong D., Leasi F., Lemke M., Lopez R., Sanchez N., Sorensen M.V., Todaro M.A., Renema W. & Fontaneto D. (2024) Enhancing metabarcoding efficiency and ecological insights through integrated taxonomy and DNA reference barcoding: a case study on beach meiofauna. *Molecular Ecology Resources* **24**, 263–294.

Magnus M., Smith J. & Johnson R. (2017) Euphausiid diversity and distribution in global oceans. *Journal of Crustacean Biology* **37**, 228–240.

Muxagata E., Barbosa C.N. & Agostini V.O. (2019) Euphausiacea diversity in a trans-oceanic transect through the South Atlantic Ocean: the first Atlantic record of Thysanopoda astylata Brinton, 1975. *Anais da Academia Brasileira de Ciências* **91**, e20180034.

Nair V.R., Kidangan F.X., Prabhu R.G., Bucklin A. & Nair S. (2015) DNA barcode of Chaetognatha from Indian Waters. *Indian Journal of Geo-Marine Sciences* **44**, 1366–1376.

Perez Y., Müller C.H.G. & Harzsch S. (2021) Chapter 15: Chaetognatha. In: Schierwater B. & DeSalle R. (eds) Invertebrate Zoology: A Tree of Life Approach. CRC Press, 231.

Pillar S.C., Stuart V., Barange M. & Gibbons M.J. (1992) Community structure and trophic ecology of euphausiids in the Benguela ecosystem. *South African Journal of Marine Science* **12**, 393–410.

Pires A.C. & Marinoni L. (2010) DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica* **10**.

Ratnasingham S. & Hebert P.D.N. (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE* **8**, e66213.

Ratnarajah L., Abu-Alhaija R., Atkinson A., Batten S., Bax N.J., Bernard K.S., Canonico G., Cornils A., Everett J.D., Grigoratou M., Ishak N.H.A., Johns D., Lombard F., Muxagata E., Ostle C., Pitois S., Richardson A.J., Schmidt K., Stemmann L., Swadling K.M., Yang G. & Yebra L. (2023) Monitoring and modelling marine zooplankton in a changing climate. *Nature Communications* **14**, 564.

Richardson A.J. (2008) In hot water: zooplankton and climate change. *ICES Journal of Marine Science* **65**, 279–295.

Sell A.F., von Maltitz G.P., Auel H., Biastoch A., Bode-Dalby M., Brandt P., Duncan S.E., Ekau W., Fock H.O., Hagen W., Huggett J.A., Koppelmann R., Körner M., Lahajnar N., Martin B., Midgley G.F., Rixen T., van der Lingen C.D., Verheye H.M. & Wilhelm M.R. (2024) Unique Southern African Terrestrial and Oceanic Biomes and Their Relation to Steep Environmental Gradients. *Sustainability of Southern African Ecosystems in Ecological Studies* **248**.

Singh S., Groeneveld J., Huggett J., Naidoo D., Cedras R. & Willows-Munro S. (2021) Metabarcoding of marine zooplankton in South Africa. *African Journal of Marine Science* **43**, 1–13.

Song C., Choi H., Jeon M., Kim E., Jeong H.G., Kim S., Kim C., Hwang H., Purnaningtyas D.W., Lee S., Eyun S. & Lee Y. (2021) Zooplankton diversity monitoring strategy for the urban coastal region using metabarcoding analysis. *Scientific Reports* **11**, 24339.

Stuart V. & Pillar S.C. (1988) Growth and production of Euphausia lucens in the southern Benguela Current. *Journal of Plankton Research* **10**, 1099–1112.

Talbot M.S. (1974) Distribution of euphausiid crustaceans from the Agulhas Current. *Zoologica Africana* **9**, 93–145.

Wright K.J. (2014) Chaetognatha (arrow worms). In: Thorp J.H. & Covich A.P. (eds) Ecology and Classification of North American Freshwater Invertebrates. 4th ed. Amsterdam: *Elsevier*, 739–756.

APPENDIX 1: Barcode-compliant sequences of mainly Euphausiid and Chaetognaths together with associated photographs and collection data generated in this study were uploaded onto BOLD under the project code ORIA (FBIP - DNA barcoding of South African Marine Euphausiacea, Chaetognatha, and Amphipoda).

	BOLD Reference	**Family**	**Genus**	**Species**	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
1	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
2	N/A	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
3	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
4	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
5	N/A	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
6	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
7	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
8	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
9	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso	-27.925	32.619	200	02-Sep-2019

	BOLD Reference	**Family**	**Genus**	**Species**	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
10	N/A	Euphausiidae	*Euphausia*	*recurva*	Durban	-29.957	31.122	200	06-Sep-2019
11	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.893	29.421	200	14-Mar-2018
12	BOLD:AAH9922	Euphausiidae	*Stylocheiron*	*N/A*	Thukela Bank	-31.893	29.421	200	14-Mar-2018
13	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Richards Bay	-28.858	31.285	100	23-Aug-2019
14	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Richards Bay	-28.858	31.285	100	23-Aug-2019
15	BOLD:AEO2674	Euphausiidae	*Euphausia*	*recurva*	Richards Bay	-28.858	31.285	100	23-Aug-2019
16	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Richards Bay	-28.858	31.285	100	23-Aug-2019
17	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Richards Bay	-28.858	31.285	100	23-Aug-2019
18	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.923	32.613	100	02-Sep-2019
19	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.923	32.613	100	02-Sep-2019
20	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.923	32.613	100	02-Sep-2019
21	BOLD:AEO6753	Sagittidae	*Sagitta*	N/A	Thukela Bank	-31.868	29.439	100	14-Sep-2018
22	N/A	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.92	32.604	20	02-Sep-2019
23	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
24	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal	-30.289	30.917	200	15-Nov-2018

	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
25	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
26	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
27	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
28	BOLD:AEO5409	Euphausiidae	*Stylocheiron*	*carinatum*	Durban	-29.947	31.099	100	08-Feb-2021
29	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Aliwal Shoal MPA	-30.289	30.917	200	15-Nov-2018
30	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
31	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
32	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
33	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
34	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019

35	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
36	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
37	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
38	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
39	BOLD:AAM4451	Euphausiidae	Euphausia	recurva	Thukela Bank	-31.894	29.422	200	14-Sep-2018
40	BOLD:AAH9921	Euphausiidae	*Stylocheiron*	*N/A*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
41	BOLD:AEO2674	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
42	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
43	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
44	BOLD:AAM4451	Euphausiidae	*Euphausia*	*diomedeae*	ISimangaliso MPA	-27.925	32.619	200	02-Sep-2019
45	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
46	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
47	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.422	200	14-Sep-2018
48	N/A	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.894	29.423	200	14-Sep-2018
49	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Thukela Bank	-31.895	29.423	200	14-Sep-2018
50	N/A	Sagittidae	*Flaccisagitta*	*enflata*	Aliwal Shoal	-30.289	30.917	200	15-Nov-2018

	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
51	BOLD:AEQ8811	N/A	N/A	N/A	Richards Bay	-28.858	31.285	100	23-Aug-2019
52	BOLD:AAM4451	Euphausiidae	N/A	N/A	Thukela Bank	-31.868	29.439	100	14-Sep-2018
53	BOLD:AAY3432	Euphausiidae	*Euphausia*	*similis*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
54	BOLD:AAJ4613	Euphausiidae	*Stylocheiron*	*elongatum*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
55	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
56	BOLD:ADZ4700	Euphausiidae	*Hansarsia*	*megalops*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
57	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
58	BOLD:AFQ8238	Euphausiidae	*Hansarsia*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
59	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
60	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
61	BOLD:ADK3324	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
62	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
63	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
64	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
65	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
66	BOLD:AFS4748	Euphausiidae	*Hansarsia*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
67	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
68	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
69	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
70	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023

	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
71	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
72	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
73	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
74	N/A	Euphausiidae	*Stylocheiron*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
75	BOLD:AAM4451	Euphausiidae	Euphausia	recurva	Agulhas Bank	35.197	23.513	200	01-Jul-2023
76	BOLD:AAM4451	Euphausiidae	Euphausia	recurva	Agulhas Bank	35.197	23.513	200	01-Jul-2023
77	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
78	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
79	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
80	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
81	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
82	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
83	BOLD:ADK3324	Euphausiidae	*Euphausia*	*mutica*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
84	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
85	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
86	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
87	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
88	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
89	BOLD:ADK3324	Euphausiidae	*Euphausia*	*mutica*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
90	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023

	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
91	BOLD:ABZ1024	Euphausiidae	*Euphausia*	*lucens*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
92	BOLD:AEH5036	Euphausiidae	*Hansarsia*	*microps*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
93	BOLD:ABZ1024	Euphausiidae	*Euphausia*	*lucens*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
94	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
95	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
96	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.197	23.513	200	01-Jul-2023
97	BOLD:AAK1821	Euphausiidae	*Thysanoessa*	*gregaria*	Agulhas Bank	35.197	23.513	200	01-Jul-2023
98	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
99	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
100	BOLD:AFQ2904	Euphausiidae	*Stylocheiron*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
101	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
102	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
103	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
104	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
105	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
106	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
107	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
108	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
109	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
110	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023

	BOLD Reference	**Family**	***Genus***	***Species***	**Geographic Collection Site**	**Latitude**	**Longitude**	**Depth**	**Collection Date**
111	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
112	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
113	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
114	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
115	BOLD:AAM4451	Euphausiidae	*Euphausia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
116	BOLD:AAM4451	Euphausiidae	*Euphausia*	*recurva*	Agulhas Bank	35.462	23.491	200	01-Jul-2023
117	BOLD:AFQ8238	Euphausiidae	*Hansarsia*	N/A	Agulhas Bank	35.462	23.491	200	01-Jul-2023
118	BOLD:AEO2674	Euphausiidae	*Euphausia*	*diomedeae*	Agulhas Bank	35.462	23.491	200	01-Jul-2023

N/A- NOT ASSIGNED

Chapter Four

General Discussion

Zooplankton play an important role in aquatic food webs, by facilitating nutrient cycling and energy transfer to higher trophic levels (Bucklin *et al*. 2021a). They also serve as indicators of ecosystem health (Atkinson *et al*. 2008), as changes in zooplankton community assemblages are linked to environmental shifts (Richardson 2008). To be useful as a biomonitoring tool, a system of rapid, identification of zooplankton species is crucial. This study explored the utility of DNA barcoding, particularly using the COI gene, for the identification of Euphausiid and Chaetognath species.

Chapter One introduces DNA barcoding as a valuable tool to improve the pace and accuracy of species identification in two key zooplankton groups, namely Euphausiids (Lowe *et al.* 2018) and Chaetognaths (Bieri and Kormanol 2020). This chapter discusses taxonomic challenges associated with these groups, such as low genetic diversity, cryptic speciation (Wei *et al.* 2018; Bucklin *et al*. 2021b), varied morphology based on developmental changes (Kürzel *et al.* 2022; Podbielski *et al.* 2024), and undescribed species together with older taxonomic descriptions due to sampling challenges (Harris and Martin 2014; Stone and Stone 2023). Euphausiids and Chaetognaths are both considered understudied in South Africa (Singh *et al.* 2021), providing the impetus for this MSc study. The study aimed to evaluate the effectiveness of the COI gene as a DNA barcode marker by statistically verifying the presence of a DNA barcode gap and improving the representation of these taxa on global databases.

Chapter two focused on statistical testing for the presence of the DNA barcode gap at species, genus and family levels using COI datasets obtained from BOLD and GenBank for Euphausiid and Chaetognath. In Euphausiids, a clear DNA barcode gap was observed at the species level (at 2-3% sequence divergence; J-M test value=1.88) and genus level (at 5-6 % sequence divergence; J-M test value = 1.99). However, intra- and interfamilial genetic distances showed a slight overlap (at 8-10% sequence divergence; J-M test value =1.99). For Chaetognaths, overlaps were found at species, genus and family levels, with species showing a divergence of 4–5% (J-M test value = 1.80) and genera at 7–8% (J-M test value = 1.80). No clear DNA barcode gap was observed at the family level for Chaetognaths, highlighting the challenges in their identification due to fragile morphology and limited diagnostic traits. The results of this study confirm that COI barcode data can be used to separate species, genera, and families belonging to Euphausiids, however the findings for Chaetognaths suggest that while COI is effective for species and genus identification, additional marker genes are needed for robust family-level delimitation.

Chapter three focused on expanding the global DNA barcode reference library. Samples were collected from five localities off South Africa's East Coast (iSimangaliso Wetland Park, Richards Bay, uThukela Marine Protected Area (MPA), Durban and Aliwal Shoal MPA (Figure 3-1)), taxonomically identified, photographed, and sequenced. All photographs, collection data and sequence data for new barcode records were uploaded to BOLD under the project code ORIA (FBIP - DNA barcoding of South African Marine Euphausiacea, Chaetognatha, and Amphipoda). This improves the available DNA reference resources for these taxa in South African waters. This study evaluated the current reference library's effectiveness in identifying species using data from a recent metabarcoding study (Govender *et al.* 2024). This study added 2 Chaetognath barcodes and 114 Euphausiid barcodes to the Barcode of Life Data System (BOLD), including new records for *Euphausia recurva*, *E. mutica*, and *Hansarsia microps*, with *E. recurva*, and *Thysanoessa gregaria* representing previously unrecorded taxa in South Africa. This enhanced reference database improves metabarcoding accuracy and facilitates better biodiversity monitoring and conservation efforts. Accurate species detection can help monitor the distribution of species and changes in environmental factors that affect zooplankton. This study, in turn demonstrates the critical role of DNA barcoding in improving metabarcoding for biodiversity monitoring, species diversity assessments, and conservation efforts. Despite the additions made by this study, there is further need for more taxonomically verified additions as many taxonomic ambiguities due to morphological variability and complex life stages hinder progress. Adding barcodes for underrepresented species will improve metabarcoding reliability and results in the future.

In conclusion, this MSc study has provided evidence that DNA barcoding is an effective tool for species identification and that the COI DNA barcode is effective for species delimitation at species and genera levels for both Euphausiids and Chaetognaths. However, Chaetognaths require additional genetic markers for family-level identification due to their morphological plasticity influenced by environmental factors (DeHart *et al.* 2020). New sequence records have been added onto BOLD with samples from South African waters, for Chaetognaths and Euphausiids, however, further sequence contributions are required to establish robust reference libraries with DNA barcode verified with taxonomic data which can be used to identify partial or degraded samples and further metabarcoding studies. The methodologies used in this study can be applied to other marine regions, contributing to the global effort to monitor biodiversity, track invasive species, and inform conservation strategies. By improving our understanding of zooplankton diversity and distribution, this research underscores the critical role of DNA barcoding in advancing marine science and ecological monitoring.

4.1. References:

Atkinson A., Siegel V., Pakhomov E.A., Rothery P., Loeb V.J. & Hewitt R. (2008) A reevaluation of the role of krill in the Southern Ocean food web. *Antarctic Science* **20**, 145–157.

Bucklin A., Peijnenburg K.T.C.A., Kosobokova K.N., O'Brien T.D., Blanco-Bercial L., Cornils A., Falkenhaug T., Hopcroft R.R., Hosia A., Laakmann S., Li C., Martell L., Questel J.M., Wall-Palmer D., Wang M., Wiebe P.H. & Weydmann-Zwolicka A. (2021a) Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology* **168**, 78.

Bucklin A., Questel J.M., Blanco-Bercial L., Frenzel A., Smolenack S.B. & Wiebe P.H. (2021b) Population connectivity of the euphausiid, Stylocheiron elongatum, in the Gulf Stream (NW Atlantic Ocean) in relation to COI barcode diversity of Stylocheiron species. ICES *Journal of Marine Science* **78**, 3464–3476.

DeHart H., Bucklin A. & Peijnenburg K.T.C.A. (2020) Integrative taxonomy of marine planktonic Chaetognatha: combining mitochondrial and nuclear markers to resolve species boundaries. *Marine Biodiversity* **50**, 1–15.

Govender A., Willows-Munro S., Singh S.P. & Groeneveld J.C. (2024) Net type, tow duration and day/night sampling effects on the composition of marine zooplankton derived from metabarcoding. *Metabarcoding and Metagenomics* **8**, 69–85.

Harris K.N.S. & Martin J.E.P. (2014) Genetic tools in taxonomy: applications for chaetognaths. *Molecular Phylogenetics and Evolution* **75**, 1–15.

Kürzel K., Kaiser S., Lörz A., Rossel S., Paulus E., Peters J., Schwentner M., Martinez A.P., Coleman C.O., Svavarsson J. & Brix S. (2022) Correct species identification and its implications for conservation using Haploniscidae (Crustacea, Isopoda) in Icelandic waters as a proxy. *Frontiers in Marine Science* **8**.

Lowe M.R., Lawson G.L. & Fogarty M.J. (2018) Drivers of euphausiid distribution and abundance in the Northeast U.S. Shelf Large Marine Ecosystem. ICES *Journal of Marine Science* **75**, 1280–1295.

Podbielski I., Hamm T. & Lenz M. (2024) Customized digestion protocols for copepods, euphausiids, chaetognaths, and fish larvae facilitate the isolation of ingested microplastics. *Scientific Reports* **14**, 19985.

Richardson A.J. (2008) In hot water: zooplankton and climate change. ICES *Journal of Marine Science* **65**, 279–295.

Singh S., Groeneveld J., Huggett J., Naidoo D., Cedras R. & Willows-Munro S. (2021) Metabarcoding of marine zooplankton in South Africa. *African Journal of Marine Science* **43**, 1–13.

Stone S.E. & Stone J.P. (2023) Diversity, seasonal abundance, and environmental drivers of chaetognath populations in North Inlet Estuary, South Carolina, USA. *Ecology and Evolution* **13**, e10151.

Wei S., Yang M., Dong Y. & Qin Q. (2018) Genetic variation analysis of the cosmopolitan chaetognath Sagitta enflata in the northern South China Sea based on mitochondrial COI gene sequences. *Mitochondrial DNA B Resources* **4**, 5–7