PhD in Aquatic Biosciences // No. 51 - 2023

Biosystematics and evolutionary genomics of deep-sea fish (lumpsuckers, snailfishes, and sculpins) (Perciformes: Cottoidei)

Likith Reddy Pinninti

FACULTY OF BIOSCIENCES AND AQUACULTURE

NORD University

www.nord.no

Biosystematics and evolutionary genomics of deep-sea fish (lumpsuckers, snailfishes, and sculpins) (Perciformes: Cottoidei)

Likith Reddy Pinninti

A thesis for the degree of
Philosophiae Doctor (PhD)

PhD in Aquatic Biosciences no. 51 (2023)
Faculty of Biosciences and Aquaculture

PhD in Aquatic Biosciences no. 51 (2023)

Likith Reddy Pinninti
Biosystematics and evolutionary genomics of deep-sea fish (lumpsuckers, snailfishes, and sculpins) (Perciformes: Cottoidei)

ISBN: 978-82-93165-50-7

Print: Trykkeriet NORD

Nord University
N-8049 Bodø
Tel: +47 75 51 72 00
www.nord.no

Preface

The thesis is submitted in fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the faculty of Biosciences and Aquaculture (FBA), Nord University. The different studies compiled in this dissertation are original research performed at Nord University, Bodø over a period of three years. The studies were funded by the Norwegian Government and Nord University.

PhD project team:

Likith Reddy Pinninti: M.Tech, FBA, Nord University: PhD candidate.

Leslie Robert Noble: Professor, FBA, Nord University: primary supervisor.

Torstein Kristensen: Associate Professor, FBA, Nord University: co- supervisor.

Jorge Manuel de Oliveira Fernandes: Professor, FBA, Nord University: co- supervisor.

Likith Reddy Pinninti

Bodø, April 2023

Acknowledgments

First of all, I would like to thank my supervisor, Leslie, for his kind support and guidance throughout the PhD period. I am very grateful to him for providing me with the opportunity to work on this challenging research project. Throughout my PhD, there were both strokes of ill-luck and celebrating moments, and I am very glad to have come across a very humble and polite PhD supervisor with whom I am able to share my thoughts. Leslie is not only my boss but also a good friend. I am very grateful to him for his overall support.

I would also like to thank my co-supervisor, Torstein Kristeinsen, for sharing his knowledge about lumpfish ecology. His ideas assisted me in realising the behaviour and ecology of lumpfish. I would also like to thank Jorge Fernandes Manuel Oliveria for his full support in structuring and writing the manuscripts and planning the experiments related to molecular biology. I also want to thank Martina and Irina for helping with the lab work on time and getting the chemicals on time.

Last but not least, I would like to thank all the members of the ecology division and the PhD fellows, both past and present, for helping and sharing their knowledge during the FBA and the ecology group meeting sessions. Also, I would like to thank our administration and technicians on campus. I am also very grateful to my collaborators in the US and Norway, who helped me finish my thesis and taught me new things.

Finally, I would like to thank all my family members and friends for their affection and support throughout my academic journey.

Table of Contents

List of abbreviations

AMH - Anti-Müllerian Hormone

CDS - Coding Sequences

COI - Cytochrome C Oxidase Subunit I

DEGs - Differentially Expressed Genes

D-loop - Displacement loop

gCF - Gene Concordance Factors

HSDs - Highly Similar Duplicates

HTS - High-Throughput Sequencing

LTP - Long Term Potentiation

mtDNA - Mitochondrial DNA

MYA - Million Years Ago

PCGs - Protein Coding Genes

PSO - Pelvic Suctorial Organ

RNA-Seq - RNA sequencing

sCF - Site Concordance Factors

WGS - Whole Genome Sequencing

List of figures and tables

List of papers

Paper I:

Pinninti, L. R., Maurstad, M. F., Hoff, S. N., Kristensen, T., Noble, L. R., Jentoft, S., & Fernandes, J. M. (2023). The complete mitochondrial genome of the Atlantic spiny lumpsucker *Eumicrotremus spinosus* (Fabricius, 1776). *Mitochondrial DNA Part B*, *8*(3), 364-367.

Paper II:

Pinninti, L.R., Fernandes, J.M.O., Smolina, I., Collard, F., Eulaers, I., Noble, L.R. (2023). Understanding the evolutionary landscape of lumpsuckers, snailfishes, and sculpins of the suborder Cottoidei (Teleostei: Perciformes).(Manuscript).

Paper III:

Pinninti, L.R., Staven, F.R., Noble, L.R., Fernandes, J.M.O., Patel D.M., Kristensen, T., (2023). A search for hormone genes in cleaner fish (*Cyclopterus lumpus* L.) during habituation. (Manuscript).

Abstract

The deep-sea fishes of the Cottoidei (Teleostei: Perciformes) comprise an extremely diverse group, having developed unique adaptations for survival in the deep ocean, with distinct morphological and genetic differences compared to their shallow-water counterparts. Elucidating the systematics of deep-water taxa is difficult, although a better appreciation of deep-sea fish evolution promises greater insights into vertebrate evolution and may support conservation of these unique ecosystems. Nevertheless, deep-sea fishes remain poorly understood in terms of their evolutionary history and behaviour, despite their commercial value. Hence, to understand their specific adaptations and evolutionary relations, genomic-based approaches were used throughout this thesis to explain the organisation of their genomes, genome history, and behaviour.

The first study presents a revised phylogeny of the suborder Cottoidei (Teleostei: Perciformes). A PacBio long-read genomics approach was used to obtain the mitogenome of a member of the Atlantic spiny lumpsucker *(Eumicrotremus spinosus)*, a member of the Cyclopteridae family. The complete single scaffold of the mitogenome amounted to 19.2 kb. Utilising available closely related mitogenomes, a revised phylogeny was constructed for the three groups of benthic fish (lumpsuckers, snailfish, sculpins) comprising the suborder Cottoidei (Teleostei: Perciformes).

The second study, based on a short-read genomics approach, produced six low-coverage genomes (30–40X) of three Cyclopteridae and three Liparidae species. These data were used in a comparative genomics study together with available genomes of nine additional closely related taxa belonging to the suborder Cottoidei (Teleostei: Perciformes). Orthologs were used to construct a species tree using super tree and super matrix-based approaches, to determine intraspecies evolutionary relationships within the Cottoidei (lumpsuckers, snailfishes, and sculpins). Additionally, 1880 loci were examined for evidence of selection. Positive selection was detected at 160 loci,

including genes for adipogenesis, DNA repair, and DNA translation in these benthic fish. Highly similar duplicates (HSDs) involved in the metabolism of carbohydrates, energy, and lipids were identified. Some of these HSDs may be artefacts. Future research should focus on generating high-quality genomes and determining the evolutionary rates of duplicated genes using validated functional genomics approaches.

A third study employed a transcriptomics-based approach (RNA-Seq) to identify sequences associated with cooperation-related behavioural traits in a cleaner fish (the lumpfish, *Cyclopterus lumpus* L.). Understanding the transcriptomics underlying such traits can be useful in the context of animal welfare. Brain transcripts were obtained from short-term (1 h 30 min) inter-species experiments designed to study cooperative behaviour (mutualism) of cleaner fish (Atlantic lumpfish, *Cyclopterus lumpus* L.) in interactions with client fish (Atlantic salmon, *Salmo salar* L.). Differential gene expression (DGEs) analysis revealed the majority of sequences to be false positives, none was identified as associated with any behavioural traits. However, correlation analysis identified genes associated with locomotion, serotonin, and pigmentation. The conclusion is based solely on correlation analysis, which is less credible than functional transcriptomic associations. Future work should focus on gene expression analysis in different parts of the brain of both species during mutualistic interactions, so that the transcriptomic synchronisation, learning, and memory during the process of mutualism can be better understood.

The thesis concludes with a discussion of the utility of a phylogenomics approach to address some novel and fundamental knowledge gaps regarding the evolution of deep-sea fish and their adaptations to benthic habitats. Finally, we reconstructed the complete mitogenome of *E. spinosus* using long reads, which will be a very useful resource for the taxonomy and systematics studies. Pacbio is better for mtDNA reconstruction, and Illumina is more cost-effective for phylogenomics.

The genome resources we created for data-deficient species are valuable and will be useful in future systematic and taxonomic studies. Furthermore, cooperation-related behavioural traits were identified from the correlation analysis between the noradrenaline hormone levels and gene expression data. Correlations identified genes associated with pigmentation, skin colour, locomotion, serotonin, and dopamine synthesis. The findings from this study may help the aquaculture community understand *C. lumpus* behaviour when they interact with the client fish during the habituation process.

Keywords: Behaviour, Genome, Lumpsuckers, Phylogenetics, Brain transcriptomics, Mitogenomics

Sammendrag på Norsk

Dyphavsfiskene til Cottoidei (Teleostei: Perciformes) utgjør en ekstremt mangfoldig gruppe, etter å ha utviklet unike tilpasninger for å overleve i dyphavet, med distinkte morfologiske og genetiske forskjeller sammenlignet med deres motparter på grunt vann. Å belyse systematikken til dypvannstaxa er vanskelig, selv om en bedre forståelse av dyphavsfiskens utvikling lover større innsikt i virveldyrenes utvikling og kan støtte bevaring av disse unike økosystemene. Likevel er dyphavsfisk fortsatt dårlig forstått når det gjelder deres evolusjonære historie og oppførsel, til tross for deres kommersielle verdi. Derfor, for å forstå deres spesifikke tilpasninger og evolusjonære relasjoner, ble genomisk-baserte tilnærminger brukt gjennom denne oppgaven for å forklare organiseringen av deres genomer, genomhistorie og atferd.

Den første studien presenterer en revidert fylogeni av underordenen Cottoidei (Teleostei: Perciformes). En PacBio langlest genomikk-tilnærming ble brukt for å oppnå mitogenomet til et medlem av den atlantiske ryggsøylen (*Eumicrotremus spinosus*), et medlem av Cyclopteridae-familien. Det komplette enkeltstillaset til mitogenomet utgjorde 19,2 kb. Ved å bruke tilgjengelige nært beslektede mitogenomer, ble en revidert fylogeni konstruert for de tre gruppene av bunnfisk (klumpsugere, sneglefisk, sculpiner) som omfatter underordenen Cottoidei (Teleostei: Perciformes).

Den andre studien, basert på en kortlest genomikk-tilnærming, produserte seks lavdekkende genomer (30–40X) av tre Cyclopteridae og tre Liparidae-arter. Disse dataene ble brukt i en komparativ genomikkstudie sammen med tilgjengelige genomer av ni ytterligere nært beslektede taxa som tilhører underordenen Cottoidei (Teleostei: Perciformes). Orthologer ble brukt til å konstruere et artstre ved å bruke supertre og supermatrisebaserte tilnærminger, for å bestemme intraarts evolusjonære forhold innenfor Cottoidei (klumpsugere, sneglefisker og sculpiner). I tillegg ble 1880 loci undersøkt for bevis på seleksjon. Positiv seleksjon ble påvist ved 160 loci, inkludert gener for adipogenese, DNA-reparasjon og DNA-translasjon i disse bunnfiskene. Svært

lignende duplikater (HSDs) involvert i metabolismen av karbohydrater, energi og lipider ble identifisert. Noen av disse HSD-ene kan være gjenstander. Fremtidig forskning bør fokusere på generering av genomer av høy kvalitet og å bestemme evolusjonsratene til dupliserte gener ved å bruke validerte funksjonelle genomiske tilnærminger.

En tredje studie brukte en transkriptomikk-basert tilnærming (RNA-Seq) for å identifisere sekvenser assosiert med samarbeidsrelaterte atferdstrekk hos en rensefisk (rognkjeksen, *Cyclopterus lumpus* L.). Å forstå transkriptomikken som ligger til grunn for slike egenskaper kan være nyttig i forbindelse med dyrevelferd. Hjernetranskripsjoner ble oppnådd fra kortsiktige (1 t 30 min) inter-arts-eksperimenter designet for å studere samarbeidsatferd (gjensidighet) av rensefisk (Atlantisk rognkjeks, *Cyclopterus lumpus* L.) i interaksjoner med klientfisk (Atlantisk laks, *Salmo salar* L. .). Differensiell genekspresjon (DGE)-analyse avslørte at flertallet av sekvensene var falske positive, ingen ble identifisert som assosiert med noen atferdstrekk. Imidlertid identifiserte korrelasjonsanalyse gener assosiert med bevegelse, serotonin og pigmentering. Konklusjonen er utelukkende basert på korrelasjonsanalyse, som er mindre troverdig enn funksjonelle transkriptomiske assosiasjoner. Fremtidig arbeid bør fokusere på genekspresjonsanalyse i forskjellige deler av hjernen til begge arter under mutualistiske interaksjoner, slik at den transkriptomiske synkroniseringen, læringen og hukommelsen under prosessen med mutualisme kan bli bedre forstått.

Avhandlingen avsluttes med en diskusjon av nytten av en fylogenomisk tilnærming for å adressere noen nye og grunnleggende kunnskapshull angående utviklingen av dyphavsfisk og deres tilpasning til bunnlevende habitater. Til slutt rekonstruerte vi det komplette mitogenomet til *E. spinosus* ved å bruke lange avlesninger, som vil være en svært nyttig ressurs for taksonomi- og systematikkstudiene. Pacbio er bedre for mtDNA-rekonstruksjon, og Illumina er mer kostnadseffektiv for fylogenomikk.

Genomressursene vi opprettet for arter som mangler data er verdifulle og vil være nyttige i fremtidige systematiske og taksonomiske studier. Videre ble samarbeidsrelaterte atferdstrekk identifisert fra korrelasjonsanalysen mellom noradrenalinhormonnivåene og genuttrykksdata. Korrelasjoner identifiserte gener assosiert med pigmentering, hudfarge, bevegelse, serotonin og dopaminsyntese. Funnene fra denne studien kan hjelpe akvakultursamfunnet til å forstå *C. lumpus* atferd når de samhandler med klientfisken under tilvenningsprosessen.

Nøkkelord: Oppførsel, Genom, Klumpsugere, Fylogenetikk, Hjernetranskriptomikk, Mitogenomikk

1 Introduction

1.1 Morphology and taxonomy of the suborder Cottoidei

The Cottoidei (lumpsuckers, snailfishes, and sculpins) is one of the most morphologically diverse teleostei suborders, consisting of 29 families classified into six infraorders based on a revised phylogeny using mtDNA markers (Smith and Busby, 2014, Maduna et al., 2022). They have adapted to low-temperature habitats (polar fronts) and are abundant in the North Atlantic, Pacific, and Arctic (Davenport, 1985, Holst, 1993). The majority live in deep-sea habitats at depths of between 200 and 8,000 m (Oguri and Noguchi, 2017, Gerringer, 2019) but some species of sculpin inhabit freshwater at depths between 200 and 1,600 m (Sideleva and Fialkov, 2015, Radnaeva et al., 2017).

Members of the lumpsucker family (Cyclopteridae) have a large pelvic suctorial organ (PSO), tubercles on their body surface, and a compressed tail, and small gills. Lumpsuckers are the closest relatives of the snailfishes (Liparidae); these taxa are characterised by a globular body shape, two short dorsal fins, and an anal fin (Figure 1). Both families have undergone convergent evolution (presence of suction disc = a synapomorphy), in their adaptation to the deep-sea benthic habitat, developing into poor swimmers lacking a swim bladder (Gerringer et al., 2021).

Sculpins (Cottoidea), one of the closest relatives of lumpsuckers and snailfishes, have developed an attachment organ in place of a PSO. Like the other two groups, sculpins also lack a swim bladder and are poor swimmers, displaying homologous structures found in both Cyclopteridae and Liparidae family members (lumpsuckers and snailfishes) (Popper and Fay, 2011). In strong ocean currents, members of all three families use an attachment organ to save energy by adhering to available surfaces.

Since 1985, the lumpfish (*Cyclopterus lumpus*) has been a primary target of a North-Atlantic fishery, in waters of the United States, Canada, Greenland, Iceland, the

United Kingdom, and Norway. It is valued as a cleaner fish and for its inexpensive caviar (Davenport, 1985, Mecklenburg and Sheiko, 2003). There are approximately 30 species in eight genera of the family Cyclopteridae *(Aptocyclus, Cyclopsis, Cyclopteropsis, Cyclopterus, Eumicrotremus, Georgimarinus, Lethotremus,* and *Proeumicrotremus*). The large number of species comprising the Liparidae (416 species) and Cottidae (289 species) families make them the most diverse taxa of the suborder Cottoidei; however, their interrelationships are difficult to comprehend. Classification has relied heavily on DNA barcoding (Knudsen et al., 2007, Smith and Busby, 2014). Consequently, the taxonomic status of a majority of species within this genus is unknown; they appear near identical as juveniles and are even difficult to distinguish as adults. Taxonomic research is retarded by a lack of recognizable and agreed diagnostic differences (anatomical and morphometric) between genera (Schmidt, 1927, Byrkjedal et al., 2007, Gerringer et al., 2021).

On the basis of morphological characteristics, *Cyclopterus* and *Eumicrotremus* (Cyclopteridae) species are easily distinguishable. There are approximately 16 species recognized in the genus *Eumicrotremus,* each sporting many bony tubercles on their body surfaces (Ueno, 1970, Mecklenburg and Sheiko, 2003, Byrkjedal et al., 2007). Based on gonadal observations, the holotypes of *E. eggivinii* and *E. spinosus* were recognized respectively as a male and a female of the same *Eumicrotremus* species following DNA barcoding of their COI DNA sequences, which were identical (Byrkjedal et al., 2007)**.** This study revealed that the taxonomy of the Cyclopteridae and Liparidae species was complicated by extreme sexual dimorphism (Byrkjedal et al., 2007, Hatano et al., 2015).

Due to an absence of scales and their gelatinous bodies, the Liparidae family is extremely difficult to study. MtDNA markers and morphological characteristics, including a reduced suction disc, pectoral girdle, tooth patterns, and cephalic pores, have been utilised to determine the relationships of taxa. Many taxonomic studies remain unresolved due to a lack of samples and the diversity of species (Chernova,

2008, Orr et al., 2019, Gerringer et al., 2021). Similarly, diversity and complexity of taxa within the Cottidae family also presents taxonomic difficulties.

Members of the Cyclopteridae are identified primarily by geographic location (Atlantic lumpsuckers versus Pacific lumpsuckers) and morphological characteristics of the fish; there is plenty of variation between them, and Pacific lumpsuckers are smaller (2.5 cm long), exhibit different colour morphs (yellow, orange, green, purple, and brown), and have more tubercles on their surfaces compared to Atlantic lumpsuckers (35-55 cm in size) (https://blogs.scientificamerican.com). This pair of lumpsuckers was likely separated aeons ago, (median time: 9.2 MYA; CI: 9.0–22.6 MYA; adjusted time: 10.8 MYA) and developed distinct characteristics (Rabosky et al., 2018). However, evolutionary and phylogenetic studies using genetic data have been scarce in the Cottoidei group to date. Historically, the classification of species has been based solely on morphological criteria. A novel departure is to use whole genome sequencing (WGS) data to reconstruct a phylogeny, to reveal the relationships of the Cottoidei subfamily, and determine the genetic basis of their evolutionary adaptations.

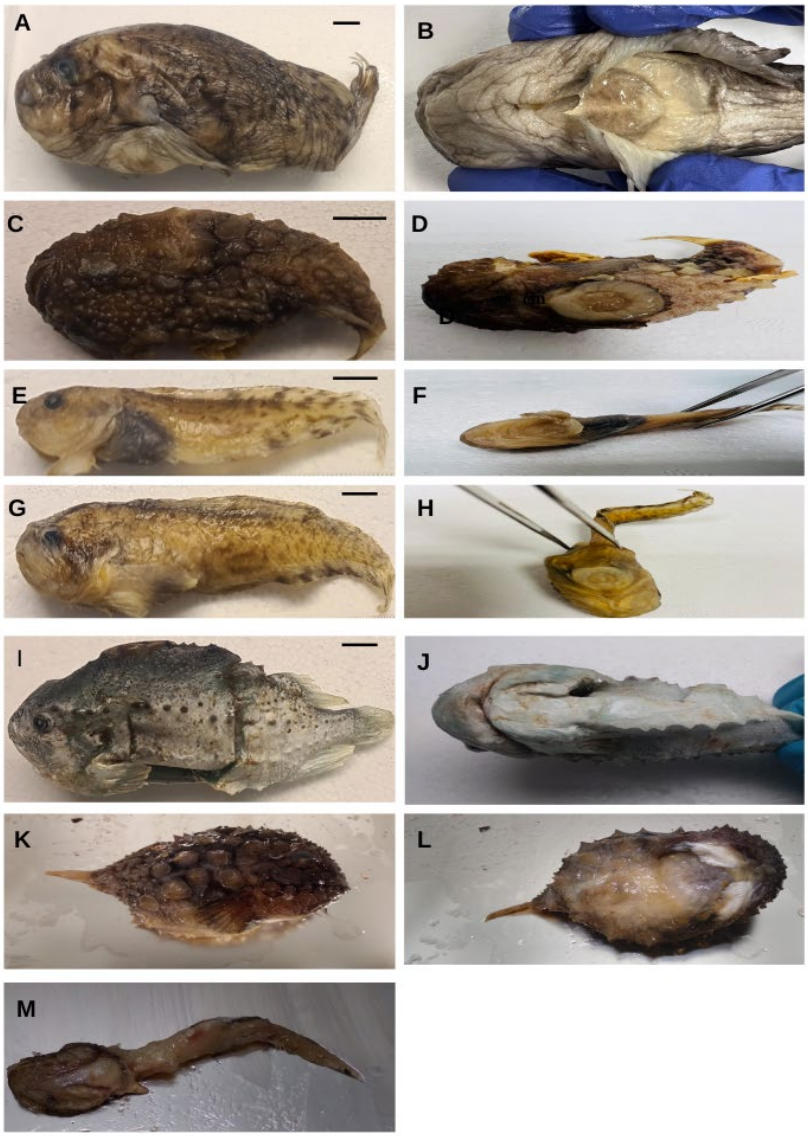


Figure 1. Images of the dorsal (cm in scale bar) and ventral surfaces of seven deep-sea fish from the Cyclopteridae and Liparidae families. There were no scales and tubercles on the dorsal side (A) of the brownish-grey smooth lumpfish (*A. ventricosus*), and a large PSO was present on the ventral side (B), which is considered to be a synapomorphic character. Spiny tubercles on the body surfaces of the *Eumicrotremus* genus (*E. derjugini*, shown in C and D, and *E. spinosus*, shown in K and L) served as an identifying feature. For the *Liparis* genus (shown in E, F, G, H, and M), they have an elongated body with a reduced suction disc on the ventral side and less pigment. The native lumpfish (*C. lumpus*, shown in I and J) are easily distinguishable based on their dark green skin colour and absence of scales on the body surface.

1.2 Biogeography of the Cottoidei (lumpsuckers, snailfishes and sculpins)

The life histories of fish species vary according to their biogeographic affiliation (Arctic, Boreal, and Arctic-boreal). Therefore, Arctic fish (*E. spinosus, E. derjugini, L. gibbus, and L. fabricii*) are smaller, mature faster, and produce more offspring than boreal fish (*C. lumpus*). Polar region species are phylogenetically more closely related and experience a harsher physical environment than boreal species (Wiedmann et al., 2014).

The majority of taxa in the Cyclopteridae, Liparidae, and Cottidea are adapted to benthic and benthopelagic environments. Many species of bottom-dwelling fish have evolved a PSO and inhabit depths between 50 and 900 meters (Chernova et al., 2014). Nevertheless, some species, such as the *(Pseudoliparis yapensis)* (Mu et al., 2021) and the *(Pseudoliparis swirei)* (Wang et al., 2019), live their entire lives on the ocean floor and have adapted to the extreme hydrostatic pressures, high temperatures around hydrothermal vents and the lower temperatures beyond, darkness, and limited food supply (Gerringer, 2019). In contrast to lumpsuckers and sculpins, Liparidae inhabit depths of 2,000 to 8,000 meters. They are widespread in the Pacific and Atlantic and are recognised as the deepest living fish (Gerringer et al., 2021).

Cottidae are also widely dispersed, particularly in boreal and cold waters; some species are marine, while others inhabit freshwater environments. Some freshwater sculpins are unique to Lake Baikal, adapted to a range of depths (50 - 1600 m). Shallow water species inhabit depths of up to 350 m, eurybathic species inhabit depths between 50 and 1300 m, and abyssal species inhabit depths between 400 and 1600 m. These fish species have developed various adaptations (abilities) to withstand the hydrostatic pressures in deep water (Sideleva, 1996). Why certain species have migrated from shallow water to greater depths is unclear. The knowledge gaps regarding distribution and biogeography are murky.

The most recent global oceanic anoxic events occurred during the Cretaceous geological period (145–66 million years ago (Mya)), when the Atlantic Ocean opened up, allowing teleost diversification (Actinopterygii), with many fish families adapting to different depths. Fish species that inhabit depths greater than 1,000 meters are regarded as deep-sea fish. Extinctions in the deep sea were consistently balanced by the migration of species from shallow to deep water (Priede and Froese, 2013).

1.3 Convergent evolution

Convergent evolution happens when natural selection occurs in unrelated or distantly related groups of organisms from different ancestral lineages. But they actually evolve similar adaptations due to similar selection pressures in their environments (Figure 2). The definitive example is the evolution of molluscan eye diversity (Serb and Eernisse, 2008). Sharks or fast-swimming fish versus dolphins bodies: unrelated or distantly related organisms evolve similar adaptations under similar selection pressures (Donley et al., 2004, Liu et al., 2010). At the level of the gene, similar phenotypes are produced by non-homologous genes. Convergent evolution suggests a mechanism underpinning the evolutionary transition (character evolution). It is very important to know the underlying pathway involved in the formation of the character evolution.

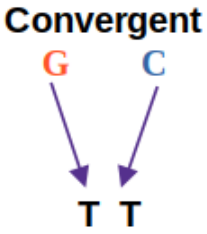


Figure 2. A schematic depiction of convergent evolution at the level of evolutionary genes shows that different non-homologous genetic factors produce similar phenotypes.

The swimming style of lumpsuckers and snail fishes prioritises mobility and the ability to swim freely. They evolved as demersal fish during the Pliocene period in the

North Sea (Ueno, 1970, Davenport and Kjørsvik, 1986). Both of these relatives seem to have evolved from demersal ancestors and achieved neutral buoyancy. The evolution of swim bladders in fish is poorly understood, and it is unknown how they became extinct in several teleost species belonging to the Cyclopteridae, Cottidae, and Liparidae families.

Another example is the diversity of attachment organs in aquatic animals that are present in five phyla of the metazoan phylogenetic tree. Single attachment organs can be seen in limpet (Mollusca, Gasteropoda) and benthic fish (Craniata, Teleostei) species (lumpsuckers, snailfishes, and sculpins). Similarly, multiple attachment organs are found in sea urchins (Echinodermata, Echinoidea) and octopuses (mollusca, cephalopoda). Due to evolutionary pressure, for their survival in the deep sea, the pelvic fins and girdle evolved into a PSO in lumpsuckers (Cyclopteridae) and snail fishes (Liparidae), which exerts stronger forces than those of other fishes and is evidence for convergent evolution in the *Liparis-Cyclopterus* group. Suction discs have additional applications. In male Cyclopterids, the suction organ serves to safeguard fertilised eggs (Davenport and Thorsteinsson, 1990). Other Cyclopteridae species, such as *Eumicrotremus spinosus* and the liparids, are likely to have homologous structures similar to those observed here (*Liparis gibbus* and *Liparis fabricii*). The attachment organs facilitate different biological functions: positional maintenance, locomotion, defence, and feeding (Delroisse et al., 2023). Single-suction circular attachment organs provide a strong mechanical attachment support that prevents leakage and lowers the internal pressure.

However, there is no solid evidence that we do not understand how PSOs originate in fish, since we do not know how they evolve or alter over time. Using advanced genomics techniques such as phylogenomics and phylotranscriptomics to study the evolution of Cottoidei could help identify the pathways and genes modified by natural selection and responsible for the phenotypic differences between these species.

1.4 Phylogenetic relationships

Phylogenetics is the study of evolutionary relationships among or within species, employing both morphological (Figure 3) and sequencing data. At the molecular level, only gene sequences can be used to determine phylogenetic relationships. Phylogenomic evidence suggests that two (2R) to three (3R) rounds of whole genome duplications (WGD) occurred around 350 million years ago in the stem lineage of ray-finned fishes (Actinopterygian) (Meyer and Van de Peer, 2005). The WGD has shaped the diversification of vertebrate genome architecture by producing many gene copies (e.g., Hox clusters). By constructing a phylogenetic tree using different types of gene sequences (orthologs and paralogs), their evolutionary history can be fundamentally depicted using these homologs (Fitch, 1970).

Orthologs are genes in different species that have evolved from an ancestral gene through the speciation process. Whereas paralogs are genes that are related via duplication events that happened in the last common ancestor. These were caused by mutations in duplicated sequences during speciation events (Fitch, 1970, Young and Gillung, 2019).

Genetic inference can be estimated at three levels: single gene, single locus (mitogenome), and multilocus (nuclear genome). The availability of free access to sequence data in biological databases and the accessibility of DNA sequencing methods provides researchers with a robust platform for examining the evolutionary events that occurred amongst numerous species.

Phylogenetic trees can be inferred from molecular sequence data using both orthologs and paralogs to assess the common ancestry. The paralog sequences can convey meaningful phylogenetic information and can sufficiently generate the full resolved phylogenetic trees (Hellmuth et al., 2015). Predicting orthologs is computationally expensive; blast-based searches have been developed and have previously been applied to biological databases by means of reverse blast searches

(Ward and Moreno-Hagelsieb, 2014). Many computational, graph-based, and phylogeny-based methods are currently available.

For comparative and evolutionary genomics research, phylogeny-based techniques demonstrate superior sensitivity and specificity in predicting orthologs from distantly related taxa (Gabaldon, 2008, Altenhoff and Dessimoz, 2009). Several computationally costly methods for detecting orthologous sequences in distantly related eukaryotic genomes have been developed thus far (Chen et al., 2007). Using genome and transcriptome data as input, orthologous and paralogous sequences can be predicted (using OrthoMCL (Li et al., 2003) and OrthoFinder (Emms and Kelly, 2019)). By applying evolutionary models, the relationships among taxa can be understood in the form of a species tree.

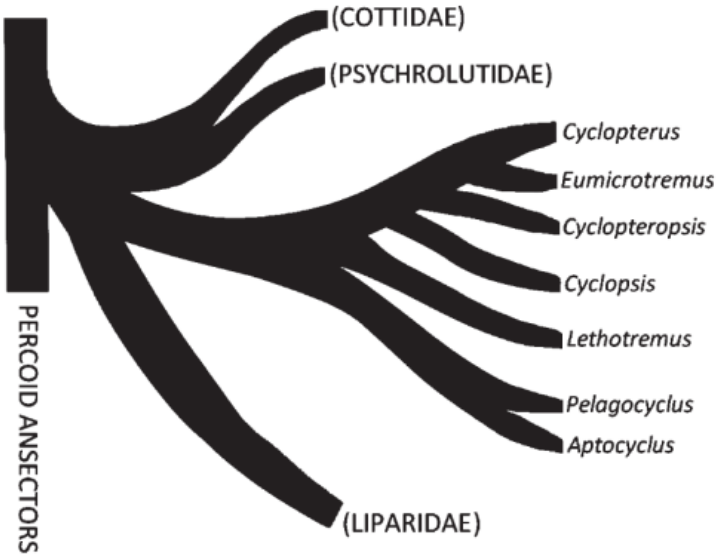


Figure 3. Phylogenetic relationships of the Cottoidei based on the synapomorphy (Ueno, 1970). A synapomorphy trait (e.g., PSO) that was present in an ancestor species and is still present (in more or less the same way) in all of its descendants. Synapomorphies, which imply homology, create evidence of historical relationships and their associated hierarchical structure. The majority of Cottoidei benthic fish have inherited ancestral features in the form of an attachment organ.

1.5 Mitochondrial genome

The mitochondrial genome of a fish is typically circular, measuring 16-20 kb in size. Typically, there are 37 genes in total (two rRNAs, 13 PCGs, and 22 tRNAs) (Boore, 1999). This genome is conserved across vertebrates and has relatively fast evolutionary rates and significant copy number variation, making it a powerful molecular tool in phylogenetics, phylogenomics, species identification, and stock evaluation (Moritz, 1994, Pereira, 2000, Ward et al., 2005, Atig et al., 2009, Galtier et al., 2009). Mitogenome research is especially useful because it is clonal (maternal), neutral, and has a constant rate of evolution (Galtier et al., 2009, Satoh et al., 2016). Nowadays, it is more common to use the complete mitochondrial genome rather than a single gene (COI) to ascertain phylogenetic relationships between fishes. Combined analysis (concatenated sequences) can enhance phylogenetic resolution and provide a greater degree of topological variation. Complete mitochondrial genomes are increasingly important in phylogeography and for determining evolutionary relationships between closely and distantly related taxa (Ward et al., 2005).

Due to the shared evolutionary history or genome history between the Liparidae and the Cyclopteridae families, lumpsuckers should be expected to share some lifestyle characteristics with deep-sea snailfish. Sculpins, sandfishes, snailfishes, and lumpsuckers represent monophyletic lineages within the Cottoidei suborder. A recent genetic investigation utilising mitogenome DNA (mtDNA) determined that snailfishes and lumpsuckers are related. They diverged 15.86 million years ago during the middle and late Miocene epochs (8.89–24.94 million years ago) (Shen et al., 2017, Maduna et al., 2022). According to divergence estimates, adaptive positive selection was detected in the *MT-ND5* gene (the enzyme complex functions as a proton pump), which is essential for energy production and responsible for deep sea snailfish adaptation to the hadal zone (Shen et al., 2017). Phylogenetic approaches using maximum likelihood and Bayesian inference both produced identical topologies for the reclassified tree. On the other hand, large-scale analyses of mitogenome selection between deep-sea fish and

freshwater sculpin lineages showed purifying selection in the fish suborder Cottoidei (Maduna et al., 2022).

These two groups of species (lumpsuckers and snailfishes) share many characteristics (life cycle, ecology, and evolution), and the majority of them appear morphologically similar; however, using the mitogenomes, they can be differentiated at the species level. Mitogenome data can easily distinguish interspecies and genus-level conflicts and provide a strong genetic inference from a species tree. However, mitochondrial DNA (mtDNA) is susceptible to complex mutations over time. It can undergo recombination, an inconsistent mutation rate, and heteroplasmy. These factors are taken into consideration when working with the mitochondrial DNA. Hence, mtDNA is not an accurate indicator of species diversity, and species trees based on evolutionary characteristics can be difficult to interpret accurately (Galtier et al. 2009). By comparing the complete genomes of taxa from both families and their close relatives (sculpins), we can deduce past selection pressures and evolutionary events. However, according to Galtier et al. (2009) and Balloux (2010), more information can be recovered from the nuclear genome than the mitogenome because its recombinant multilocus structure records more detail of a taxon´s history than the effectively single locus of the mitogenome.

1.6 The era of phylogenomics

Phylogenomics can be defined as the evaluation of gene function based on high-throughput sequencing (HTS) of genome data in relation to phylogenetic inference (Eisen, 1998). To infer phylogenetic relationships, numerous types of molecular homology (homolog, ortholog, paralog, and xenolog) are utilised (Eisen, 1998, Young and Gillung, 2019). In studies of evolutionary biology or systematics, reconstructing phylogenetic trees to infer genealogical links between distinct species is a significant undertaking. Rüber et al. (2004) utilised mitochondrial (*COI, COII*) and nuclear (*Tmo-4C4*) markers to examine interspecies relationships in fish. As a result of the low cost of

sequencing and the availability of genome and transcriptome data in public databases such as NCBI, genome and transcriptome data are now widely accessible. Using genomic data to construct a phylogeny and infer genetic relationships between genera, species, or operational taxa is a difficult task. Nonetheless, the only viable alternative is to infer the species tree on a larger scale. There are two approaches: 1) phylogenomics, which uses DNA-seq data, and 2) phylotranscriptomics, which uses RNA-seq data.

The sequencing of entire genomes of extremely diverse organisms, such as plants and fish, is a difficult task due to ploidy, often enormous size of the nuclear genome, and high sequencing costs. Phylotranscriptomics is an alternative method. It is cost-effective to sequence the transcriptome in the absence of genome information in order to identify orthologous genes. Cheon et al. (2020) state that phylotranscriptomics is a sensitive method for discovering orthologous genes because it examines only exon sequences.

Because phylogenomic data is computationally intensive to analyse, it is necessarily difficult to interpret. Additionally, low sequence read quality can result in poor genome assemblies and partitions with faint phylogenetic signals, causing a species conflict (Spillane et al., 2021). Taxon sampling and selection is the first step in phylogenomic inference; increasing taxon sampling by obtaining sequencing data from the UCSC genome browser, or adding a large repository of vertebrate sequence data can improve the phylogenetic accuracy of species trees (Heath et al., 2008, Young and Gillung, 2019). The two primary techniques for phylogenomic sequencing are target enrichment sequencing (UCE (Ultra Conserved Elements) or AHE (Anchor Hybrid Enrichment)) and shotgun sequencing (Figure 4). Target enrichment is a cost-effective sequencing strategy for non-model organisms that uses a set of primers or probes to sequence only regions of the genome with low or moderate sequence coverage. In contrast, shotgun sequencing randomly scans the entire fragmented genome (Young and Gillung, 2019).

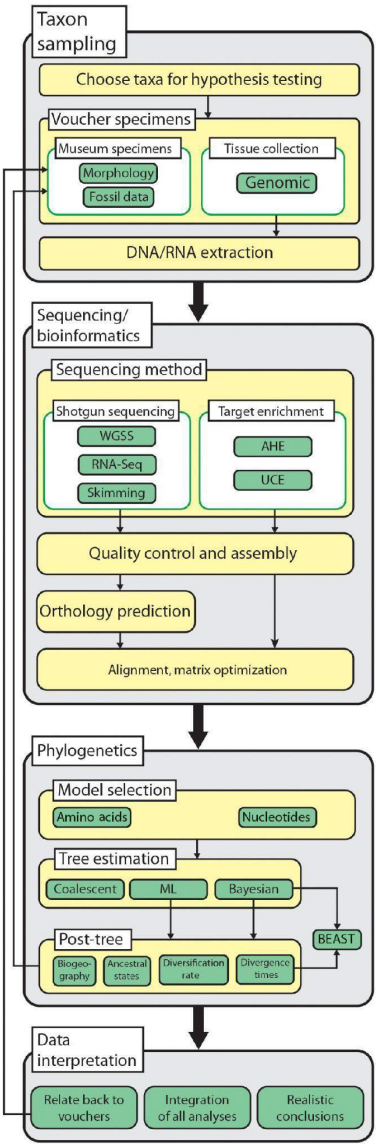


Figure 4. A typical phylogenomics workflow. In the current study the shotgun sequencing approach was adapted for sequencing whole genomes (Young and Gillung, 2019).

It is evident that each phylogenetic technique has its own set of benefits and drawbacks. The requirement for extremely high-quality RNA, which can only be obtained from fresh tissue samples, is one of the disadvantages of transcriptome sequencing (Young and Gillung, 2019). RNA sequencing from museum specimens stored in ethanol is not a feasible approach due to poor RNA quality and highly fragmented or degraded samples are only suitable for targeted approaches (RT-PCR and qPCR) (Worobey et al., 2016). As yet there are not a lot of RNA isolation methods that have produced promising results. The RNA in old museum specimens degrades considerably more quickly than fresh specimens. The most reliable alternative is a phylogenomics approach, which permits the extraction of DNA from historic specimens (Ruane and Austin, 2017). Techniques for obtaining high-quality DNA from very old specimens (aDNA) that have not been properly preserved have been refined. DNA extracted from museum specimens enables researchers to elucidate the history of a species (Burrell et al., 2015, Derkarabetian et al., 2019). Evolutionary biologists are interested in combining traditional genotyping approaches with restriction enzyme-based protocols, to generate genome-scale genotype data from hundreds of individuals at a time. Using SNP genotyping methods, such as RAD-seq and DArT-seq-based technologies, we can learn more about the demography and population structure of extinct and extant species (Jaccoud et al., 2001, Ruane and Austin, 2017, Raxworthy and Smith, 2021).

Both genomics-based methodologies will yield results that are comparable, and extensive comparative studies have revealed little topological distinction. However, regardless of the type of tissue used from a given species (Cheon et al., 2020), the difficulty of inferring phylogeny from transcriptome data sets is higher. Gene expression varies between tissues, and highly expressed transcripts evolve slowly (Zhang and Yang, 2015). Thus, phylogenomics is considered a more reliable approach than transcriptomics. Moreover, the final genome sequence data includes both coding (CDS) and non-coding (intron) sequences, offering both ortholog and paralog sequences for species tree construction, whereas phylotranscriptomics is more sensitive to

identification of orthologs. Most likely, this means that genetic differences and levels of selection among taxa can be extrapolated on a broader scale using a phylogenomics approach.

Due to various technological advances in sequencing, genomics researchers can rapidly and accurately sequence genomes of any size using long-read (PacBio) and short-read (Illumina) approaches (Figure 5). These two sequencing technologies are often coupled to provide high quality complete genomes. Such hybrid-based approaches are advantageous for genome assembly, allowing coverage of the entire genome, detecting structural variation as precisely as possible (Miller et al., 2017, Amarasinghe et al., 2020). In this project, PacBio Sequel II was used to sequence mitogenomes, while Illumina NextSeq 500 was used for low-coverage of nuclear genomes. Genome assembly of mitogenomes using Illumina alone is challenging because the structure of the vertebrate mitogenome contains a non-coding (D-loop) region that is difficult to assemble and recover from short reads. Additionally, identifying true duplicates is difficult (Satoh et al., 2016). Long-reads by Sequel II are automatically corrected for precision (99.9%) and have an advantage over short-reads by Illumina. However, PacBio sequencing of whole genomes is prohibitively expensive, a limitation meaning the majority of genome projects must rely on Illumina (Amarasinghe et al., 2020).

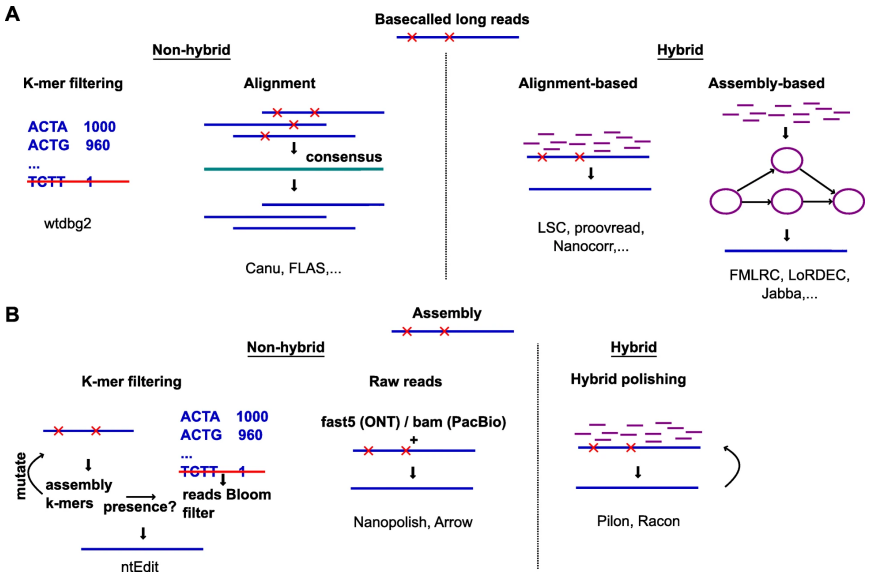


Figure 5. Paradigms of error correction (Amarasinghe et al., 2020), genome polishing, and assembly (long and short reads). Genome polishing is a procedure to fix errors in the draft genomes generated from short and long read sequence data. Additionally, improving the reliability of genomic analysis during genome assembly is nothing more than assembling the preprocessed short or long reads based on specific K-mers, into a consensus sequence by applying *de* novo assemblers designed for a specific read length (150 bp and 2 kb). Genome assemblies generated only from short read data are called "non-hybrid assemblies." If it is a combination of both (short and long reads), it is referred to as a "hybrid genome assembly." The hybrid method of filling gaps in the genome results in a nearly complete genome.

1.7 The importance of lumpsuckers in aquaculture

In aquaculture, cleaner fish are used to remove sea lice from farmed salmon. For instance, ballan wrasse (*Labris bergylta*) and goldsinny wrasse (*Ctenolabrus rupestris*) are frequently utilised as cleaners to eradicate sea lice (Deady et al., 1995, Treasurer, 2018), despite the fact that they are unsuitable for delousing below 6 °C (Sayer and Reader, 1996). Several wrasse species have been introduced into aquaculture; however, at low temperatures, lumpfish have been shown to be more effective and robust delousers (Imsland et al., 2018, Powell et al., 2018).

In reality, lumpsuckers use shelters to hide from predators in coastal areas with strong ocean currents, adhering to surfaces with their powerful suction disc. This strategy allows them to conserve energy as their spherical bodies prevent them from swimming efficiently (Imsland et al., 2014, Imsland et al., 2015, Hvas et al., 2018). In recent years, lumpfish have been targeted for their roe, an inexpensive substitute for sturgeon caviar. However, by-catch in the North Atlantic and Gulf of St. Lawrence has precipitously reduced their populations in recent years (Gauthier et al., 2017). Lumpfish bodies can be a variety of hues, including red, blue, and green. The blue-green colouring of lumpfish is caused by biliverdin, a product of heme catabolism. Additionally, lumpfish can alter their coloration based on the surroundings. However, few studies have determined the molecular pathways activating body colour adaptation to the surrounding environment in this species (Davenport and Thorsteinsson, 1989, Davenport and Bradshaw, 1995).

1.7.1 Lumpfish breeding and broodstock

Effective broodstock management points the way to long-term aquaculture sustainability. This has been demonstrated in salmon, tilapia, and trout. Broodstock management in cleaner fish was largely ignored until recently (Migaud et al., 2013). Better knowledge of cleaner fish and their welfare can facilitate production efforts.

Scientists in the aquaculture industry, particularly in Norway and other regions of Europe, such as Iceland, Greenland, the United Kingdom, and Canada, have lately begun breeding programmes for large-scale sustainable production of lumpfish. There are currently no published scientific reports available online about lumpfish breeding efforts, but in Norway, cleaner fish breeding programmes are at the initial stages.(https://thefishsite.com/articles/how-breeding-can-improve-lice-eating-efficacy-of-lumpfish-in-salmon-farms). Understanding the life cycle of lumpfish species in captivity will assist management of broodstock and the optimization of hatcheries in the near future (Pountney et al., 2020). Genotyping-based approaches paired with next-generation sequencing have yielded promising findings in the establishment of

sea lice resistance strains of salmon, growth-related features in rainbow trout (Vallejo et al., 2018), and Nile tilapia (Yoshida et al., 2019). Species more recently introduced into aquaculture, like lumpfish, are still in the early stages of domestication, which means that genetic improvements to growth-related traits remain to be made (Houston et al., 2020).

Large-scale sustainable broodstock production demands effective environmental management strategies (temperature, pH, and photoperiod) (Houston et al., 2020). The manipulation of gametes is the first stage in lumpfish stock management; according to Pountney et al. (2020), optimal rearing conditions for captive lumpfish broodstock should be below 10 °C. Temperatures above 14 °C were found to be detrimental to both milt and sperm (Pountney et al., 2020). Lumpfish milt has a longer motility survival duration than other marine teleost species. Using extender solutions, lumpfish milt can remain viable for up to two weeks. This is important for improving hatchery management and domesticating lumpfish in the near future (Pountney et al., 2020).

1.7.2 Challenges in lumpfish welfare

With respect to lumpfish well-being, Garcia de Leaniz et al. (2021) briefly highlight the information gaps and obstacles. Behaviour, mortality, and captive breeding are all important aspects of increasing the sustainability of fish welfare. Conducting productive R&D, creating collaborative animal research culture, and introducing more training programs are all possible solutions to these problems (e.g., the Delphi approach, (Garcia de Leaniz et al., 2021)). Because lumpfish and salmon have different habits and occupy different habitats they seldom come into contact in natural environments (Leclercq et al., 2018). Hence, special lumpfish husbandry procedures are needed to ensure lumpfish survive and remain healthy during the salmon rearing cycle.

In captivity, behavioural traits play a critical role in sustaining fish welfare. Lumpfish behaviour (Staven et al., 2019, Staven et al., 2021) and broodstock management have both been investigated (Pountney et al., 2020). Loss of appetite, lethargy, erratic

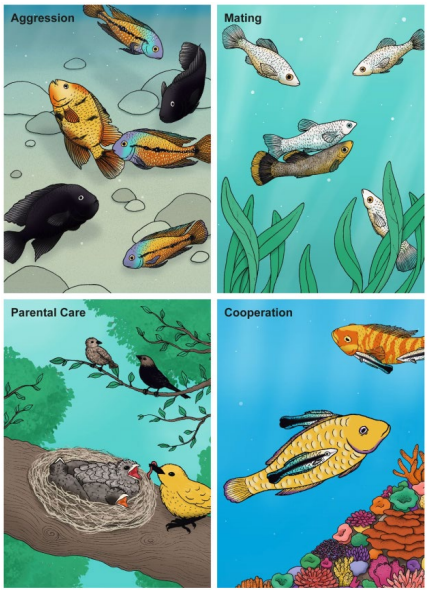
swimming, hostility, a change in skin colour, and gasping are some of the behavioural indicators reported in lumpfish management.

1.7.3 Lumpfish behaviour

Understanding behaviour is crucial to improving animal welfare in aquaculture. The genotype of an animal interacts with its environment to influence behaviour (Martins et al., 2012). Insufficient research has been conducted at the genetic level to fully comprehend lumpfish-salmon mutualistic cooperative behaviour. Through habituation-based studies, it was possible to comprehend the behavioural change that occurs when lumpfish interact with other species. When naive lumpfish were exposed to Atlantic salmon, their swimming activity increased, inter-species distance decreased, and plasma cortisol levels rose (Staven et al., 2021). Staven et al. (2019) observed the exact opposite to be true for properly habituated lumpfish.

Lumpfish exhibit strong sensory cues (swimming activity and skin colour) in the presence of salmon, with skin pigmentation and colour changes observed in stressed fish (Staven et al., 2021). In the aquaculture industry, the colour of lumpfish can be used to determine their well-being. In the majority of fish species, Mauthner cells (M-cells) initiate C-start behaviour, which is characterised by a strong flight response to threatening stimuli. The absence of M-cells in lumpfish (*Cyclopterus lumpus*) and the Pacific spiny lumpsucker (*Eumicrotremus orbis*), (Cyclopteridae), offers an exciting opportunity to learn more about the evolutionarily conserved neuronal circuits that trigger the fast startle response in lumpsuckers (Zottoli, 1978, Hale, 2000).

A Social heterospecific interactions



B Mutualism

Figure 6. (A) Social heterospecific interactions in different functional domains (Oliveira and Bshary, 2021). Heterospecific sociality can be found across various groups of animals. It can be aggression behaviour in cichlids, mating behaviour in Amazon mollies, parental care behaviour by yellow warblers, and cooperative behaviour by cleaner wrasse species. (B) In this study, the cooperative behaviour between the lumpfish and salmon was examined by performing interspecies experiments (https://www.aquaculturenorthamerica.com/).

2 Objectives

This thesis aims to address questions relevant to aquaculture and which are related to the evolution of deep-sea fish and the genetic basis of their behavioural traits as cleaner fish, using genomic approaches (DNA-Seq and RNA-Seq). The major goals of this dissertation are to contribute to the data-deficient lumpsucker mitochondrial genome database, and to utilise a phylogenomics approach to study deep-sea fish evolution. All species used in this analysis belong to the suborder Cottoidei, representing three different family groups (lumpsuckers, snailfishes, and sculpins). In the context of animal welfare, it was expected that a transcriptomics approach (RNA-Seq) might identify some sequences involved in the control of lumpfish behaviour.

Goals:

1. The complete mitochondrial genome of the Atlantic spiny lumpsucker, *Eumicrotremus spinosus* (Fabricius, 1776) **(Paper I)**
2. Understanding the evolutionary landscape of lumpsuckers, snailfishes, and sculpins of the suborder Cottoidei (Teleostei: Perciformes) **(Paper II)**
3. A search for hormone genes in cleaner fish (*Cyclopterus lumpus* L.) during habituation **(Paper III)**

The taxonomic status of deep-sea fish is not well understood, especially that of Arctic species which live in deep-sea benthic habitats which are difficult to access, requiring expeditions to the North Atlantic and Barents Sea. Taxonomic confusion exists among the Cyclopteridae, Liparidae, and Cottidae. In particular, the sexually dimorphic Cyclopteridae taxa have been mistaken for different species (Byrkjedal et al., 2007). However, recently efforts have been made to increase availability of mitogenomic resources (Maduna et al., 2022) for taxonomic and systematic studies of benthic fish.

Mitochondrial DNA is the first-choice marker to resolve phylogenetic relationships among closely related taxa, to infer genetic relatedness and intraspecific genetic

variation. To date, using only mtDNA markers, DNA barcoding has been used for phylogenetic inference in the suborder Cottoidei (Byrkjedal et al., 2007, Smith and Busby, 2014, Maduna et al., 2022). Availability of Cyclopteridae and Liparidae mitogenomes is limited, so as a contribution to the growing database this study contributed the complete mitogenome of the deep-sea Atlantic spiny lumpsucker *(Eumicrotremus spinosus),* using PacBio sequencing technology, detailed in Paper I. This mitogenomic data will be a useful resource for population genomics and systematic studies. Using the 17 mitogenome resources available, a revised phylogeny was constructed to determine a phylogenetic tree at the single locus level.

Further, this evolutionary exploration was continued in nuclear genes using 15 genomes of the Cottoidei. To understand the species tree at the multi-locus level requires identifying genes under selection and relating their potential function to the aquatic habitat. Using museum samples, we addressed these questions by applying a phylogenomics approach (Paper II).

For comparative genomics, the genome set comprised Cyclopteridae (4), Liparidae (6), and Cottidae (5) genomes. Six genomes of museum specimens (Cyclopteridae-3 and Liparidae-3) were sequenced with low-coverage (30–40X) using Next-Seq 500 for this study. Using available open source genomic data processing tools, an end-to-end pipeline was constructed where all six genomes were pre-processed, assembled, and annotated. A species tree was estimated based on single-copy orthologs. Different selection models (branch and branch-site) were tested to see if there was evidence of positive selection in genes likely to be involved in adaptation of benthic fish. These are elaborated on in Paper II.

Lumpfish are a fisheries target for inexpensive caviar and, in recent years, as cleaner fish. To understand the mutualism between both cleaner and client fish, habituation experiments were conducted to highlight the sensory cues in cleaner fish while they participated in inter-species interactions (Staven et al., 2019, Staven et al., 2021). To understand the complex social (mutualistic) behaviour and how it develops

with the client fish (salmon), lumpfish brain transcriptomes were sequenced using the NovaSeq 6000, in order to identify sequences involved in social and cognitive behavioural traits in cleaner fish after exposure to different stresses associated with their role. The transcriptomic resources generated in this study provide some reliable gene associations related to production of the stress hormone noradrenaline, which is discussed in Paper III. Understanding the genomic basis of these behavioural traits could enhance the welfare of cleaner fish.

3 Main findings

Paper I: The complete mitochondrial genome of the Atlantic spiny lumpsucker, *Eumicrotremus spinosus* (Fabricius, 1776)

To better understand the evolution and phylogenetic relationships of deep-sea fish species, we sequenced the complete mitogenome of the spiny lumpsucker (*E. spinosus*) using long-read (PacBio Sequel II) sequencing to avoid nuclear-derived mitochondrial sequences and identify any true duplications in the mitogenome. The mitogenome (19.2 kb, GenBank-OP728784) generated in this way has 13 PCGs, 22 tRNA genes, three non-coding regions, (Figure 7), 2 rRNA genes (Table 1) and one putative control region (Figure 8). The *E. spinosus* mitogenome has a similar sequence order to that of its group members (including *C. lumpus*). We used this resource, along with other freely available mitogenomes, to generate a phylogeny for close relatives of lumpsuckers that include snailfishes and sculpins, 17 species in total. There was strong evidence suggesting that Liparidae family members are the closest relatives of lumpsuckers (Paper I). The mitogenome annotation shows an extra non-coding area on the H-strand. The presence of a duplicated control region (CR2), which is rare in teleosts, located in between the two tRNA genes (tRNA-Thr-tRNA-Pro), was a new finding derived from this study. Duplicated control regions can evolve rapidly in a concerted fashion in animals but are rarely found in the teleostei. Future studies of this group need to use phylogenetic and ancestral state reconstruction analyses, to provide firm evidence of concerted evolution. This finding will assist researchers performing comparative genomics and taxonomic studies of these and related groups in the future.

E. spinosus has a typical vertebrate mitogenome structure (Figure 7) and was used to investigate evolutionary relationships at the single locus level by phylogenetic analysis of 16 additional mitogenomes of species from the suborder Cottoidei. The taxonomic placement of the lumpsuckers, snail fishes, and sculpins all derives from the

same common ancestor and are very closely related. All three family members—snail fishes, lumpsuckers, and sculpins—form individual groups.

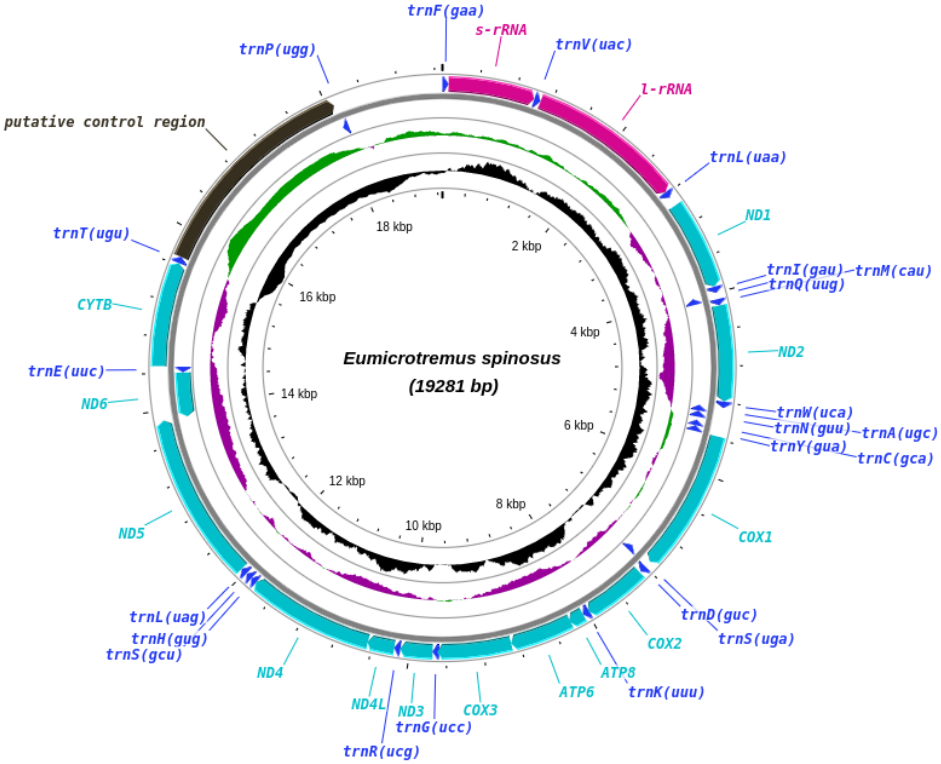


Figure 7. Circular plot of the spiny lumpsucker *(E. spinosus)* mitogenome displaying both strands H (+) and L (-) representing the typical vertebrate mitochondrial structure.

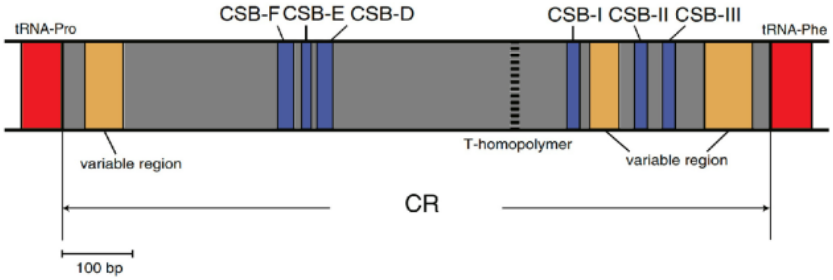


Figure 8. The control region (CR) in spiny lumpsucker (*E. spinosus*) resembles the similar type of control region in lumpfish (*C. lumpus*) (Maduna et al., 2022).

Table 1. Organisation and features of the *E. spinosus* mitogenome.

Gene	Strand	Position		Size		Codon		Anticodon
		Start	End	Nucleotide (bp)	Amino acid	Start	Stop	
trnF	H	1	67	67				ttc
rrnS (12 S rRNA)	H	68	1011	944				
trnV	H	1012	1083	72				gta
rrnL (16S rRNA)	H	1085	2774	1691				
trnL2	H	2775	2848	74				tta
3rd non coding (Homopolymer region)	H	2849	2967	119				
nad1	H	2968	3921	954	318	ATA	CAA	
trnI	H	3932	4001	80				atc
trnQ	L	4001	4071	70				caa
trnM	H	4071	4139	68				atg
nad2	H	4140	5177	1038	345	ATG	TTA	
trnW	H	5186	5256	79				tga
trnA	L	5258	5326	70				gca
trnN	L	5328	5400	74				aac
OL	L	5401	5437	37				
trnC	L	5438	5503	66				tgc

trnY	L	5504	5571	68				tac
cox1	H	5579	7108	1537	510	ATC	TTA	
trnS2	L	7124	7194	86				tca
trnD	H	7198	7270	76				gac
cox2	H	7285	7968	698	228	ATG	GAA	
trnK	H	7976	8049	81				aaa
atp8	H	8051	8212	163	54	ATG	TGA	
atp6	H	8209	8889	667	227	ATG	GTT	
cox3	H	8892	9674	785	261	ATG	TCC	
trnG	H	9676	9748	74				gga
nad3	H	9749	10096	348	116	ATG	GAA	
trnR	H	10098	10166	70				cga
nad4L	H	10167	10460	294	98	ATG	TGC	
nad4	H	10457	11830	1370	457	ATG	TGG	
trnH	H	11838	11905	75				cac
trnS1	H	11906	11973	68				agc
trnL1	H	11978	12050	77				cta
nad5	H	12072	13877	1827	602	ATA	CTA	
nad6	L	13889	14407	530	173	ATG	GTT	
trnE	L	14408	14475	68				gaa
Cob (Cyt b)	H	14481	15611	1136	377	ATG	TTA	
trnT	H	15622	15693	82				aca
CR region2 (duplicated)	H	15694	18077	2384				
trnP	L	18078	18147	70				cca
CR region 1	H	18148	19281	1134				

Paper II: Understanding the evolutionary landscape of lumpsuckers, snailfishes, and sculpins of the suborder Cottoidei (Teleostei: Perciformes)

To obtain a more thorough understanding of the evolution of benthic fish adaptations, six deep-sea (specimens) from museum samples were sequenced using the Next-Seq 500 (Illumina), to produce low-coverage (30–40X) genomes. The final dataset used for comparative genomics contains 15 complete genomes from the families Cottidae (5), Liparidae (6), and Cyclopteridae (4).

Using standard pipelines for genome analysis, we processed the data with good quality output (Phread = 20, 99%) and performed genome assemblies with different pipelines (SPAdes, Velvet, and AbySS). Only the SPAdes assemblies were chosen as the best (N50) for downstream analysis. Following genome annotation, the coding and protein sequences were predicted from the respective genomes, using the vertebrate protein sequences as a reference set. The 1880 loci were considered as single-copy orthologs from 15 genomes. Using these, we estimated the phylogenetic species trees using maximum likelihood and coalescent-based methods. The topology remained stable across analyses; no conflicting loci were detected. Furthermore, the values of concordance factors (gCF and sCF) strongly supported the species tree.

Using the simple-site models (M0, M8, and M8a), 1880 loci were tested, and a significant number of the genes (160) under positive selection were identified. These were involved in the protein folding, ribosome translation, and adipogenesis. Adipogenesis is one of the primary processes by which vertebrates store energy in adipose tissue**.** The majority of benthic fish employ their body fat and lipid layers to maintain neutral buoyancy in the deep sea. This phylogenomic study enables better comprehension of the genetic changes underlying phenotypic adaptations of aquatic Cottoidei species. This study is one of relatively few, as far as we are aware, to suggest some of the mechanisms of adaptation at the molecular level in benthic fish. Our research was consistent with earlier studies on deep-sea species.

Paper III: A search for hormone genes in cleaner fish (*Cyclopterus lumpus* L.) during habituation

Behavioural traits in cleaner fish were studied by conducting inter-species behavioural interactions experiments between lumpfish (cleaner) and salmon (client). Gene expression profiling was carried out on the forebrain regions of cleaner fish after inducing different stress responses (based on conditions in model, salmon, olfaction, and control groups). From the differential gene expression analysis, we were able to capture only very low gene expression signals. Most of the genes found to have significant expression (p-value < 0.05) turn out to be false positives, with very low fold change values (log FC > 0.5). As a result, only noradrenaline (NordAd) hormone levels have a positive relationship with gene count data (i.e., expression), as found in 1586 genes. This work resolved the significance of noradrenaline involvement in various functions. Basically stress, emotion, behaviour, learning and memory are maintained by noradrenaline, which in turn mobilises the different body functions (Tully and Bolshakov, 2010, Hussain et al., 2022). Our results were purely based on correlation analysis, which is considered to be less reliable. Direct evidence from brain expression analysis was lacking. We cannot be certain of the reasons why we were unable to find genes related to these behavioural traits. We expected gene expression to vary greatly among different areas of the brain, not just in the telencephalon region alone. To figure out the changes in the transcriptome, gene expression analysis needs to be done on both cleaner and client fish samples.

Improvements which could be considered in future studies: interspecies experiments should be conducted for longer durations (more than 2 h, every 30 min interval monitoring) to understand the brain transcriptome synchronisation process happening in both the brains of both participants (cleaner fish and client fish). This can be achieved by targeting the different portions of the brain tissues (endbrain, forebrain, hindbrain, and midbrain) to interpret the social decision-making signals that may exist in them. Other researchers who used similar approaches revealed the molecular

genetic basis of the aggressive behaviour in Siamese fighting fish (Vu et al., 2020) and zebrafish (Thornqvist et al., 2019). They also suggested that the complex social behaviour (mutualism) of cleaner fish is caused by dopaminergic and synaptic pathways.

4 General discussion

4.1 Life in a deep-sea environment

The deep sea is diverse and largely unexplored. It hosts a specialised fauna, mostly heterotrophs compared to that of shallow waters. The deep-sea megafauna experience high hydrostatic pressures, extremely cold water, and an absence of sunlight (Ramirez-Llodra et al., 2010). The deep-sea environment consists of five oceanic layers (epipelagic, mesopelagic, bathypelagic, abyssopelagic and hadal) (Figure 9). It is difficult to study the deep-sea, leading to a lack of knowledge about its habitats and communities. The scarce knowledge of deep-sea species that live in this challenging environment created a taxonomic impediment for classifying the intraspecific variation among the species, making it a problematic task to recognise species boundaries (Frutos et al., 2022). Ecosystem monitoring and conservation are complex due to the remoteness of these habitats, which makes them a challenge to study. Technological advances (deep-sea landers) have been developed to enable more systematic study of the deep-sea habitat, and may be useful in implementing conservation measures in the near future (Aguzzi et al., 2022).

The deep-sea environment is dark and exhibits high pressures at greater depths (>200 m) (pressure increases by one atmosphere every 10.06 metres) (Jamieson, 2015). There is no sunlight for photosynthesis and very limited food resources, in contrast to shallow waters, where there is an abundance of food and it is warmer. Faster rates of speciation have happened in the shallow water lineages, and adaptive radiation has led to this environment becoming species-rich (Miller et al., 2022). Colonisation by deep-sea fish happened during the colder periods of earth history, which provided a favourable circumstance for settlement in the deeper ocean.

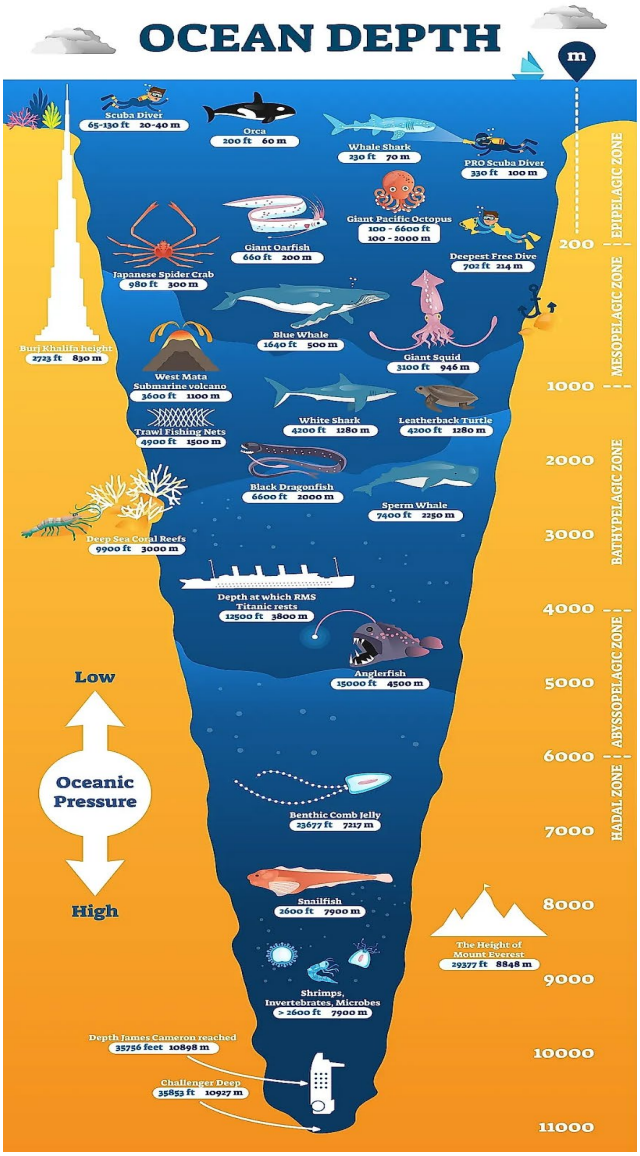


Figure 9. Different layers of the ocean and deep-sea creatures that inhabited different zones: shallow water (epipelagic zone - the sunlight zone), mid water (mesopelagic - the twilight zone, and bathypelagic - the midnight zone), benthic (abyssopelagic zone), and hadal (hadalpelagic zone). (The figure was adopted from shutterstock.com.)

According to Priede and Froese (2013), the evolutionary origin of deep-sea fishes occurred around 200 million years ago. The last time this happened on a global scale was during the Cretaceous geological period (145–66 million years ago (Mya)), when the Atlantic Ocean opened. Eventually, the Actinopterygii (teleosts) began to diversify, with many fish families adapted to different depths (200–9000 metres). Species migration into the deep-sea was constantly balanced throughout the earth's history. Understanding the history of Atlantic deep-sea colonisation could provide useful context for the evolutionary origin and speciation events of deep-sea fish.

4.2 Deep-sea fish phylogenetics

The genera belong to the lumpsucker, snailfish, and sculpin groups distributed throughout the North Atlantic and North Pacific oceans. Skeletal morphology and body density in deep-sea and shallow-water fishes can be easily differentiated (Gerringer et al., 2021). Especially in deep-sea fish like snailfish, lumpsuckers, and sculpins, they have poorly mineralized bones and skeletal reductions. In both lumpsucker and snailfish groups, reductions in ventral suction discs can be observed based on the natural habitats they live in (Budney and Hall, 2010, Gerringer et al., 2021). Sculpin and snail fishes are a very diverse group; some of the sculpin species are demersal, some are shallow marine, and some are freshwater. Morphologically, all three groups differ markedly. Lumpsuckers have a globular body shape, while snail fishes have an elongated body and less skin pigmentation; whereas sculpins also have an elongated body but have a modified webbed pectoral fin that they use for gripping seafloor substrate (Gosline, 1994, Kane and Higham, 2012).

Relationships between the three benthic fish groups were classified based on synapomorphic characteristics that showed that lumpsuckers, snailfishes, and sculpins have been classified as three different families that are closely related and belong to the single Cottoidei suborder. However, their divergence from common ancestors was not known. Skeletal reductions in deep-sea habitat fish are very common; they are

driven by hydrostatic pressures and the influence of other environmental factors (Gerringer et al., 2021), as well as the strong selection pressure on deep-sea fish lineages (Wang et al., 2019, Mu et al., 2021). The need for taxonomic studies at the morphological and genetic levels in deep-sea species will permit a better understanding of the evolution and diversification of different fish lineages.

Taxonomic confusion exists because of species diversity (Voskoboinikova et al., 2020). There is very little data available at the taxonomic and genome levels (Maduna et al., 2022). Efforts have been made to understand their taxonomy using morphological characters (modified pectoral fishes, tubercules, and suction discs); hence, it is reported that *Eumicrotremus* is a young group of weakly differentiated species (Knudsen et al., 2007, Voskoboinikova et al., 2020) and the existence of holotypes and extreme sexual dimorphism in this genus (Byrkjedal et al., 2007, Hatano et al., 2015). To address this issue, we have sequenced the complete mitogenome of the Atlantic spiny lumpsucker, *E. spinosus, a* deep-sea Arctic fish. In this genus, the genome resources are very limited; so far, only the *E. asperrimus* mitogenome has been reported. The revised phylogeny and their evolutionary relationships with their close relatives, the lumpsuckers and sculpins, were shown in Paper I. *Eumicrotremus* genus is a data-deficient group; getting samples of this group is challenging; they live in very hostile and extreme cold temperatures at the Ny-Ålesund settlement in Svalbard (Berge and Nahrgang, 2013). Moreover, we have sequenced its draft genome using the Illumina NextSeq 500 and discussed the relationships at the genome level in Paper II. For comparative genome-based studies in Paper II, we have analysed genomes belonging to members of the deep-sea lineages.

4.3 Ecology of lumpfish

Solitary species like lumpfish prefer to live in much colder environments in the North-Atlantic region (4°C) and in the Barents Sea (Pampoulie et al., 2014). During breeding seasons, they form spawning aggregates and prefer to live in shallow, warm

waters and participate in annual long-distance migrations (Schopka, 1974, Kennedy et al., 2015). Females carry the eggs and deposit a batch in the nest; males use the suction disc to guard them until they hatch (Goulet et al., 1986). While during the warmer temperatures the majority of them prefer to live in 4–7 °C. The number of lumpfish in the sea depends on oceanographic conditions. Lumpfish abundance totally depends on the temperatures. Their feeding preference varies depending on the environment they live in (sea lice or crustacean species) (Imsland et al., 2015). Recent reports suggest that the increase in temperatures has affected the lumpfish biomass. Monitoring the fish stocks in these regions will improve our knowledge of the population structure and management of these cleaner fish populations, which is very useful for aquaculture purposes (Eriksen et al., 2014). Moreover, tagging experiments could assist in understanding fish behaviour in aquaculture.

4.4 Ecological and environmental drivers in deep-sea fishes

The key ecological and environmental drivers in deep-sea fish vary greatly in comparison to those experienced by shallow water teleost species. Deep-sea fish live in conditions of low temperatures, low ambient light, and food scarcity. After evaluating different ecological and environmental drivers, depth was identified as a significant factor for growth rate (late maturation and slower life histories) in deep-sea fishes (Black et al., 2021). Deep-sea teleosts generally mature later, have greater longevity, slower growth, and lower metabolic rates. Life history parameters vary greatly across the depth continuum (Drazen and Haedrich, 2012). Similarly, organic carbon is the primary basis of the food web in the ocean, and their abundance also decreases with the depth factor (Lutz et al., 2007). The scarcity of food in the benthic zones is obvious, causing deep-sea fishes to develop a slow metabolic rate.

Snail fishes at depths of 400–9000 metres (Figure 9) vary considerably in terms of body structure, metabolic rates, and skin pigmentation relative to their closer relatives, lumpsuckers and sculpins. Depth, temperature, and food supply strongly influenced

their evolution, and is reflected in their phylogeny (Black et al., 2021). This was noted in the phylogenomics study (Papers I and II). Hadal snailfish (*Pseudoliparis swirei* and *Pseudoliparis amblystomopsis*) and benthic snailfish *(L. gibbus, L. liparis, and L. fabricii*) may differ in this regard. Similarly, in the case of sculpins, this would apply to the demersal group (*M. scorpius, 0–451 metres; T. bubalis, 0–300 metres; and C. analis, 0–1800 metres*) and the shallow-water fishes (*C. gobio* and *C. rhenanus*) as well. But in the case of lumpsuckers, more or less, all of them are expected to have similar life histories; coming to shores and shallow waters for breeding. Most of their time is spent in the benthic zones. North Atlantic (*E. derjugini, 1038 metres; E. spinosus, 1000 metres;* and *C. lumpus, 868 metres*) and the North Pacific species (*A. ventricosus, 612 to 1700 metres*) occupy different depths. Most benthic fish acquire similar adaptations: a slow metabolic rate, development of a pelvic suction disc or an attachment organ, a high content of fat deposits as an energy source, which in turn helps in neutral buoyancy, and no swim bladder.

4.5 Implications for taxonomy and phylogenomics

We have generated genome datasets for data-deficient deep-sea Arctic fishes. The complete mitochondrial genome of *E. spinosus* was generated using long-read sequencing, which resembles the similar mitochondrial structure of vertebrates. However, the presence of a duplicated control region was identified; maybe due to the concerted evolutionary processes, which must be evaluated in future studies by ancestral construction analysis (Li et al., 2015). Previous reports on the *C. lumpus* mitogenome (Maduna et al., 2022) reported that there was a third non-coding sequence on the H-strand present in three species of lumpsuckers (*E. spinosus, C. lumpus, and A. ventricosus)* and that it was found in a few species of snailfish (*L. ochotensis*) as well. But this region is absent in the hadal snail fishes; this has not been reported yet in other benthic snail fishes.

The presence of a third non-coding region is rare in teleosts, as reported earlier by Maduna et al. (2022). Although the reason for this remains unknown, this non-coding region could be a useful marker for species delineation purposes, since lumpsuckers have the GC-tract homopolymer region. Similarly, the liparis have a G-tract homopolymer, whereas in hadal snailfish, this region appears to be absent. This GC tract in lumpsuckers is an ancestral inversion event. All three species belong to the Cyclopteridae family, and phylogenetic analysis revealed that the three species were closely related. All of them represent the same mitogenome structure with 13 PCGS, 2 rRNA, and 22 tRNA genes. The putative control region is well annotated; containing two variable regions, conserved sequence blocks (CSB-F,CSB-E,CSB-D,CSB-I, CSB-II, and CSB-III), and a T-homopolymer region (Figure 8). Only in the Atlantic spiny lumpsucker, *E. spinosus, was* a duplicated control region identified using the long reads that would not have been possible using short read data (Paper I). Highly variable regions in the CR region do not permit sequencing using Illumina platforms, requiring instead long read approaches (Nanopore and Pacbio). Nowadays, a lot of sequencing projects, especially for studying the organelle genomes, are adapting long-read-based sequencing technologies.

The *E. spinosus* mitogenome was used to investigate evolutionary relationships at the single locus level by phylogenetic analysis of 72 additional mitogenomes of species from the suborder Cottoidei. We generated a revised phylogeny (Figure 10) depicting the relationships between the cottales, hexagrammelers, gasterosteales, zoarcales, and anoplopomatales. More branch-site and site models were tested on the protein-coding genes of Cottoidei lineages from benthic (deep-sea) and pelagic (fresh water) fish groups. The analysis showed that all genes were under neutral (purifying) selection (Pinninti et al., *unpublished*).

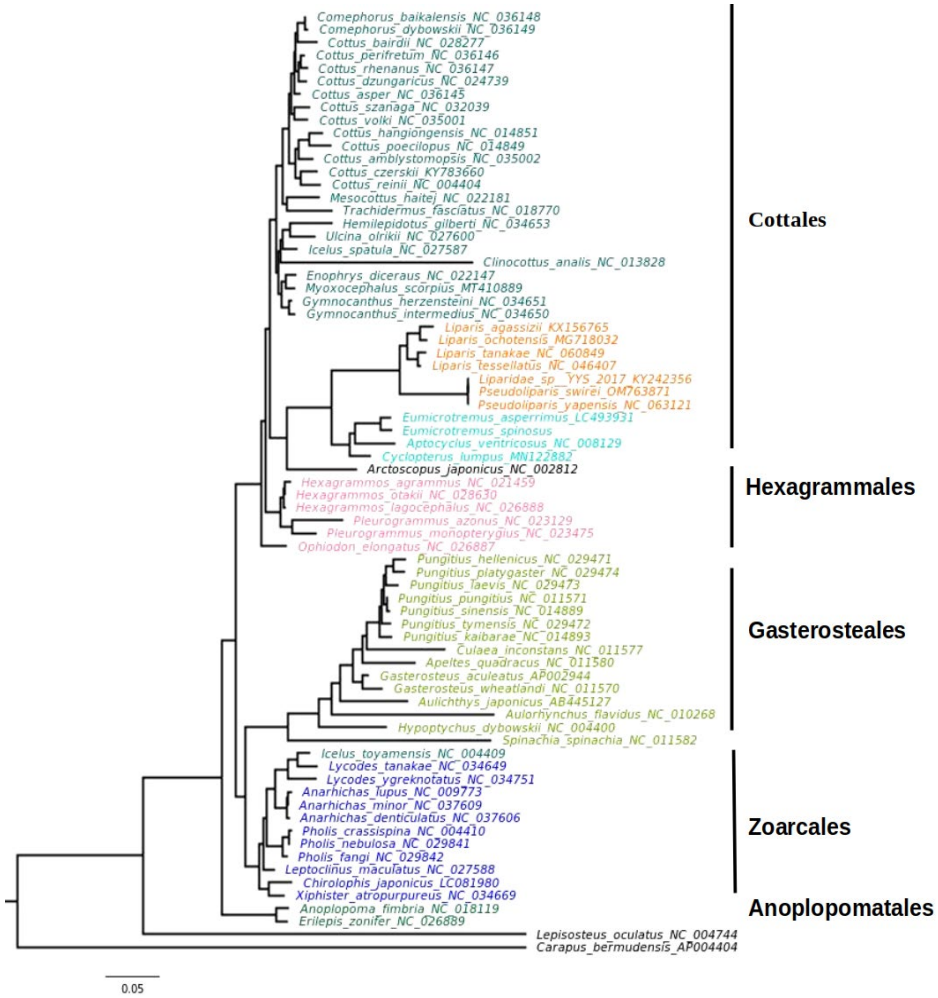


Figure 10. A revised maximum likelihood phylogenetic tree with 100 % bootstrap support was generated using the 13 protein coding genes (PCGs) of 73 mitochondrial genomes belonging to the Cottoidei suborder. Only the *Eumicrotremus spinosus* mitogenome generated from this study was closely clustered with the other Cyclopteridae family members (*E. asperrimus, A. ventricosus,* and *C. lumpus*) labelled in blue.

For this data-deficient group, we have generated six draft nuclear genomes for these Arctic fishes from museum specimens (lumpsuckers: 3 and snail fishes: 3), which allowed better understanding of evolutionary relationships between the 15 taxa at the multilocus level and showed a similar phylogeny to that from the mitogenomes (Paper I). Previous reports on DNA-barcoding using museum specimens have also shown the

similar phylogeny for lumpsuckers, snail fishes (Byrkjedal et al., 2007), and sculpins (Knope, 2013). From our analysis, there were no topological differences noticed in the robust ML tree, and the coalescent-based ML tree was clearly able to show the phylogenetic placement of the species into three different individual groups. In Liparidae, there is clearly a separation of the hadal snail fish and benthic snail fish, forming individual clades. Similarly, in Cyclopteridae, species belonging to the *Eumicrotremus* genus represent a single clade, and *A. ventricosus* and *C. lumpus* represent a single individual clade. All of them represent a single group. Moreover, marine sculpins form one clade, and freshwater sculpins another. All of them represent the Cottidae family. Our results and the genome datasets we generated will be useful for future systematic studies and large-scale comparative genomics projects. In this study, we have shown the evolutionary relationships among the three family members at the genome scale for the first time (Paper II).

In the process of natural selection, natural forces determine an organism's characteristics. A variation, or allele, of a trait makes certain individuals better able to thrive in their environment. It is generally referred to as genetic selection. In our analysis using different selection models, we have identified some key genes involved in deep-sea fish adaptation to the benthic zone. Such genes' roles were described in Paper II. We have shown the genes involved in adipogenesis, protein folding, translation, and DNA repair kit genes identified from our analysis were under positive selection. It is anticipated that pressure-adapted proteins will exhibit greater stability to counterbalance the consequences of denaturation. Such genes as chaperonin-containing TCP-1 subunit 5 (CCT5) are essential for protein folding in deep-sea aquatic species (Weber et al., 2020). Selection can favour the organism in adaptation, that is sure. It can be a positive or relaxed selection. In the case of annual killifish species, they revealed a pervasive relaxation of purifying selection, which preferentially targets ageing-related pathways predicted by the mutation accumulation theory of ageing (Gavrilov and Gavrilova, 2002). Relaxed selection alone contributed to the majority of the deleterious mutations affecting lifespan across the environmental gradient (Cui et

al., 2019). In contrast, longevity of deep-sea fishes may be attributable to a robust DNA-repair mechanism.

Some benthic fish lack a swim bladder; to maintain neutral buoyancy, they utilise the high content of stored fats or lipid layers in the body (Davenport and Kjørsvik, 1986). We reported positive selection genes involved in adipogenesis for the first time in benthic fish (Paper II). The putative genes were involved in adipocyte differentiation, proliferation, and inhibition functions. Moreover, due to evolutionary pressure, they have developed a specialised organ on their ventral side, the pelvic suction disc, which is also referred to as an attachment organ. Most deep-sea fish use the suction disc for attachment to the adjacent surface or the sea bed to save energy instead of swimming. Similarly, cartilaginous fish (sharks and rays) also lack a swim bladder, instead using fats and oils to maintain buoyancy (Davenport and Kjørsvik, 1986). Dynamic lift is supplied by their large pectoral fins and heterocercal tails (Thomson and Simanek, 1977).

Neutral processes drive the evolution of duplicate genes in eukaryotic genomes, their appearance and absence in the genomes through genetic drift (Brunet and Doolittle, 2018). Some duplications are fixed under certain conditions, often allowing the species to adapt to new environmental conditions (Kondrashov, 2012). Recently revealed by Zhang et al. (2021), HSDs have a potential role in the Antarctic green algae *Chlamydomonas* sp. UWO241 survival in harsh environments. Researchers have been studying the divergent duplications in the teleosts for a long time (Volff, 2005), and in Paper II, we have shown the HSDs role in deep-sea fishes. From the 15 genomes, we have predicted the highly similar duplicates (HSDs) based on percentage of sequence identity of amino acid sequences and with various length cut offs. The annotation indicated a lot of the HSDs were associated with carbohydrate and lipid metabolism. Some of these may be artefacts; in future studies, the HSDs possible functions need to be explained in deep-sea fishes by performing the selection analysis and functional studies.

4.6 Lumpfish in aquaculture

Lumpfish are in heavy demand in salmon aquaculture for delousing fish in pens (Powell et al., 2018), reducing reliance on chemical treatments. Studies have shown that it can lower sea lice infestations very effectively (Imsland et al., 2018). At present, the farming industry is heavily dependent on wild broodstocks. Annually in Norway alone, around 50 million cleaner fish are used for experimental and industrial purposes (Barrett et al., 2020). Currently, a lot of research programmes on the selective breeding of cleaner fish to improve sea-lice efficacy are in their initial stages in Norway, conducted by GIFAS and Akvaplan-Niva (https://thefishsite.com/articles/).

The welfare of domesticated fish species may be improved by appreciating their behavioural traits. With regard to cleaner fish, how they behave in client-cleaner interactions, their performance in captivity, and what welfare needs they demand. This is especially true for cleaner fish that have a high commercial value. Habituation studies in cleaner fish with salmon interactions were conducted by Staven et al. (2019) and Staven et al. (2021) to understand the behavioural traits at the physiological level by monitoring their locomotion before and after stress treatment in juveniles with mature salmon. It was revealed that lumpfish have high sensory cues (swimming and dark skin colour) towards the Atlantic salmon in the odour and salmon treatment groups. These studies may help the aquaculture industry devise better welfare practices.

4.7 Social heterospecific interactions (cooperation)

Basically, interspecies interactions, between two similar or dissimilar species in nature, can benefit from each other. It can be a process of aggression, parental care, mating, or cooperation (mutualism) (Oliveira and Bshary, 2021)**.** Here, we have studied the mutualism between lumpfish and salmon. We conducted habituation-based experiments to understand the behavioural traits at the genetic level (Staven et al., 2021). The traits that help in establishing the social complex behaviour between two interacting species have been thoroughly investigated by conducting inter-species

experiments for a short-term duration (1 h 30 min). We were unable to find genes which were significantly differentially expressed in the brain transcriptome analysis of multiple treatment groups (model, odour, salmon, and control) due to experimental, sampling, and sequencing difficulties.

This negative finding could be attributed to many reasons, such as the fact that while the telencephalon can be an important brain region involved in neurotransmitter synthesis, a cerebrum hemisphere is involved in central nervous system regulation. Gene expression varies widely among different brain regions (i.e., the hindbrain, midbrain, and forebrain) and may have impeded identification of the gene expression related to learning and memory formation in cleaner fish. A major drawback of the study is that gene expression analysis is based on cleaner fish samples only (Paper III). Additionally, gene expression profiling of different parts of the brain tissue could really improve studies attempting to identify the behaviour-specific cognitive and social traits and their links with the transcriptome synchronisation process and genes responsible for learning and memory in great detail. Our findings imply that genes linked to noradrenaline hormone levels may contribute to the behaviour of cleaner fish. The analysis suggests important genes related to neurotransmitter synthesis, cell signalling, pigmentation, and stress-related genes involved in locomotion. Hence, we concluded that noradrenaline has a major role as a neurotransmitter, in pigmentation (Aspengren et al., 2003, Nilsson Sköld et al., 2013), brain function, and body for action (fight-or-flight response) (Hussain et al., 2022).

Recent fish behaviour studies have shown that noradrenaline, calcium (Vu et al., 2020), dopaminergic pathways, and the synapse pathways (Ramirez-Calero et al., 2022) play a major role in aggression and cooperation in different teleost species, such as zebrafish (Ncube et al., 2022), and Siamese fighting fish (Vu et al., 2020). Ramrez-Calero et al. (2022) performed an experiment similar to ours. Their experimental setup involved species interactions (40 min trail) between the bluestreak cleaner wrasse *(L. dimidiatus)* and surgeonfish *(A. leucosternon)* with a smaller sample size (n = 6) and

only two treatment groups: control (no interaction, n = 6) and the treatment group (interaction, n = 6). They performed the gene expression analysis on different brain regions, so that they could detect how learning and memory have a direct influence on behavioural components.

Future studies should focus on conducting long-term behavioural trial experiments that would improve understanding and identify the behavioural traits involved in social cognition, which would provide insights into cooperative behaviour in interacting partners. In particular, such successful experiments would greatly benefit cleaner fish aquaculture in developing animal welfare programmes. Based on our experiences, we would suggest using single-cell RNA-seq for behaviour studies (Bentzur et al., 2022), which can reveal transcriptional shifts across cleaner fish treatment groups and possibly suggest behaviour traits related to learning and memory in both the cleaner and client transcriptomic synchronisation processes. This advanced technology can enable the finding of stress-related markers with precision.

4.8 Implications of climate change and habitat disturbance

At present, the salmon farming industry is heavily dependent on wild lumpfish broodstocks. Climate change has a necessary effect on deep-sea species, but so far no such reports on lumpfish populations are apparent in the literature. There are some studies explaining the population structure of lumpfish in the North Atlantic (such as (Pampoulie et al., 2014). All the studies were based only on microsatellite markers. To-date there are no reports using advanced genotyping approaches. Studies based on EST-STRs analysis revealed a lack of population differentiation along the Norwegian coast (Jónsdóttir et al., 2022). Similarly, a study conducted by Whittaker et al. (2018) revealed that some populations are small and have low genetic diversity. Future studies should be conducted using the genome-based approaches (RADSeq and DArTSeq), along the Norwegian coast where most lumpfish farms are located (Barrett et al., 2020). That would indicate the potential impacts of introgression of commercial stocks into

natural populations on a broader scale. It could provide diagnostic markers for differentiating wild and farmed cleaner fish on a broader scale, eventually helping in conservation management of this species.

Similarly, introgression studies conducted on salmonids to understand the introgression of farmed and wild salmon populations by Bolstad et al. (2021), revealed the ecological and evolutionary consequences of introgression on life histories and growth-related markers (*vgll3* and *six6),* useful in aquaculture production. In addition, extensive studies on introgression have revealed that the offspring of farmed salmon have a lower lifetime fitness in the wild than wild salmon, resulting in a decrease in the production of genetically wild salmon. Long-term consequences of introgression lead to changes in life-history traits, reduced population productivity, and decreased durability (Glover et al., 2017). Diagnostic SNP panels were designed that can discriminate between farmed and wild Atlantic salmon (Karlsson et al., 2011, Karlsson et al., 2016). This marker approach has become a more useful way of quantifying gene flow for conservation purposes.

4.9 Limitations of the study

Sampling deep-sea fish was a challenging task due to the environment they live in. Deep-sea taxa are extremely diverse, and most researchers lack a thorough understanding of the various niches that exist in the deep-sea. Deep-sea life remains largely unexplored, although there are an increasing number of expeditions to understand the marine life of various habitats.

As a consequence of these practical difficulties we relied on historic museum specimens (15- to 2-year-old samples). The other major issue was that the DNA quality was poor, and was likely to degrade very quickly from historic ethanol preserved samples. We managed to sequence the degraded DNA for some samples to obtain good-quality (Phred = 20; base call accuracy 99%) genome data. We recovered sufficient sequence data to understand the evolutionary relationships in the Cottoidei

lineages, which represent three major bottom-dwelling fish groups (lumpsuckers, snail fishes, and sculpins) (Paper I and Paper II), despite the fact that some of the species' genome assemblies (*Eumicrotremus spinosus*) were of poor quality. We managed to identify important loci that regulate some key functions in deep-sea benthic fish. Future phylogenomics research should combine the most recent sequencing technologies — Pacbio Sequel II and Nanopore—with Illumina to allow the majority of the genomes to be covered.

Understanding/Inferring complex social behaviour between two interacting partners (cleaner and client fish) at the genetic level is challenging because two interacting partners behave differently. After a series of species interaction experiments, we were able to measure the individual fish hormone response before and after stress induction. Physiological observations have produced a clear understanding of cleaner fish sensory cues (skin colour and swimming activity) in relation to their client fish (salmon). However, it was not possible to identify the underlying genetic traits involved in social and cognitive behaviours that promote learning and memory. Improvements to the protocol might involve gene expression profiling performed for specific and identifiable candidate brain regions (endbrain, forebrain, hindbrain, and midbrain) thought to be associated with the behaviours of interest in both interacting species. Our study focused on only one specific region (the forebrain) of the brain in only the cleaner fish (Paper III). The results based on correlation analysis of gene expression data with noradrenaline hormone levels; significant DEGs (logFC > 0.5) related to the behaviour traits (neurotransmitters or dopamine) were not found from the gene expression data. However, transcriptome-based methods are ideal for investigating gene expression, particularly in behavioural genomics, where more advanced genomic approaches (scRNA-seq) can be employed in future work. That would provide a more comprehensive interpretation of behavioural transcripts responsible for stress responses in cleaner fish.

5 Conclusions

The current thesis adds to our understanding of deep-sea fish evolution and cleanerfish cooperation behaviour at the genetic level. In the current PhD project, our main focus was to contribute to the genome resources for the deep-sea fish to explore the evolutionary relationships and the adaptations they have acquired. Analysis of 73 mitogenomes of suborder Cottoidei, including deep-sea and fresh water lineages, suggested sequences that were subject to purifying selection. A revised phylogeny at the single locus level provided some insights to the relationship of this group (Paper I).

In addition, a contribution was made to deep-sea fish genomics of the Cottoidei suborder by providing low-coverage genomes of six lumpsucker species (Cyclopteridae-3, Liparidae-3). Using available closely genomes of closely related species, we provided genetic inference at the genomic level using single-copy orthologs. A species tree was inferred for the first time in the Cottoidei suborder, a data-deficient genome resource. Some genes were shown to be under positive selection, such as those involved in adipogenesis, protein folding, DNA repair, and ribosome translation. Such adaptive changes are likely important mechanisms in deep-sea fish, maintaining neutral buoyancy using lipids and carbohydrates. This possibility was supported by genes thought to be involved in lipid and carbohydrate metabolism also appearing to be under selection. Freshwater species may show similar adaptations and would benefit from investigation. Genes involved in DNA repair mechanisms also show signatures of positive selective, perhaps ensuring these species adapt to a slower lifestyle requiring greater longevity under the challenging physical conditions of the deep-sea (Paper II). However, further experimental validation is needed to establish their functionality. Here we provide genome exploration of several species using low quality genome data. Future genomic-based studies should be on a sufficiently large scale to address the questions raised here and to provide markers which may contribute to conservation of the deep-sea benthic ecosystem.

Cleaner fish behavioural traits related to social and cognitive processes could be used as stress markers to determine the best conditions for lumpfish. Understanding their stress behaviour while in captivity could allow better framing of cleaner fish welfare. We considered our experimental strategy and suggested improvements for future studies (Paper III).

This thesis has explored the use of genomic-based methods to address knowledge gaps in the evolution of deep-sea fish. An attempt was made to identify markers for social and cognitive behaviours in lumpsuckers, potentially useful in assessing their welfare during rearing and when used as cleaner fish in the aquaculture industry. In the process, contributions have been made of genomic and transcriptomic data useful for future systematic, phylogenomic, and behavioural studies.

6 Future perspectives

The description of the deep-sea fish phylogeny is both a valuable resource and a challenging task. Future research should focus on applying comparative genomics approaches to other deep-sea species, utilising very high-quality whole genome datasets and various phylogenomics approaches capable of explaining selection pressures in great detail. Genes suggested to show positive selection pressure must be validated using experimental methodologies that would allow their functionality to be assessed, determining whether or not the amino acid modifications in the protein sequences are both factual and functional.

In the future, large-scale population genetics studies employing advanced genotypic methods (RAD-Seq and DArTSeq™), will be required to identify genetic markers that can be used for association mapping. QTL mapping can identify commercially important markers in aquaculture species**.** Population genetic studies are necessary to determine gene flow, genetic diversity, population structure, and resilience to climate change. It is hoped the findings from this work will assist marine ecologists in understanding aquatic species populations (stocks) and their adaptations to changing climates, allowing scientists to devise preventative measures that benefit marine conservation.

Finding transcriptomic and perhaps ultimately genotype markers, for cleaner fish behaviour would allow researchers to understand the basis of interactions between two individuals, and the changes during synchronization of brain transcriptomes. For behavioural studies involving more advanced molecular approaches (single-cell RNA-seq), the design and execution of the experimentation are vital to enable understanding of the molecular pathways involved in triggering different behaviours (aggression, swimming, and foraging). Future studies should consider the limitations of this study (Paper III), improving the experimental procedure and molecular approach.

7 References

Aguzzi, J., Flögel, S., Marini, S., Thomsen, L., Albiez, J., Weiss, P., Picardi, G., Calisti, M., Stefanni, S., Mirimin, L., Vecchi, F., Laschi, C., Branch, A., Clark, E.B., Foing, B., Wedler, A., Chatzievangelou, D., Tangherlini, M., Purser, A., Dartnell, L. & Danovaro, R. (2022). Developing technological synergies between deep-sea and space research. *Elementa: Science of the Anthropocene,* 10. DOI 10.1525/elementa.2021.00064

Altenhoff, A.M. & Dessimoz, C. (2009). Phylogenetic and functional assessment of orthologs inference projects and methods. *PLoS Comput Biol,* 5**:** e1000262. DOI 10.1371/journal.pcbi.1000262

Amarasinghe, S.L., Su, S., Dong, X., Zappia, L., Ritchie, M.E. & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol,* 21**:** 30. DOI 10.1186/s13059-020-1935-5

Aspengren, S., Sköld, H.N., Quiroga, G., Mårtensson, L. & Wallin, M. (2003). Noradrenaline- and melatonin-mediated regulation of pigment aggregation in fish melanophores. *Pigment Cell Research,* 16**:** 59-64.

Atig, R.K., Hsouna, S., Beraud-Colomb, E. & Abdelhak, S. (2009). [Mitochondrial DNA: properties and applications]. *Arch Inst Pasteur Tunis,* 86**:** 3-14.

Balloux, F. (2010). The worm in the fruit of the mitochondrial DNA tree. *Heredity*, *104*(5), 419.

Barrett, L.T., Overton, K., Stien, L.H., Oppedal, F. & Dempster, T. (2020). Effect of cleaner fish on sea lice in Norwegian salmon aquaculture: a national scale data analysis. *Int J Parasitol,* 50**:** 787-796. DOI 10.1016/j.ijpara.2019.12.005

Bentzur, A., Alon, S. & Shohat-Ophir, G. (2022). Behavioral Neuroscience in the Era of Genomics: Tools and Lessons for Analyzing High-Dimensional Datasets. *Int J Mol Sci,* 23. DOI 10.3390/ijms23073811

Berge, J. & Nahrgang, J. (2013). The Atlantic spiny lumpsucker *Eumicrotremus spinosus*: life history traits and the seemingly unlikely interaction with the pelagic amphipod *Themisto libellula*.

Black, J.A., Neuheimer, A.B., Horn, P.L., Tracey, D.M. & Drazen, J.C. (2021). Environmental, evolutionary, and ecological drivers of slow growth in deep-sea demersal teleosts. *Marine Ecology Progress Series,* 658**:** 1-26. DOI 10.3354/meps13591

Bolstad, G. H., Karlsson, S., Hagen, I. J., Fiske, P., Urdal, K., Sægrov, H., ... & Hindar, K. (2021). Introgression from farmed escapees affects the full life cycle of wild Atlantic salmon. *Science Advances*, *7*(52), eabj3397. DOI: 10.1126/sciadv.abj3397

Boore, J.L. (1999). Animal mitochondrial genomes. *Nucleic Acids Res,* 27**:** 1767-80. DOI 10.1093/nar/27.8.1767

Brunet, T.D.P. & Doolittle, W.F. (2018). The generality of Constructive Neutral Evolution. *Biology & Philosophy,* 33. DOI 10.1007/s10539-018-9614-6

Budney, L.A. & Hall, B.K. (2010). Comparative morphology and osteology of pelvic fin-derived midline suckers in lumpfishes, snailfishes and gobies. *Journal of Applied Ichthyology,* 26**:** 167-175. DOI 10.1111/j.1439-0426.2010.01398.x

Burrell, A.S., Disotell, T.R. & Bergey, C.M. (2015). The use of museum specimens with high-throughput DNA sequencers. *J Hum Evol,* 79**:** 35-44. DOI 10.1016/j.jhevol.2014.10.015

Byrkjedal, I., Rees, D.J. & Willassen, E. (2007). Lumping lumpsuckers: molecular and morphological insights into the taxonomic status of *Eumicrotremus spinosus* (Fabricius, 1776) and *Eumicrotremus eggvinii* Koefoed, 1956 (Teleostei: Cyclopteridae). *Journal of Fish Biology,* 71**:** 111-131. DOI 10.1111/j.1095-8649.2007.01550.x

Chen, F., Mackey, A.J., Vermunt, J.K. & Roos, D.S. (2007). Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS One,* 2**:** e383. DOI 10.1371/journal.pone.0000383

Cheon, S., Zhang, J. & Park, C. (2020). Is Phylotranscriptomics as Reliable as Phylogenomics? *Mol Biol Evol,* 37**:** 3672-3683. DOI 10.1093/molbev/msaa181

Chernova, N. (2008). Systematics and phylogeny of fish of the genus Liparis (Liparidae, Scorpaeniformes). *Journal of Ichthyology,* 48**:** 831-852.

Chernova, N.V., Friedlander, A.M., Turchik, A. & Sala, E. (2014). Franz Josef Land: extreme northern outpost for Arctic fishes. *PeerJ,* 2**:** e692. DOI 10.7717/peerj.692

Cui, R., Medeiros, T., Willemsen, D., Iasi, L.N.M., Collier, G.E., Graef, M., Reichard, M. & Valenzano, D.R. (2019). Relaxed Selection Limits Lifespan by Increasing Mutation Load. *Cell,* 178**:** 385-399 e20. DOI 10.1016/j.cell.2019.06.004

Davenport, J. (1985). Synopsis of biological data on the lumpsucker, *Cyclopterus lumpus* (Linnaeus, 1758): Food & Agriculture Org.

Davenport, J. & Bradshaw, C. (1995). Observations on skin colour changes in juvenile lumpsuckers. *Journal of fish biology,* 47**:** 143-154.

Davenport, J. & Kjørsvik, E. (1986). Buoyancy in the lumpsucker Cyclopterus lumpus. Journal of the Marine Biological Association of the United Kingdom, 66**:** 159-174.

Davenport, J. & Thorsteinsson, V. (1989). Observations on the colours of lumpsuckers, *Cyclopterus lumpus* L. *Journal of fish biology,* 35**:** 829-838.

Davenport, J. & Thorsteinsson, V. (1990). Sucker action in the lumpsucker *Cyclopterus lumpus* L. *Sarsia,* 75**:** 33-42.

Deady, S., Varian, S.J. & Fives, J.M. (1995). The use of cleaner-fish to control sea lice on two Irish salmon (*Salmo salar*) farms with particular reference to wrasse behaviour in salmon cages. *Aquaculture,* 131**:** 73-90.

Delroisse, J., Kang, V., Gouveneaux, A., Santos, R. & Flammang, P. (2023). Convergent evolution of attachment mechanisms in aquatic animals. *Convergent Evolution: Animal Form and Function.* Springer.

Derkarabetian, S., Benavides, L.R. & Giribet, G. (2019). Sequence capture phylogenomics of historical ethanol-preserved museum specimens: Unlocking the rest of the vault. *Mol Ecol Resour,* 19**:** 1531-1544. DOI 10.1111/1755-0998.13072

Donley, J.M., Sepulveda, C.A., Konstantinidis, P., Gemballa, S. & Shadwick, R.E. (2004). Convergent evolution in mechanical design of lamnid sharks and tunas. *Nature,* 429**:** 61-5. DOI 10.1038/nature02435

Drazen, J.C. & Haedrich, R.L. (2012). A continuum of life histories in deep-sea demersal fishes. *Deep Sea Research Part I: Oceanographic Research Papers,* 61**:** 34-42. DOI 10.1016/j.dsr.2011.11.002

Eisen, J.A. (1998). Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res,* 8**:** 163-7. DOI 10.1101/gr.8.3.163

Emms, D.M. & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol,* 20**:** 238. DOI 10.1186/s13059-019-1832-y

Eriksen, E., Durif, C.M.F. & Prozorkevich, D. (2014). Lumpfish (*Cyclopterus lumpus*) in the Barents Sea: development of biomass and abundance indices, and spatial distribution. *ICES Journal of Marine Science,* 71**:** 2398-2402. DOI 10.1093/icesjms/fsu059

Fitch, W.M. (1970). Distinguishing homologous from analogous proteins. *Syst Zool,* 19**:** 99-113.

Frutos, I., Kaiser, S., Pułaski, Ł., Studzian, M. & Błażewicz, M. (2022). Challenges and Advances in the Taxonomy of Deep-Sea Peracarida: From Traditional to Modern Methods. *Frontiers in Marine Science,* 9. DOI 10.3389/fmars.2022.799191

Gabaldon, T. (2008). Large-scale assignment of orthology: back to phylogenetics? *Genome Biol,* 9**:** 235. DOI 10.1186/gb-2008-9-10-235

Galtier, N., Nabholz, B., Glemin, S. & Hurst, G.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol,* 18**:** 4541-50. DOI 10.1111/j.1365-294X.2009.04380.x

Garcia De Leaniz, C., Gutierrez Rabadan, C., Barrento, S.I., Stringwell, R., Howes, P.N., Whittaker, B.A., Minett, J.F., Smith, R.G., Pooley, C.L., Overland, B.J., Biddiscombe, L., Lloyd, R., Consuegra, S., Maddocks, J.K., Deacon, P.T.J., Jennings, B.T., Rey Planellas, S., Deakin, A., Moore, A.I., Phillips, D., Bardera, G., Castanheira, M.F., Scolamacchia, M., Clarke, N., Parker, O., Avizienius, J., Johnstone, M. & Pavlidis, M. (2021). Addressing the welfare needs of farmed lumpfish: Knowledge gaps, challenges and solutions. *Reviews in Aquaculture,* 14**:** 139-155. DOI 10.1111/raq.12589

Gauthier, J., Grégoire, F. & Nozères, C. (2017). Assessment of Lumpfish *(Cyclopterus lumpus)* in the Gulf of St. Lawrence (3Pn, 4RS) in 2015: Fisheries and Oceans Canada, Ecosystems and Oceans Science.

Gavrilov, L.A. & Gavrilova, N.S. (2002). Evolutionary theories of aging and longevity. *ScientificWorldJournal,* 2**:** 339-56. DOI 10.1100/tsw.2002.96

Gerringer, M.E. (2019). On the Success of the Hadal Snailfishes. *Integr Org Biol,* 1**:** obz004. DOI 10.1093/iob/obz004

Gerringer, M.E., Dias, A.S., Von Hagel, A.A., Orr, J.W., Summers, A.P. & Farina, S. (2021). Habitat influences skeletal morphology and density in the snailfishes (family Liparidae). *Front Zool,* 18**:** 16. DOI 10.1186/s12983-021-00399-9

Glover, K.A., Solberg, M.F., Mcginnity, P., Hindar, K., Verspoor, E., Coulson, M.W., Hansen, M.M., Araki, H., Skaala, Ø. & Svåsand, T. (2017). Half a century of genetic interaction between farmed and wild Atlantic salmon: status of knowledge and unanswered questions. *Fish and Fisheries,* 18**:** 890-927.

Gosline, W.A. (1994). Function and structure in the paired fins of scorpaeniform fishes. *Environmental Biology of Fishes,* 40**:** 219-226.

Goulet, D., Green, J.M. & Shears, T.H. (1986). Courtship, spawning, and parental care behavior of the lumpfish, *Cyclopterus lumpus* L., in Newfoundland. *Canadian journal of zoology,* 64**:** 1320-1325.

Hale, M.E. (2000). Startle responses of fish without Mauthner neurons: escape behavior of the lumpfish (*Cyclopterus lumpus*). *The Biological Bulletin,* 199**:** 180-182.

Hatano, M., Abe, T., Wada, T. & Munehara, H. (2015). Ontogenetic metamorphosis and extreme sexual dimorphism in lumpsuckers: *Eumicrotremus asperrimus, Cyclopteropsis bergi* and *Cyclopteropsis lindbergi*, may be synonymous. *J Fish Biol,* 86**:** 1121-8. DOI 10.1111/jfb.12627

Heath, T.A., Hedtke, S.M. & Hillis, D.M. (2008). Taxon sampling and the accuracy of phylogenetic analyses. *Journal of systematics and evolution,* 46**:** 239-257.

Hellmuth, M., Wieseke, N., Lechner, M., Lenhof, H.P., Middendorf, M. & Stadler, P.F. (2015). Phylogenomics with paralogs. *Proc Natl Acad Sci U S A,* 112**:** 2058-63. DOI 10.1073/pnas.1412770112

Holst, J.C. (1993). Observations on the distribution of lumpsucker (*Cyclopterus lumpus*, L.) in the Norwegian Sea. *Fisheries Research,* 17**:** 369-372. DOI 10.1016/0165-7836(93)90136-u

Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Selly, S.L.C., Martin, S.a.M., Stevens, J.R., Santos, E.M., Davie, A. & Robledo, D. (2020). Harnessing genomics to fast-track genetic improvement in aquaculture. *Nat Rev Genet,* 21**:** 389-409. DOI 10.1038/s41576-020-0227-y

Hussain, L.S., Reddy, V. & Maani, C.V. (2022). Physiology, noradrenergic synapse. *StatPearls [Internet].* StatPearls Publishing.

Hvas, M., Folkedal, O., Imsland, A. & Oppedal, F. (2018). Metabolic rates, swimming capabilities, thermal niche and stress response of the lumpfish, *Cyclopterus lumpus*. *Biol Open,* 7. DOI 10.1242/bio.036079

Imsland, A.K., Reynolds, P., Eliassen, G., Hangstad, T.A., Foss, A., Vikingstad, E. & Elvegård, T.A. (2014). The use of lumpfish (*Cyclopterus lumpus* L.) to control sea lice (*Lepeophtheirus salmonis* Krøyer) infestations in intensively farmed Atlantic salmon (*Salmo salar* L.). *Aquaculture,* 424**:** 18-23.

Imsland, A.K., Reynolds, P., Eliassen, G., Hangstad, T.A., Nytrø, A.V., Foss, A., Vikingstad, E. & Elvegård, T.A. (2015). Feeding preferences of lumpfish (*Cyclopterus lumpus* L.) maintained in open net-pens with Atlantic salmon (*Salmo salar* L.). *Aquaculture,* 436**:** 47-51. DOI 10.1016/j.aquaculture.2014.10.048

Imsland, A.K.D., Hanssen, A., Nytrø, A.V., Reynolds, P., Jonassen, T.M., Hangstad, T.A., Elvegård, T.A., Urskog, T.C. & Mikalsen, B. (2018). It works! Lumpfish can significantly lower sea lice infestation in large-scale salmon farming. *Biology Open,* 7**:** bio036301.

Jaccoud, D., Peng, K., Feinstein, D. & Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res,* 29**:** E25. DOI 10.1093/nar/29.4.e25

Jamieson, A. (2015). *The hadal zone: life in the deepest oceans*: Cambridge University Press.

Jónsdóttir, Ó.D.B., Gíslason, D., Ólafsdóttir, G., Maduna, S., Hagen, S.B., Reynolds, P., Sveinsson, S. & Imsland, A.K.D. (2022). Lack of population genetic structure of lumpfish along the Norwegian coast: A reappraisal based on EST-STRs analyses. *Aquaculture,* 555. DOI 10.1016/j.aquaculture.2022.738230

Kane, E.A. & Higham, T.E. (2012). Life in the flow lane: differences in pectoral fin morphology suggest transitions in station-holding demand across species of marine sculpin. *Zoology,* 115**:** 223-232.

Karlsson, S., Diserud, O.H., Fiske, P. & Hindar, K. (2016). Widespread genetic introgression of escaped farmed Atlantic salmon in wild salmon populations. *ICES Journal of Marine Science,* 73**:** 2488-2498. DOI 10.1093/icesjms/fsw121

Karlsson, S., Moen, T., Lien, S., Glover, K.A. & Hindar, K. (2011). Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Mol Ecol Resour,* 11 Suppl 1**:** 247-53. DOI 10.1111/j.1755-0998.2010.02959.x

Kennedy, J., Jónsson, S.Þ., Kasper, J.M. & Ólafsson, H.G. (2015). Movements of female lumpfish (*Cyclopterus lumpus*) around Iceland. *ICES Journal of Marine Science,* 72**:** 880-889. DOI 10.1093/icesjms/fsu170

Knope, M.L. (2013). Phylogenetics of the marine sculpins (Teleostei: Cottidae) of the North American Pacific Coast. *Mol Phylogenet Evol,* 66**:** 341-9. DOI 10.1016/j.ympev.2012.10.008

Knudsen, S.W., Moller, P.R. & Gravlund, P. (2007). Phylogeny of the snailfishes (Teleostei: Liparidae) based on molecular and morphological data. *Mol Phylogenet Evol,* 44**:** 649-66. DOI 10.1016/j.ympev.2007.04.005

Kondrashov, F.A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc Biol Sci,* 279**:** 5048-57. DOI 10.1098/rspb.2012.1108

Leclercq, E., Zerafa, B., Brooker, A.J., Davie, A. & Migaud, H. (2018). Application of passive-acoustic telemetry to explore the behaviour of ballan wrasse (*Labrus bergylta*) and lumpfish (*Cyclopterus lumpus*) in commercial Scottish salmon sea-pens. *Aquaculture,* 495**:** 1-12. DOI 10.1016/j.aquaculture.2018.05.024

Li, D.H., Shi, W., Munroe, T.A., Gong, L. & Kong, X.Y. (2015). Concerted Evolution of Duplicate Control Regions in the Mitochondria of Species of the Flatfish Family Bothidae (Teleostei: Pleuronectiformes). *PLoS One,* 10**:** e0134580. DOI 10.1371/journal.pone.0134580

Li, L., Stoeckert, C.J., Jr. & Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res,* 13**:** 2178-89. DOI 10.1101/gr.1224503

Liu, Y., Cotton, J.A., Shen, B., Han, X., Rossiter, S.J. & Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. *Current Biology,* 20**:** R53-R54.

Lutz, M.J., Caldeira, K., Dunbar, R.B. & Behrenfeld, M.J. (2007). Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *Journal of Geophysical Research,* 112. DOI 10.1029/2006jc003706

Maduna, S.N., Vivian-Smith, A., Jonsdottir, O.D.B., Imsland, A.K.D., Klutsch, C.F.C., Nyman, T., Eiken, H.G. & Hagen, S.B. (2022). Mitogenomics of the suborder Cottoidei (Teleostei: Perciformes): Improved assemblies, mitogenome features, phylogeny, and ecological implications. *Genomics,* 114**:** 110297. DOI 10.1016/j.ygeno.2022.110297

Martins, C.I., Galhardo, L., Noble, C., Damsgard, B., Spedicato, M.T., Zupa, W., Beauchaud, M., Kulczykowska, E., Massabuau, J.C., Carter, T., Planellas, S.R. & Kristiansen, T. (2012). Behavioural indicators of welfare in farmed fish. *Fish Physiol Biochem,* 38**:** 17-41. DOI 10.1007/s10695-011-9518-8

Mecklenburg, C. & Sheiko, B. (2003). Family Cyclopteridae Bonaparte 1831—lumpsuckers. *Annotated Checklists of Fishes***:** 17.

Meyer, A. & Van De Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays,* 27**:** 937-45. DOI 10.1002/bies.20293

Migaud, H., Bell, G., Cabrita, E., Mcandrew, B., Davie, A., Bobe, J., Herráez, M.P. & Carrillo, M. (2013). Gamete quality and broodstock management in temperate fish. *Reviews in Aquaculture,* 5**:** S194-S223. DOI 10.1111/raq.12025

Miller, E.C., Martinez, C.M., Friedman, S.T., Wainwright, P.C., Price, S.A. & Tornabene, L. (2022). Alternating regimes of shallow and deep-sea diversification explain a species-richness paradox in marine fishes. *Proc Natl Acad Sci U S A,* 119**:** e2123544119. DOI 10.1073/pnas.2123544119

Miller, J.R., Zhou, P., Mudge, J., Gurtowski, J., Lee, H., Ramaraj, T., Walenz, B.P., Liu, J., Stupar, R.M., Denny, R., Song, L., Singh, N., Maron, L.G., Mccouch, S.R., Mccombie, W.R., Schatz, M.C., Tiffin, P., Young, N.D. & Silverstein, K.a.T. (2017). Hybrid assembly with long and short reads improves discovery of gene family expansions. *BMC Genomics,* 18**:** 541. DOI 10.1186/s12864-017-3927-8

Moritz, C. (1994). Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology,* 3**:** 401-411.

Mu, Y., Bian, C., Liu, R., Wang, Y., Shao, G., Li, J., Qiu, Y., He, T., Li, W., Ao, J., Shi, Q. & Chen, X. (2021). Whole genome sequencing of a snailfish from the Yap Trench (~7,000 m) clarifies the molecular mechanisms underlying adaptation to the deep sea. *PLoS Genet,* 17**:** e1009530. DOI 10.1371/journal.pgen.1009530

Ncube, D., Tallafuss, A., Serafin, J., Bruckner, J., Farnsworth, D.R., Miller, A.C., Eisen, J.S. & Washbourne, P. (2022). A conserved transcriptional fingerprint of multi-neurotransmitter neurons necessary for social behavior. *BMC genomics,* 23**:** 675.

Nilsson Sköld, H., Aspengren, S. & Wallin, M. (2013). Rapid color change in fish and amphibians–function, regulation, and emerging applications. *Pigment cell & melanoma research,* 26**:** 29-38.

Oguri, K. & Noguchi, T. (2017). Deepest fish ever recorded, documented at depths of 8,178 m in Mariana Trench. JAMSTEC Press Release.

Oliveira, R.F. & Bshary, R. (2021). Expanding the concept of social behavior to interspecific interactions. *Ethology,* 127**:** 758-773. DOI 10.1111/eth.13194

Orr, J.W., Spies, I., Stevenson, D.E., Longo, G.C., Kai, Y., Ghods, S. & Hollowed, M. (2019). Molecular phylogenetics of snailfishes (Cottoidei: Liparidae) based on MtDNA and RADseq genomic analyses, with comments on selected morphological characters. *Zootaxa,* 4642**:** zootaxa 4642 1 1. DOI 10.11646/zootaxa.4642.1.1

Pampoulie, C., Skirnisdottir, S., Olafsdottir, G., Helyar, S.J., Thorsteinsson, V., Jónsson, S.Þ., Fréchet, A., Durif, C.M.F., Sherman, S., Lampart-Kałużniacka, M., Hedeholm, R., Ólafsson, H., Daníelsdóttir, A.K. & Kasper, J.M. (2014). Genetic structure of the lumpfish *Cyclopterus lumpus* across the North Atlantic. *ICES Journal of Marine Science,* 71**:** 2390-2397. DOI 10.1093/icesjms/fsu071

Pereira, S.L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular biology,* 23**:** 745-752.

Popper, A.N. & Fay, R.R. (2011). Rethinking sound detection by fishes. *Hear Res,* 273**:** 25-36. DOI 10.1016/j.heares.2009.12.023

Pountney, S.M., Lein, I., Migaud, H. & Davie, A. (2020). High temperature is detrimental to captive lumpfish (*Cyclopterus lumpus*, L) reproductive performance. *Aquaculture,* 522. DOI 10.1016/j.aquaculture.2020.735121

Powell, A., Treasurer, J.W., Pooley, C.L., Keay, A.J., Lloyd, R., Imsland, A.K. & Garcia De Leaniz, C. (2018). Use of lumpfish for sea-lice control in salmon farming: challenges and opportunities. *Reviews in Aquaculture,* 10**:** 683-702.

Priede, I.G. & Froese, R. (2013). Colonization of the deep sea by fishes. *J Fish Biol,* 83**:** 1528-50. DOI 10.1111/jfb.12265

Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M. & Alfaro, M.E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature,* 559**:** 392-395. DOI 10.1038/s41586-018-0273-1

Radnaeva, L.D., Popov, D.V., Grahl-Nielsen, O., Khanaev, I.V., Bazarsadueva, S.V. & Käkelä, R. (2017). Fatty acid composition in the white muscle of Cottoidei fishes of Lake Baikal reflects their habitat depth. *Environmental Biology of Fishes,* 100**:** 1623-1641.

Ramirez-Calero, S., Paula, J.R., Otjacques, E., Rosa, R., Ravasi, T. & Schunter, C. (2022). Neuro-molecular characterization of fish cleaning interactions. *Sci Rep,* 12**:** 8468. DOI 10.1038/s41598-022-12363-6

Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., Levin, L.A., Martinez Arbizu, P., Menot, L., Buhl-Mortensen, P., Narayanaswamy, B.E., Smith, C.R., Tittensor, D.P., Tyler, P.A., Vanreusel, A. & Vecchione, M. (2010). Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. *Biogeosciences,* 7**:** 2851-2899. DOI 10.5194/bg-7-2851-2010

Raxworthy, C.J. & Smith, B.T. (2021). Mining museums for historical DNA: advances and challenges in museomics. *Trends Ecol Evol,* 36**:** 1049-1060. DOI 10.1016/j.tree.2021.07.009

Ruane, S. & Austin, C.C. (2017). Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Mol Ecol Resour,* 17**:** 1003-1008. DOI 10.1111/1755-0998.12655

Rüber, L., Britz, R., Kullander, S. O., & Zardoya, R. (2004). Evolutionary and biogeographic patterns of the Badidae (Teleostei: Perciformes) inferred from mitochondrial and nuclear DNA sequence data. *Molecular phylogenetics and evolution*, *32*(3), 1010-1022. DOI 10.1016/j.ympev.2004.04.020

Satoh, T.P., Miya, M., Mabuchi, K. & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. *BMC Genomics,* 17**:** 719. DOI 10.1186/s12864-016-3054-y

Sayer, M. & Reader, J. (1996). Exposure of goldsinny, rock cook and corkwing wrasse to low temperature and low salinity: survival, blood physiology and seasonal variation. *Journal of fish biology,* 49**:** 41-63.

Schmidt, P. (1927). A revision of the cottoid fishes of the genus Artediellus. *Proceedings of the United States National Museum*.

Schopka, S.A. (1974). Preliminary results from tagging of lumpsucker *(Cyclopterus lumpus). Icelandic waters 1971-3*.

Serb, J.M. & Eernisse, D.J. (2008). Charting evolution's trajectory: using molluscan eye diversity to understand parallel and convergent evolution. *Evolution: Education and Outreach,* 1**:** 439-447.

Shen, Y., Dai, W., Gao, Z., Yan, G., Gan, X. & He, S. (2017). Molecular phylogeny and divergence time estimates using the mitochondrial genome for the hadal snailfish from the Mariana trench. *Sci Bull (Beijing),* 62**:** 1106-1108. DOI 10.1016/j.scib.2017.07.010

Sideleva, V. (1996). Comparative character of the deep-water and inshore cottoid fishes endemic to Lake Baikal. *Journal of fish biology,* 49**:** 192-206.

Sideleva, V. & Fialkov, V. (2015). Cottoid fishes (Cottoidei) in deep-water hydrothermal vent community in Frolikha Bay, Lake Baikal. *Aquatic biological resources*.

Smith, W.L. & Busby, M.S. (2014). Phylogeny and taxonomy of sculpins, sandfishes, and snailfishes (Perciformes: Cottoidei) with comments on the phylogenetic significance of their early-life-history specializations. *Molecular Phylogenetics and Evolution,* 79**:** 332-352.

Spillane, J.L., Lapolice, T.M., Macmanes, M.D. & Plachetzki, D.C. (2021). Signal, bias, and the role of transcriptome assembly quality in phylogenomic inference. *BMC Ecol Evol,* 21**:** 43. DOI 10.1186/s12862-021-01772-2

Staven, F.R., Nordeide, J.T., Gesto, M., Andersen, P., Patel, D.M. & Kristensen, T. (2021). Behavioural and physiological responses of lumpfish (*Cyclopterus lumpus*) exposed to Atlantic salmon (*Salmo salar*) sensory cues. *Aquaculture,* 544**:** 737066.

Staven, F.R., Nordeide, J.T., Imsland, A.K., Andersen, P., Iversen, N.S. & Kristensen, T. (2019). Is Habituation Measurable in Lumpfish *Cyclopterus lumpus* When Used as Cleaner Fish in Atlantic Salmon *Salmo salar* Aquaculture? *Front Vet Sci,* 6**:** 227. DOI 10.3389/fvets.2019.00227

Thomson, K.S. & Simanek, D.E. (1977). Body form and locomotion in sharks. *American Zoologist,* 17**:** 343-354.

Thornqvist, P.O., Mccarrick, S., Ericsson, M., Roman, E. & Winberg, S. (2019). Bold zebrafish (*Danio rerio*) express higher levels of delta opioid and dopamine D2 receptors in the brain compared to shy fish. *Behav Brain Res,* 359**:** 927-934. DOI 10.1016/j.bbr.2018.06.017

Treasurer, J. (2018). An introduction to sea lice and the rise of cleaner fish. *Cleaner fish biology and aquaculture applications***:** 3-25.

Tully, K. & Bolshakov, V.Y. (2010). Emotional enhancement of memory: how norepinephrine enables synaptic plasticity. *Molecular brain,* 3**:** 1-9.

Ueno, T. (1970). *Fauna Japonica: Cyclopteridae (Pisces)*: Academic Press of Japan.

Vallejo, R.L., Silva, R.M.O., Evenhuis, J.P., Gao, G., Liu, S., Parsons, J.E., Martin, K.E., Wiens, G.D., Lourenco, D.a.L., Leeds, T.D. & Palti, Y. (2018). Accurate genomic predictions for BCWD resistance in rainbow trout are achieved using low-density SNP panels: Evidence that long-range LD is a major contributing factor. *J Anim Breed Genet*. DOI 10.1111/jbg.12335

Volff, J.N. (2005). Genome evolution and biodiversity in teleost fish. *Heredity (Edinb),* 94**:** 280-94. DOI 10.1038/sj.hdy.6800635

Voskoboinikova, O.S., Kudryavtseva, O.Y., Orlov, A.M., Orlova, S.Y., Nazarkin, M.V., Chernova, N.V. & Maznikova, O.A. (2020). Relationships and Evolution of Lumpsuckers of the Family Cyclopteridae (Cottoidei). *Journal of Ichthyology,* 60**:** 154-181. DOI 10.1134/s0032945220020204

Vu, T.-D., Iwasaki, Y., Shigenobu, S., Maruko, A., Oshima, K., Iioka, E., Huang, C.-L., Abe, T., Tamaki, S. & Lin, Y.-W. (2020). Behavioral and brain-transcriptomic synchronization between the two opponents of a fighting pair of the fish *Betta splendens*. *PLoS genetics,* 16**:** e1008831.

Wang, K., Shen, Y., Yang, Y., Gan, X., Liu, G., Hu, K., Li, Y., Gao, Z., Zhu, L., Yan, G., He, L., Shan, X., Yang, L., Lu, S., Zeng, H., Pan, X., Liu, C., Yuan, Y., Feng, C., Xu, W., Zhu, C., Xiao, W., Dong, Y., Wang, W., Qiu, Q. & He, S. (2019). Morphology and genome of a snailfish from the Mariana Trench provide insights into deep-sea adaptation. *Nat Ecol Evol,* 3**:** 823-833. DOI 10.1038/s41559-019-0864-8

Ward, N. & Moreno-Hagelsieb, G. (2014). Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: how much do we miss? *PloS one,* 9**:** e101850.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. & Hebert, P.D. (2005). DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B Biol Sci,* 360**:** 1847-57. DOI 10.1098/rstb.2005.1716

Weber, A.A., Hugall, A.F. & O'hara, T.D. (2020). Convergent Evolution and Structural Adaptation to the Deep Ocean in the Protein-Folding Chaperonin CCTalpha. *Genome Biol Evol,* 12**:** 1929-1942. DOI 10.1093/gbe/evaa167

Whittaker, B. A., Consuegra, S., & de Leaniz, C. G. (2018). Genetic and phenotypic differentiation of lumpfish (*Cyclopterus lumpus*) across the North Atlantic: implications for conservation and aquaculture. *PeerJ*, *6*, e5974. DOI 10.7717/peerj.5974.

Wiedmann, M.A., Primicerio, R., Dolgov, A., Ottesen, C.A. & Aschan, M. (2014). Life history variation in Barents Sea fish: implications for sensitivity to fishing in a changing environment. *Ecol Evol,* 4**:** 3596-611. DOI 10.1002/ece3.1203

Worobey, M., Watts, T.D., Mckay, R.A., Suchard, M.A., Granade, T., Teuwen, D.E., Koblin, B.A., Heneine, W., Lemey, P. & Jaffe, H.W. (2016). 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature,* 539**:** 98-101. DOI 10.1038/nature19827

Yoshida, G.M., Lhorente, J.P., Correa, K., Soto, J., Salas, D. & Yanez, J.M. (2019). Genome-Wide Association Study and Cost-Efficient Genomic Predictions for Growth and Fillet Yield in Nile Tilapia (*Oreochromis niloticus*). *G3 (Bethesda),* 9**:** 2597-2607. DOI 10.1534/g3.119.400116

Young, A.D. & Gillung, J.P. (2019). Phylogenomics — principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology,* 45**:** 225-247. DOI 10.1111/syen.12406

Zhang, J. & Yang, J.R. (2015). Determinants of the rate of protein sequence evolution. *Nat Rev Genet,* 16**:** 409-20. DOI 10.1038/nrg3950

Zhang, X., Cvetkovska, M., Morgan-Kiss, R., Hüner, N. P., & Smith, D. R. (2021). Draft genome sequence of the Antarctic green alga *Chlamydomonas* sp. UWO241. *Iscience*, *24*(2), 102084. DOI 10.1016/j.isci.2021.102084

Zottoli, S.J. (1978). Comparison of Mauthner cell size in teleosts. *J Comp Neurol,* 178**:** 741-57. DOI 10.1002/cne.901780409

Paper I

MITOCHONDRIAL DNA PART B: RESOURCES
2023, VOL. 8, NO. 3, 364–367
https://doi.org/10.1080/23802359.2023.2184649

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

The complete mitochondrial genome of the Atlantic spiny lumpsucker *Eumicrotremus spinosus* (Fabricius, 1776)

Likith R. Pinninti[a]*, Marius F. Maurstad[b]*, Siv N.K Hoff[b], Torstein Kristensen[a], Leslie R. Noble[a], Sissel Jentoft[b] and Jorge M. O. Fernandes[a]

[a]Faculty of Biosciences and Aquaculture (FBA), Nord University, Bodø, Norway; [b]Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo, Norway

ABSTRACT

The complete mitogenome of the Atlantic spiny lumpsucker (*Eumicrotremus spinosus)* was generated using the PacBio Sequel II HiFi sequencing platform. The mitogenome assembly has a length of 19,281 bp and contains 13 protein-coding sequences, 22 tRNA genes, 2 rRNA genes, one control region containing the D-loop (2383 bp) and a duplicate control region (1133 bp) Phylogenetic analysis using maximum likelihood revealed that *E. spinosus* is closely related to the Siberian lumpsucker (*E. asperrimus*). The mitogenome of the spiny lumpsucker will be useful in population genomics and systematic studies of Cyclopteridae, Liparidae, and Cottidae.

ARTICLE HISTORY
Received 5 January 2023
Accepted 21 February 2023

KEYWORDS
Cyclopteridae; *E. spinosus*; Mitogenome; PacBio

The Atlantic spiny lumpsucker, *Eumicrotremus spinosus* (Fabricius, 1776), is a deep-sea benthic fish native to the Arctic and coastal North Atlantic. Its morphological characteristics and particularly the presence of spiny tubercles and the ventral suction disk confirmed its taxonomic identification (Figure 1). To understand the phylogeny of Cyclopteridae family members, we sequenced the mitochondrial genome of *E. spinosus*. The specimen of *E. spinosus* was collected from the Barents Sea during an expedition as part of The Nansen Legacy project (https://arvenetternansen.com/), which was conducted by trained scientists and according to the European Union animal experimentation guidelines on the protection of animals used for scientific purposes (directive 2010/63/UE). The *E. spinosus* specimen was deposited at the University of Oslo (contact person: Sissel Jentoft; email: sissel.jentoft@ibv.uio.no) under the voucher number 787912a2-9c81-11e8-9126-8c164557e466.

The whole fish was soaked in 96% ethanol and placed at 4 °C for two weeks; then, the ethanol was decanted before storage at −80 °C. DNA isolation and sequencing were performed by the Norwegian Sequencing Center. Genomic DNA was extracted from heart tissue using the nucleated tissue/-blood protocol from the Circulomics Nanobind BIG DNA kit (Circulomics Inc.). The libraries were prepared using the Pacific Biosciences protocol for HiFi library prep using SMRTbell® Express Template Prep Kit 2.0. A total of 5.77 µg DNA was sheared into 15–20 kb fragments using Megaruptor 3. After clean-up, we had ∼ 3 µg of fragmented DNA, which was used for library preparation. The final library was size selected using BluePippin with an 11 kb cutoff, resulting in ∼0.9 µg of genomic DNA. The library was sequenced on one 8 M SMRT cell on a Sequel II instrument using Sequel II Binding kit 2.2 and Sequencing chemistry v2.0. Loading was performed by adaptive loading, with a movie time of 30 h.

Mitochondrial reads were extracted from the whole genome data using MitoHiFi v2.2 (Uliano-Silva et al. 2021) with the mitogenome of *Cyclopterus lumpus* (MN122882.1) as a mapping reference. Extracted reads were assembled with Flye v2.9.1 (Kolmogorov et al. 2019) to obtain a full-length mitogenome assembly. The genome was annotated using MITOS (Bernt et al. 2013) and MitoFish annotator (Iwasaki et al. 2013). The complete *E. spinosus* mitogenome consisted of 19.2 kb (GenBank accession no. OP728784) and had a typical mitochondrial genome structure consisting of 13 protein-coding genes PCGs, 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes, and three non-coding regions (ncDNA) (Figure 2). Additionally, the control region (CR) was duplicated.

The *E. spinosus* mitogenome was similar to that of *C. lumpus*, which also comprises three non-coding regions, two *trnL* and two *trnS* genes (Maduna et al. 2022). In the H-strand, two CR regions, one non-coding region (the homo polymer region), 12 PCGs, and 14 tRNA genes were present. The two control regions were adjacent to each other, with the CR1 region (2383 bp) located between the *trnT* and *trnP*, and the CR2 region (1133 bp) found between the *trnP* and *trnF* genes.

CONTACT Sissel Jentoft sissel.jentoft@ibv.uio.no Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo, Norway; Leslie R. Noble leslie.r.noble@nord.no; Jorge M. O. Fernandes jorge.m.fernandes@nord.no Faculty of Biosciences and Aquaculture (FBA), Nord University, Bodø, Norway
*Equal contribution.

The presence of two CR regions is linked to concerted evolutionary events in non-coding regions (Li et al. 2015). Also, a third non-coding region (intergenic spacer) of 133 bp with a C-tract homopolymer was present between *trnL* and the *ND1* gene on the H-strand. Three non-coding regions have also been observed in other fish species (Maduna et al. 2022).

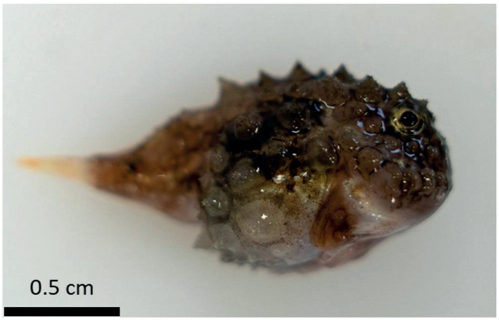


Figure 1. Photograph of the *E. spinosus* specimen used in the present study. The morphological traits typical of this species are spiny tubercles and a suction disk on the ventral side.

The L-strand encoded the origin of replication (OL) of 43 bp, 8 tRNA genes and the *ND6* protein-coding gene. The mitogenome arrangement and structure were very similar to a typical vertebrate mitogenome, but the presence of three non-coding regions contribute to our knowledge about the diversity and organization of mitochondrial genomes.

To understand the phylogenetic relationships among the members of Cyclopteridae, Liparidae, and their closest Cottidae relatives, all their mitochondrial PCGs were aligned using MAFFT v7.471 (Katoh and Standley 2013), trimmed with trimAL (Capella-Gutierrez et al. 2009) to remove poorly aligned, divergent, and ambiguous regions, and then a maximum likelihood phylogenetic tree was generated using IQ-TREE2 v2.1.2 (Minh et al. 2020). The best-fitting evolution model was automatically selected (JC + SYM + K80) and the phylogenetic reconstruction was performed with the parameters -B 1000, -wbt and 1000 bootstrap replicates (Nguyen et al. 2015; Hoang et al. 2018). Our phylogeny followed the currently accepted taxonomic relationships between all the closely related group members of the Cyclopteridae, Liparidae, and Cottidae families (Maduna et al. 2022) with more than

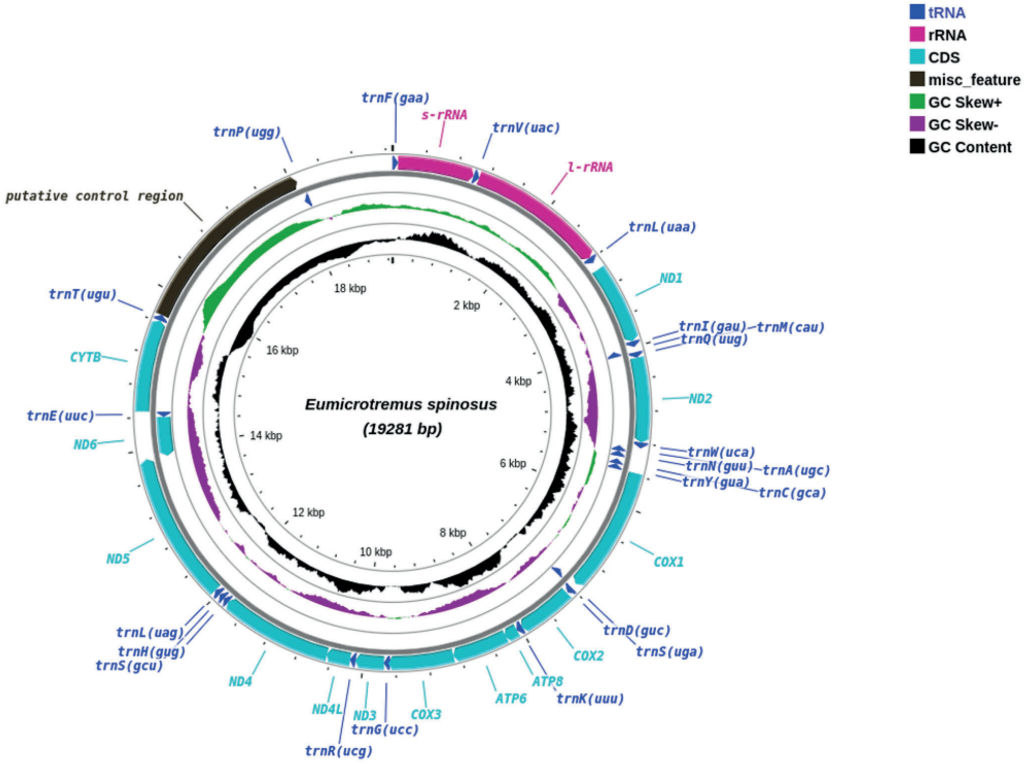


Figure 2. Circular plot of the Atlantic spiny lumpsucker (*E. spinosus*) mitogenome displaying both heavy (outer circle) and light (inner circle) strands. It has a typical vertebrate mitochondrial structure with 13 PCGs, two rRNA genes and 22 tRNA genes.

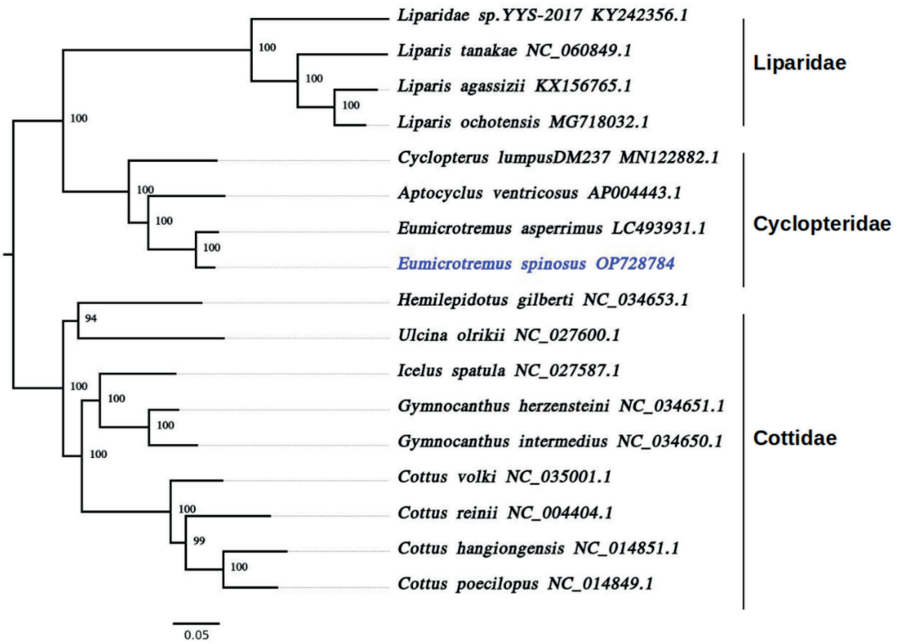


Figure 3. Maximum likelihood phylogeny of 17 infraorder Cottales (Teleostei: Perciformes) species based on their mitogenomes. The tree has more than 94% bootstrap support for each node. Accession numbers for each species are shown after the name of the species. The genome sequence generated in this study (OP728784) is labeled in violet and the branches are indicated in black.

94% bootstrap support for each node, as shown in Figure 3. Moreover, it confirmed that Cyclopteridae and Liparidae are sister clades and *E. spinosus* is most closely related to the Siberian lumpsucker (*E. asperrimus*). The data generated from this study will be a valuable resource for future comparative genomics and phylogenetics studies of molecular evolution and systematics in deep-sea fish.

Ethics statement

The specimen used in this study was collected in a responsible manner, i.e. in connection to the first joint research cruise of the Nansen Legacy project (https://arvenetternansen.com/nb/arven-etter-nansen/), as part of larger hauls for stock assessments. The fish were humanely sacrificed before sampling in accordance with the guidelines set by national and international animal welfare laws (e.g. www.norecopa.no), and thus no specific legislation were needed.

Authors' contributions

The study was planned by S.J., J.M.O.F., and L.R.N. L.R.P. and M.F.M. performed the analysis, and L.R.P. wrote the manuscript. We especially thank S.J. and S.N.K.H. who sampled the specimen used and initiated the sequencing. Finally, I thank J.M.O.F. for his overall support. S.J., S.N.K.H, M.F.M., J.M.O.F., and L.R.N. have reviewed the manuscript.

Disclosure statement

No conflict of interest was reported by the author(s).

Funding

We thank Nord University for their financial support. Sampling was conducted as part of the Nansen Legacy project [RCN# 272721] funded by the Research Council of Norway (RCN). The sequencing was financed and provided by the Norwegian Sequencing Center (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo which are supported by RCN over the "Functional Genomics" and "Infrastructure" programs and the Southeastern Regional Health Authorities.

Data availability statement

The mitogenome sequence data can be found in GenBank under the accession number (OP728784).

References

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2):313–319.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25(15):1972–1973.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Molecular Biology and Evolution. 35(2):518–522.

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, Sado T, Mabuchi K, Takeshima H, Miya M. 2013. MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol Biol Evol. 30(11): 2531–2540.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 37(5):540–546.

Li DH, Shi W, Munroe TA, Gong L, Kong XY. 2015. Concerted evolution of duplicate control regions in the mitochondria of species of the flatfish family Bothidae (Teleostei: Pleuronectiformes). PLoS ONE. 10(8): e0134580.

Maduna SN, Vivian-Smith A, Jónsdóttir ÓDB, Imsland AK, Klütsch CF, Nyman T, Eiken HG, Hagen SB. 2022. Mitogenomics of the suborder Cottoidei (Teleostei: Perciformes): improved assemblies, mitogenome features, phylogeny, and ecological implications. Genomics. 114(2): 110297.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Molecular Biology and Evolution. 37(5):1530–1534.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Uliano-Silva M, Nunes JGF, Krasheninnikova K. 2021. marcelauliano/MitoHiFi: mitohifi_v2.0. https://zenodo.org/record/5205678#.Y_cvAXZBzIU

List of previously published theses for PhD in Aquaculture / PhD in Aquatic Biosciences, Nord University

No. 1 (2011)
PhD in Aquaculture
Chris André Johnsen
Flesh quality and growth of farmed Atlantic salmon (*Salmo salar* L.) in relation to feed, feeding, smolt type and season
ISBN: 978-82-93165-00-2

No. 2 (2012)
PhD in Aquaculture
Jareeporn Ruangsri
Characterization of antimicrobial peptides in Atlantic cod
ISBN: 978-82-93165-01-9

No. 3 (2012)
PhD in Aquaculture
Muhammad Naveed Yousaf
Characterization of the cardiac pacemaker and pathological responses to cardiac diseases in Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-02-6

No. 4 (2012)
PhD in Aquaculture
Carlos Frederico Ceccon Lanes
Comparative Studies on the quality of eggs and larvae from broodstocks of farmed and wild Atlantic cod
ISBN: 978-82-93165-03-3

No. 5 (2012)
PhD in Aquaculture
Arvind Sundaram
Understanding the specificity of the innate immune response in teleosts: Characterisation and differential expression of teleost-specific Toll-like receptors and microRNAs
ISBN: 978-82-93165-04-0

No. 6 (2012)
PhD in Aquaculture
Teshome Tilahun Bizuayehu
Characterization of microRNA during early ontogeny and sexual development of Atlantic halibut (*Hippoglossus hippoglossus* L.)
ISBN: 978-82-93165-05-7

No. 7 (2013)
PhD in Aquaculture
Binoy Rajan
Proteomic characterization of Atlantic cod skin mucosa – Emphasis on innate immunity and lectins
ISBN: 978-82-93165-06-04

No. 8 (2013)
PhD in Aquaculture
Anusha Krishanthi Shyamali Dhanasiri
Transport related stress in zebrafish: physiological responses and bioremediation
ISBN: 978-82-93165-07-1

No. 9 (2013)
PhD in Aquaculture
Martin Haugmo Iversen
Stress and its impact on animal welfare during commercial production of Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-08-8

No. 10 (2013)
PhD in Aquatic Biosciences
Alexander Jüterbock
Climate change impact on the seaweed *Fucus serratus*, a key foundational species on North Atlantic rocky shores
ISBN: 978-82-93165-09-5

No. 11 (2014)
PhD in Aquatic Biosciences
Amod Kulkarni
Responses in the gut of black tiger shrimp *Penaeus monodon* to oral vaccine candidates against white spot disease
ISBN: 978-82-93165-10-1

No. 12 (2014)
PhD in Aquatic Biosciences
Carlo C. Lazado
Molecular basis of daily rhythmicity in fast skeletal muscle of Atlantic cod *(Gadus morhua)*
ISBN: 978-82-93165-11-8

No. 13 (2014)
PhD in Aquaculture
Joanna Babiak
Induced masculinization of Atlantic halibut (*Hippoglossus hippoglossus* L.): towards the goal of all-female production
ISBN: 978-82-93165-12-5

No. 14 (2015)
PhD in Aquaculture
Cecilia Campos Vargas
Production of triploid Atlantic cod: A comparative study of muscle growth dynamics and gut morphology
ISBN: 978-82-93165-13-2

No. 15 (2015)
PhD in Aquatic Biosciences
Irina Smolina
Calanus in the North Atlantic: species identification, stress response, and population genetic structure
ISBN: 978-82-93165-14-9

No. 16 (2016)
PhD in Aquatic Biosciences
Lokesh Jeppinamogeru
Microbiota of Atlantic salmon (*Salmo salar L.*), during their early and adult life
ISBN: 978-82-93165-15-6

No. 17 (2017)
PhD in Aquatic Biosciences
Christopher Edward Presslauer
Comparative and functional analysis of microRNAs during zebrafish gonadal development
ISBN: 978-82-93165-16-3

No. 18 (2017)
PhD in Aquatic Biosciences
Marc Jürgen Silberberger
Spatial scales of benthic ecosystems in the sub-Arctic Lofoten-Vesterålen region
ISBN: 978-82-93165-17-0

No. 19 (2017)
PhD in Aquatic Biosciences
Marvin Choquet
Combining ecological and molecular approaches to redefine the baseline knowledge of the genus Calanus in the North Atlantic and the Arctic Oceans
ISBN: 978-82-93165-18-7

No. 20 (2017)
PhD in Aquatic Biosciences
Torvald B. Egeland
Reproduction in Arctic charr – timing and the need for speed
ISBN: 978-82-93165-19-4

No. 21 (2017)
PhD in Aquatic Biosciences
Marina Espinasse
Interannual variability in key zooplankton species in the North-East Atlantic: an analysis based on abundance and phenology
ISBN: 978-82-93165-20-0

No. 22 (2018)
PhD in Aquatic Biosciences
Kanchana Bandara
Diel and seasonal vertical migrations of high-latitude zooplankton: knowledge gaps and a high-resolution bridge
ISBN: 978-82-93165-21-7

No. 23 (2018)
PhD in Aquatic Biosciences
Deepti Manjari Patel
Characterization of skin immune and stress factors of lumpfish, *Cyclopterus lumpus*
ISBN: 978-82-93165-21-7

No. 24 (2018)
PhD in Aquatic Biosciences
Prabhugouda Siriyappagouder
The intestinal mycobiota of zebrafish – community profiling and exploration of the impact of yeast exposure early in life
ISBN: 978-82-93165-23-1

No. 25 (2018)
PhD in Aquatic Biosciences
Tor Erik Jørgensen
Molecular and evolutionary characterization of the Atlantic cod mitochondrial genome
ISBN: 978-82-93165-24-8

No. 26 (2018)
PhD in Aquatic Biosciences
Yangyang Gong
Microalgae as feed ingredients for Atlantic salmon
ISBN: 978-82-93165-25-5

No. 27 (2018)
PhD in Aquatic Biosciences
Ove Nicolaisen
Approaches to optimize marine larvae production
ISBN: 978-82-93165-26-2

No. 28 (2019)
PhD in Aquatic Biosciences
Qirui Zhang
The effect of embryonic incubation temperature on the immune response of larval and adult zebrafish (*Danio rerio*)
ISBN: 978-82-93165-27-9

No. 29 (2019)
PhD in Aquatic Biosciences
Andrea Bozman
The structuring effects of light on the deep-water scyphozoan *Periphylla periphylla*
ISBN: 978-82-93165-28-6

No. 30 (2019)
PhD in Aquatic Biosciences
Helene Rønquist Knutsen
Growth and development of juvenile spotted wolffish (*Anarhichas minor*) fed microalgae incorporated diets
ISBN: 978-82-93165-29-3

No. 31 (2019)
PhD in Aquatic Biosciences
Shruti Gupta
Feed additives elicit changes in the structure of the intestinal bacterial community of Atlantic salmon
ISBN: 978-82-93165-30-9

No. 32 (2019)
PhD in Aquatic Biosciences
Peter Simon Claus Schulze
Phototrophic microalgal cultivation in cold and light-limited environments
ISBN: 978-82-93165-31-6

No. 33 (2019)
PhD in Aquatic Biosciences
Maja Karoline Viddal Hatlebakk
New insights into Calanus glacialis and C. finmarchicus distribution, life histories and physiology in high-latitude seas
ISBN: 978-82-93165-32-3

No. 34 (2019)
PhD in Aquatic Biosciences
Arseny Dubin
Exploration of an anglerfish genome
ISBN: 978-82-93165-33-0

No. 35 (2020)
PhD in Aquatic Biosciences
Florence Chandima Perera Willora Arachchilage
The potential of plant ingredients in diets of juvenile lumpfish (*Cyclopterus lumpus*)
ISBN: 978-82-93165-35-4

No. 36 (2020)
PhD in Aquatic Biosciences
Ioannis Konstantinidis
DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication
ISBN: 978-82-93165-36-1

No. 37 (2021)
PhD in Aquatic Biosciences
Youngjin Park
Transcriptomic and cellular studies on the intestine of Atlantic salmon
Discovering intestinal macrophages using omic tools
ISBN: 978-82-93165-34-7

No. 38 (2021)
PhD in Aquatic Biosciences
Purushothaman Kathiresan
Proteomics of early embryonic development of zebrafish (Danio rerio)
ISBN: 978-82-93165-37-8

No. 39 (2021)
PhD in Aquatic Biosciences
Valentin Kokarev
Macrobenthic communities of sub-Arctic deep fjords: composition, spatial patterns and community assembly
ISBN: 978-82-93165-38-5

No. 40 (2021)
PhD in Aquatic Biosciences
Aurélien Delaval
Population genomics of a critically endangered data-deficient elasmobranch, the blue skate *Dipturus batis*
ISBN: 978-82-93165-39-2

No. 41 (2021)
PhD in Aquatic Biosciences
Isabel Sofía Abihssira García
Environmental impact of microplastics in relation to Atlantic salmon farming
ISBN: 978-82-93165-40-8

No. 42 (2022)
PhD in Aquatic Biosciences
Yousri Abdelmutalab Ahmed Abdelhafiz
Insights into the bacterial communities of Nile tilapia – core members and intergenerational transfer
ISBN: 978-82-93165-41-5

No. 43 (2022)
PhD in Aquatic Biosciences
Fredrik Ribsskog Staven
Interaction studies on lumpfish exposed to Atlantic salmon: behavioural observations and the underlying physiological and neurobiological mechanisms
ISBN: 978-82-93165-42-2

No. 44 (2022)
PhD in Aquatic Biosciences
Solveig Lysfjord Sørensen
Influence of feed ingredients and additives on mucosal health with focus on the intestine of Atlantic salmon (*Salmo salar*)
ISBN: 978-82-93165-43-9

No. 45 (2022)
PhD in Aquatic Biosciences
Apollo Marco Dalonos Lizano
Examining challenges in species-level taxonomy among Calanus copepods in the Northern seas using genome and transcriptome data
ISBN: 978-82-93165-44-6

No. 46 (2022)
PhD in Aquatic Biosciences
Sowmya Ramachandran
Ribosomal RNA, ribose methylation, and box C/D snoRNAs during embryonic development of teleosts zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*)
ISBN: 978-82-93165-45-3

No. 47 (2023)
PhD in Aquatic Biosciences
William John Hatchett
The brown algal genus Fucus: A unique insight into reproduction and the evolution of sex-biased genes
ISBN: 978-82-93165-46-0

No. 48 (2023)
PhD in Aquatic Biosciences
Ying Yen
Application of ensiled *Saccharina latissima* and *Alaria esculenta* as feed: ensilbility, digestibility and bioactivity
ISBN: 978-82-93165-47-7

No. 49 (2023)
PhD in Aquatic Biosciences
Deepak Pandey
Marine macroalgae as an alternative, environment-friendly, and bioactive feeding resource for animals
ISBN: 978-82-93165-48-4

No. 50 (2023)
PhD in Aquatic Biosciences
Nimalan Nadanasabesan
Dietary approaches to improve mucosal health of Atlantic salmon (*Salmo salar*)
ISBN: 978-82-93165-49-1

Many deep-sea fish are diverse and their taxonomy complex and difficult to understand due to a lack of prior knowledge and genomic resources. To survive the demands of their extreme environments, particularly the high hydrostatic pressure and low temperatures, they have undergone a variety of adaptations. In this thesis, genomic-based approaches were harnessed to address knowledge gaps in deep-sea fish evolution of the suborder Cottoidei, and their mutualistic behaviour when used as cleaner fish. For systematic and comparative genomic resources, this study has contributed the complete mitogenome sequence of one deep-sea species, the Atlantic spiny lumpsucker. An examination of selection on mitogenomes of protein-coding genes (PCGs) in 73 species provided no evidence of positive selection on PCGs. Second, in order to explore the nuclear genome, the near-complete genomes of six species of deep-sea fish were sequenced. This permitted construction of a species tree of the suborder Cottoidei, which indicates patterns of positive selection acting on key genes involved in adipogenesis, DNA repair, and translation. These positive selection genes (PSGs) control fundamental biological functions of deep-sea fish. Finally, gene expression analysis was used to investigate the role of noradrenaline in the complex social behaviour of cleaner fish and clients. Overall, this thesis emphasizes the utility of genomic approaches in studies of deep-sea fish taxonomy and systematics, evolution, and mutualistic behaviour.

NORD University

ISBN: 978-82-93165-50-7
Trykk: Trykkeriet, Nord universitet
www.nord.no