Sandra Raposo Costa

Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to Assemble Mitochondrial Genomes

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências e Tecnologia

2022

Sandra Raposo Costa

Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to Assemble Mitochondrial Genomes

Mestrado em Biologia Marinha
Masters in marine biology

Supervisors:
Dr. Márcio A. G. Coelho, Centre of Marine Sciences (CCMAR) University of Algarve, macoelho@ualg.pt

Co-supervisor

Prof. Dr. Ester Serrão, University of Algarve & CCMAR, eserrao@ualg.pt

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências
e Tecnologia 2022

Declaração de autoria de trabalho

Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to Assemble Mitochondrial Genomes

Declaro ser o(a) autor(a) deste trabalho, que é original e inédito.
Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

_____

Sandra Raposo Costa

Resumo

As florestas de corais de zonas mais profundas ou de águas temperadas são caracterizadas por comunidades de elevada diversidade comummente denominadas por “florestas animais” pois são dominadas por espécies filtradoras não fotossintéticas que formam habitats tridimensionalmente complexos. Octocorallia é uma vasta e diversificada subclasse de antozoários que incluem os corais moles e gorgónias (Alcyonacea), os corais azuis (Helioporacea) e as penas marinhas (Pennatulacea). Estas espécies também apresentam um grande valor ecológico pois atuam como engenheiros do ecossistema com a capacidade de manter um habitat fornecendo abrigo e alimento para várias espécies de peixe, incluindo espécies com elevado interesse para a pesca comercial. Infelizmente os impactos antropogénicos persistem devido ao crescimento populacional exponencial e ao desenvolvimento da indústria global criando uma enorme pressão sobre estes ecossistemas marinhos. Felizmente, o nosso conhecimento tem crescido acerca da gravidade de como os impactos estão a afetar a integridade dos ecossistemas em zonas mais profundas. Tais impactos antropogénicos incluem poluição como derrames de petróleo, acidificação dos oceanos e atividade intensas de pesca. Corais, como as gorgónias, têm um crescimento lento e alta longevidade tornando-os mais vulneráveis a distúrbios no ecossistema pois apresentam uma baixa taxa de recuperação natural. Devido à sua importância ecológica, existe um grande interesse na comunidade científica de desenvolver estudos e implementar medidas de conservação e restauração destes habitats. A perda da biodiversidade, causada por impactos antropogénicos, tem sido um motivo preocupante globalmente. O desenvolvimento de ferramentas genéticas inovadores que facilitam na identificação de espécies, especialmente de espécies que são morfologicamente muito semelhantes, mas geneticamente diferentes, têm vindo a complementar os métodos de monitoração de biodiversidade existentes. A next-generation sequencing (NGS) é

uma plataforma recente de sequenciação que fornece elevadas leituras de sequenciação numa única execução, pois fornecem informação rápida e massiva por sequenciação com esforços e custos mínimos em comparação com as técnicas tradicionais de sequenciação de Sanger. Esta técnica promove o aperfeiçoamento e complementação dos métodos tradicionais, mas apresenta algumas limitações. O primeiro indica que a região que é amplificada por PCR e sequenciada está limitada a um tamanho pequeno do genoma mitocondrial total como um fragmento do COI (marcador genético tradicional). O segundo indica que, para alguns grupos de invertebrados, estes marcadores genéticos mitocondriais não possuem polimorfismo suficiente para distinguir espécies próximas. Por fim, o *barcoding* de ADN requer uma biblioteca de referência completa e com qualidade. Estas bibliotecas contêm dados genéticos que são fundamentais para estudos ecológicos e também são ferramentas valiosas para a avaliação da biodiversidade como "eDNA metabarcoding" que posteriormente podem ser incorporados em estratégias de gestão e conservação. Alguns grupos de corais exigem um sistema de classificação baseado em informação genética que complemente a classificação através de morfológica. A utilização de genomas mitocondriais completos também tem sido uma ferramenta emergente e eficaz em taxonomia e estudos filogenéticos, pois são bastante úteis para avaliar relações filogenéticas ao nível de espécies porque têm uma taxa evolutiva muito mais rápida do que o genoma nuclear. Pois requer mais que um gene, ou preferencialmente, o mitogenoma completo para resolver relações evolutivas entre espécies. O genoma mitocondrial em antozoários, em particular na subclasse Octocorallia, exibe uma taxa de evolução molecular muito lenta em comparação com outros grupos taxonómicos. Como tal, o *barcoding* de ADN através da sequenciação de apenas um gene mitocondrial (normalmente o tradicional COI) é inadequado para distinguir muitas espécies de corais. A Oxford Nanopore Techniques (ONT) é uma técnica de sequenciação de terceira geração capaz de produzir longas,

fáceis e rápidas leituras de sequenciação. Uma tecnologia ideal para a construção de mitogenomas devido à sua capacidade de realizar leituras longas e repetitivas de sequenciação com uma contiguidade muito maior que garante a integridade da informação genética e construção completa do mitogenoma. Enquanto que os métodos tradicionais de sequenciação de mitogenomas completo, como o primer-walk (i.e., sequenciação de Sanger) sequenciam apenas um único fragmento de ADN de cada vez a um custo muito mais elevado. Para além disso, os octocorais exibem cinco ordens de genes diferentes em que podem ter arranjos de ordem genética diferentes ou existe um bloco de cinco ou mais genes que foram invertidos, ou seja, a codificação da cadeia é revertida. O objetivo geral desta tese foi desenvolver uma abordagem de sequenciação, construção e anotação de mitogenomas completos de diferentes grupos de corais com aplicação universal e que permita sequenciação em larga escala. O estudo é focado em espécies de octocorais maioritariamente presentes na costa portuguesa e alguns representantes de Scleractinians de Cabo Verde e de Espanha. Como tal, os objetivos específicos foram: 1) desenvolver *primers* específicos para corais para amplificar o mitogenoma completo usando dados de referências que estão disponíveis publicamente; 2) expandir as bibliotecas de referencia disponíveis através de sequenciação, *assembly* e anotação dos mitogenomas de espécies que existem em Portugal; 3) confirmar as ordens de genes das espécies-alvo e ver se estão em conformidade com o que foi descrito na literatura; 4) Fazer uma reconstrução filogenética com base nos dados da sequenciação do mitogenoma para inferir a localização filogenética e a afinidade genética das espécies-alvo dentro de Octocorallia, incluindo espécies que não foram identificadas de forma conclusiva com base na morfologia. A abordagem desenvolvida consistiu em desenhar primers universais que permitiram a amplificação do mitogenoma por PCR de longo alcance (produtos de PCR > 8000 pb), seguido de sequenciação com tecnologia de terceira geração, ou seja, dados de “long reads”

obtidos com Oxford Nanopore Technologies. Uma vez que o genoma mitocondrial é circular, procurou-se desenhar primers em duas regiões opostas e equidistantes de maneira a amplificar o mitogenoma em duas reações de PCR, uma para cada metade da molécula. Um total de 17 mitogenomas de octocorais foram construídos com sucesso através de *de novo assembly*. Os mitogenomas circulares codificam 14 genes codificadores de proteínas (*Nad1-6, Nad4L, cox1-3, Cytb, mtMutS, Atp6 e Atp8*), dois genes de RNA ribossómico (*r12S e r16S*) e um RNA de transferência (*trnM*). Das cinco ordens de genes existentes em octocorais, foram sequenciadas espécies que representavam três tipos de organização do mitogenoma: *Isidella elongata* exibindo a ordem de genes B, *Corallium rubrum* exibindo a ordem de genes C e as restantes espécies exibiram a ordem de genes A, embora não tenhamos conseguido reconstruir o mitogenoma da *Isidella elongata* devido a um erro na combinação de primers durante a amplificação. A reconstrução filogenética com base em 16 genes mitocondriais permitiu inferir a localização filogenética e a afinidade genética das espécies-alvo dentro de Octocorallia, incluindo espécies que não foram identificadas de forma conclusiva com base na morfologia. Posicionamos e identificamos, pelo menos ao nível de genérico, todos os 17 mitogenomas reconstruídos na árvore filogenética de Octocorallia e recuperamos três ramos principais que correspondem a estudos filogenéticos anteriores e um ramo que aparenta ser basal ao ramo *Holaxonia-Alcyoniina*. Esta abordagem com sequenciação de terceira geração com a Oxford Nanopore Technology permitiu a sequenciação de “long-read” fornecendo sequenciação em larga escala de muitas amostras numa única execução de sequenciação. Essa abordagem que inclui PCR de longo alcance e sequenciação de “long-read” poderá ter a desvantagem de introduzir erros durante a amplificação (PCR), mas com elevada cobertura de sequenciação é possível escapar a esta limitação. A amplificação por PCR aumenta o número de cópias do mitogenoma do extrato de ADN purificado em multiplex

juntamente com várias amostras agrupadas com *barcoding*. Esse processo reduz a proporção de ADN genómico (não mitocondrial) nos extratos de ADN e permite sequenciar de uma maneira bastante mais direcionada no qual permite agrupar várias amostras na mesmo processo de sequenciação. Esta abordagem possibilitou construir bibliotecas de referência completas cobrindo todos os mitogenomas, no qual melhorou assim as bibliotecas de referência disponíveis que posteriormente poderão auxiliar na avaliação da biodiversidade de corais com base no metabarcoding de eDNA (ADN ambiental). Mas o mais importante foi que esta abordagem pode ser altamente aplicável para sequenciar o mitogenoma completo de amostras ambientais (eDNA) por meio de metabarcoding. Desta forma, esta abordagem, juntamente com o eDNA, é valiosa para catalogar a biodiversidade e auxiliar na avaliação da biodiversidade e possíveis medidas de conservação.

Palavras-chave: Octocorallia, genomas mitocondriais, sequenciação com Oxford Nanopore Technologies, barcoding de ADN, organização de genes mitocondriais.

Abstract

Coral communities found either in tropical or deeper locations or in temperate waters can be classified as “marine animal forests” as they are dominated by habitat-forming suspension feeders, creating three-dimensional forest-like structures. Octocorallia is a wide and diverse subclass of Anthozoa that includes soft corals and gorgonians (Alcyonacea), blue corals (Helioporacea), and sea feathers (Pennatulacea). Their ecological importance, slow growth, and susceptibility to degradation caused by anthropogenic impacts make them vulnerable marine ecosystems. This leads to an interest of the scientific community to develop studies and implement conservation and restoration measures. Yet for many of these organisms, their identity is uncertain or debatable. The development of genetic tools that facilitate the identification of species, especially species that are morphologically identical but genetically different, has been complementing existing biodiversity monitoring methods. DNA barcoding allows high throughput multispecies identification but using sequencing of just one mitochondrial gene (typically the COI) is inadequate to distinguish many coral species. This study aims to develop an approach to barcode the mitochondrial genomes of octocorals by coupling long-range PCR with a 3rd generation sequencing platform (i.e., long-read sequencing) by designing coral-specific primers to amplify the mitogenomes, expanding the available reference library, confirming the gene orders arrangements of the target species and their placement in the Octocorallia phylogeny tree. We successfully identified and placed, at least at the genus level, all 17 reconstructed mitogenomes in the Octocorallia phylogenetic tree and we were able to identify three of the five existing gene orders within octocorals. This approach complemented and expanded the reference libraries that are applicable for eDNA metabarcoding to catalog biodiversity and assist in biodiversity assessment and possible conservation measures.

Keywords: Coral biodiversity, Octocorallia, DNA barcoding, Mitochondrial genomes, Third generation sequencing.

Acknowledgments

I would like to kindly thank my mentor and supervisor Dr. Márcio Coelho, for all his knowledge, dedication, and support. His support was fundamental for the development of this thesis, and I am very grateful for all his support and dedication.

I would like also to acknowledge Dr. Aschwin Engelen for all his knowledge and guidance and for providing scleractinian samples for this work. I would like to thank Dr. Ester Serrão for her fundamental comments during writing revision. Finally, I would like to kindly thank Marta Valente for all her help and availability during the sequencing periods

This research was funded by project HABMAR (Grant No. MAR-01.04.02-FEAMP-0018) co-financed by the European Maritime and Fisheries Fund of the Operational Program MAR 2020 for Portugal (Portugal 2020). Furthermore, this study received Portuguese national funds from FCT—Foundation for Science and Technology through projects UIDB/04326/2020, as well as samples from project RESTORSEAS (Biodiversa/ DivRestore/0013/2020) and FCT AGA-KHAN / 540316524 / 2019.

I would like to also give special thanks to my family, especially my husband, for their continuous support and constant encouragement throughout this period.

Table of contents

Index of figures

Chapter 1: Introduction

Chapter 2: Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to assemble Mitochondrial Genomes

Index of tables

Chapter 2: Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to assemble Mitochondrial Genomes

1. Chapter 1: Introduction

1.1. Ecological and biological aspects of coral communities

Deep sea locations around the world are poorly understood and largely unexplored (Costello et al., 2010; Oliveira et al., 2021). It was not until recently that technological advancements opened new doors to underwater exploitation that renewed interest by the scientific community (Roberts et al., 2009; Oliveira et al., 2021). Deep sea studies have revealed a huge variety of habitats, including coral-dominated communities that share many similarities with shallow-water tropical coral reefs (Rogers, 1999; Buhl-Mortensen and Mortensen, 2004; Henry and Roberts, 2016). Coral habitats are generally known to be complex, containing a high diversity of invertebrates and fish. Corals are ecosystem engineers that provide a wide array of ecosystem services and goods that fall under a few main categories (e.g., provisioning, regulating, cultural, and supporting services) that are connected to the well-being of millions of people (Woodhead et al., 2019). Their complex tridimensional morphology that builds habitats directly benefits associated organisms by also providing a great set of services such as structure, food, shelter, and nursing grounds to many different species (Buhl-Mortensen et al., 2009). This includes species important for commercial fisheries, which are in popular demand for human consumption (Roberts and Hirshfield, 2004). Corals belong to the Cnidarian Phylum, which encompasses several taxonomic groups, including stony corals (order Scleractinia); soft corals, sea pens and gorgonians (subclass Octocorallia); black corals (order Antipatharia); zoanthids (order Zoantharia) (Carreiro-Silva et al., 2017) and hydrocorals (family Stylasteridae) (Roberts et al., 2006; Sundahl et al., 2020).

In general, the great majority of tropical corals are located in regions with strong light penetration, they are light-dependent (i.e., autotrophic) due to their reliance on a symbiotic relationship with the dinoflagellate *Symbiodinium spp*., for which they are responsible for

providing most of the energy that the corals need to meet their metabolic demands to survive in these shallow oligotrophic ecosystems (Roth, 2014). Therefore, the distribution of tropical corals is confined by light (Rossi et al., 2017). In contrast, corals from temperate and cold-water habitats, including deep communities, are mostly heterotrophic and do not depend on sunlight energy so they can occur at deeper habitats, down to 6 km below the surface of the oceans (Roberts and Hirshfield, 2004). These habitats can be composed of single isolated colonies, or they can form large carbonate mounds that can extend up to hundreds of meters (Roberts et al., 2006). They can be found in the deep sea around the world in continental shelves, slopes, seamounts, and ridges. Locations with specific hydrodynamics, characterized by strong currents, enable the success of settlement and food supply allowing the corals to successfully grow (Patterson, 1984; Sponaugle and LaBarbera, 1991; Roberts et al., 2006). Either in the depths of the ocean, in temperate waters, or tropical regions coral habitats hold areas of great biodiversity resulting in high heterogeneity with different colors, shapes, and sizes (Roberts and Hirshfield, 2004; Rossi et al., 2017).

In recent years, the term “marine animal forests” has been introduced to describe marine mega-benthic communities formed by suspension feeders such as sponges, bivalves, and corals, communities with three-dimensional structures. This term highlights the structural and functional resemblance between these marine ecosystems and terrestrial forests, which are dominated by animals instead of plants (Rossi et al., 2017; Dias et al., 2020). There are many marine animal forests represented by octocorals around the world (Rossi et al., 2017). Octocorallia is a wide diverse Subclass of Anthozoa named after its octoradial symmetry due to its 8 tentacle polyps that include soft corals and gorgonians (Alcyonacea), blue corals (Helioporacea), and sea pens (Pennatulacea) (Cairns et al., 2009; Altura e Poliseno, 2019; Oliveira et al., 2021). Just like trees, octocorals are generally slow-growing and live through long periods of time forming dense forests

along rocky grounds such as cliffs, outcrops, and biogenic substrates (Ballesteros, 2006; Palma et al., 2018) they can be found across a wide range of depths between 15m- 6000m (Grasshoff, 1981; Gori et al., 2017; Palma et al., 2018). Here, we will classify the corals used in this study as representatives of marine animal forests.

1.2. Impacts and vulnerability of deep coral communities

Anthropogenic impacts on marine ecosystems have persisted due to the exponential population growth of humans and the development of the global industry since industrial times. Global climate change caused by increasing anthropogenic greenhouse gases due to, mostly, the rising CO_2 concentrations in the atmosphere has led to a decrease in the ocean's pH, turning the oceans more acidic and leading to an increase in ocean temperatures (IPCC, 2019). Coral-dominated habitats can be subject to human impacts directly, such as destructive fishing activities (Ragnarsson et al., 2017), and indirectly, such as coral bleaching caused by the rising ocean temperature (Wilkinson, 2000). The main indirect impacts include a wide range of pollution (Weinnig et al., 2020), fishing activities such as bottom trawling (Clark et al., 2016; Ragnarsson et al., 2017), deep-sea mining (Dover et al., 2017), oil and gas operations (Cordes et al., 2016), among others related to climate change. The life-history traits of corals, such as gorgonians, include slow growers, high longevity, low recruitment rates, and high post-settlement mortality which makes them vulnerable to any of these impacts, especially to destructive fishing activities such as bottom trawling (Andrews et al., 2002; Lacharité and Metaxas, 2013). These impacts are compromising the long-term integrity and functioning of these ecosystems which highlights the need for conservation and restoration measures.

Oil operations have been increasing lately causing a greater risk of accidental release of oil spills leading to lethal damages to the environment and directly affecting deep-sea locations (Cordes et al., 2016; Weinnig et al., 2020). Hydrocarbon operations can temporarily contaminate the deep sea with discharges of drill cuttings and drill muds, operations that are always at risk of accidental spills directly into the ocean. The *Deepwater Horizon* (DWH) oil spill was so far the largest accidental oil spill in history that caused sublethal and lethal impacts to various marine species including deep-sea coral communities (Dubansky et al., 2013; DeLeo et al., 2018). A recent study was able to provide evidence of the cellular stress responses to oil of deep-sea corals by examining changes in their gene expression and demonstrated the utility of next-generation sequencing for monitoring anthropogenic impacts in deep waters by finding that genes involved in immunity and regenerate responses to stressors were expressed differently in impacted corals stress that were subject to oil exposure. (DeLeo et al., 2018). Local and deep-sea coral communities, who are exposed to these contaminants and drill cuttings, are induced to stress affecting their behavior, fitness, and survival (Ragnarsson et al., 2017).

Bottom fishing operations such as bottom trawling along with bottom set longlines and gillnets are known to dramatically impact deep ecosystems. These ecosystems hold many commercially important fish species, thereby attracting bottom fishing activities, especially trawling, which can cause high impacts on marine animal forests. In extreme cases, trawlers can destroy the reefs with the intention of capturing only a few commercial fish species (Armstrong and Hove, 2008; Ragnarsson et al., 2017). Recent studies on the Portuguese coast also suggested that artisanal fisheries using bottom-set gillnets generate a substantial amount of coral bycatch that can threaten the ecological integrity and functioning of circalittoral coral gardens, (Dias et al., 2020). All these studies indicate that there is an urgent need to take action and apply appropriate conservation

measures to protect these slow-growing and fragile organisms (Armstrong and Hove, 2008; Ragnarsson et al., 2017; Dias et al., 2020). In some locations, off the coast of Ireland, bottom trawling has been banned as a fishing utility after scientists exposed video footage of the damaged trawled reef to the authorities revealing the severity of the impacts of trawling (Grehan et al., 2004; Ragnarsson et al., 2016). Although several measures have already been taken to protect these reefs, there is still much work to be done especially in providing evidence to take appropriate conservation actions.

1.3. Genetic tools for species identification

The loss of biodiversity has been a matter of concern in the scientific community worldwide (Thomsen and Willerslev, 2015). The impacts of human activities have been aggravating and leaving an increasing mark on marine (and terrestrial) ecosystems posing one of the major challenges of the century. Yet, our current knowledge about the biodiversity of corals reefs is still limited, even for shallow reefs that are much more accessible to investigate compared to deep-sea corals (McFadden et al., 2014). Therefore, the need to expand our knowledge about the existing biodiversity has been increasing, and this was previously done by surveying and monitoring species in their natural environment. But this is not an easy task, it requires expertise and time (Creer et al., 2016; Hopkins and Freckleton, 2002), in some cases, it can be intrusive and invasive, and the identification of closely related species can be difficult, especially the ones that share identical morphological traits (Thomsen and Willerslev, 2015). It is common, especially in deeper locations, to misinterpret the true identity of corals due to their morphological similarities. Species misidentification can lead to biased assessments of biodiversity, thus influencing the research fields that rely on species as units of analysis (Oury et al., 2020).

Innovative genetic tools open new doors to opportunities to find cryptic species, species that are morphologically identical but genetically different (Appeltans et al., 2012), to supplement existing biodiversity monitoring methods (Holman et al.,2019). DNA barcoding allows high-throughput species identification, the DNA from the entire community across taxonomic groups can be analyzed at the same time using a single standardized DNA sample (Thomsen and Willerslev, 2015; Taberlet et al., 2012). The analysis process usually starts with DNA amplification using *Polymerase Chain Reaction* (PCR) with species-specific primers if focused on a single-species or with generic primers if using the multiple species approach followed by DNA sequencing (Thomsen and Willerslev, 2015). Next-generation sequencing (NGS) is a sequencing platform that generates millions of sequence reads in a single run with at least five orders of magnitude of improved reads compared to traditional Sanger sequencing (Taberlet et al., 2012). The NGS platforms can obtain thousands of sequence reads per amplicon providing fast and massive information per sequencing run for limited effort and cost (Taberlet et al., 2012; Thomsen and Willerslev, 2015). This technique promotes the improvement and complementation of working methods along with traditional methods. But this barcoding approach contains some limitations. The first indicates that the taxonomic resolution of the marker genes such as COI, 12S ribosomal RNA (rRNA), 16S rRNA, and CytB is limited to a small size of the overall mitochondrial genome(Schroeter et al., 2020; Taberlet et al., 2012). The second indicates that for some invertebrate groups these mitochondrial gene markers do not have enough polymorphism to distinguish closely related species (McFadden et al.,2010; Hebert et al., 2003). For example, the COI barcode is virtually invariant across all Mediterranean species of gorgonians of the genus *Eunicella* (Calderón et al., 2006). Finally, DNA barcoding requires a high-quality reference database/library. It is a taxonomic library containing the nucleotide sequences of the target sample

(Taberlet et al., 2012). These genetic databases are crucial in ecological studies and a valuable tool in biodiversity assessment as they can be used in DNA barcoding approaches if containing the desired sequences, therefore it requires a continuation in the development of these genetic databases, which later can be incorporated into management strategies (Schroeter et al., 2020). This third-generation sequencing approach with Oxford Nanopore Technology enabled long-read sequencing by providing large-scale sequencing of many samples in a single sequencing run and ensuring the presence of complete genetic information (Baeza, 2020).

1.4. Gene order rearrangements of octocorals

Some groups of corals require a more reliable classification system based on genetic information to complement the morphological approach (Iwasaki and Suzuki, 2010; Uda et al., 2011). It is common to use mitochondrial gene sequences to assess phylogenetic relationships at the species level because it has a much faster evolutionary rate than the nuclear genome (Wolstenholme, 1992; Uda et al., 2011). In general, the animal mitochondrial genome is composed of one double-stranded circular DNA that encodes 14 protein-coding genes, two ribosomal RNAs, and 22 transfer RNAs. It also contains one major non-coding region, the control region, believed to be used in the initiation of transcription and replication (Uda et al., 2013; Wolstenholme, 1992). The rearrangement of gene order within mitochondrial genomes is relatively uncommon, and when the rearrangement is established, usually remains unchanged over long periods of evolutionary time (Uda et al., 2013). Consequently, it is common to use the information on gene order arrangement to construct phylogenies in major metazoan groups (Uda et al., 2013; Boore, 1999).

Recent studies suggest that using the whole mitochondrial genome for phylogenetic analysis provides better support of the tree topology (Figueroa and Baco, 2014). They also revealed five

different gene orders in octocorals, one gene order is shared by most octocorals and is present in all clades of octocorals while the other four gene arrangements are only found within the Calcaxonia-Pennatulacea clade and the Anthomastus-Corallium clade of Octocorallia (McFadden et al. 2006; Figueroa and Baco, 2014; Hogan et al., 2019) (Figure 1.1). To date, all octocorals possess a unique protein-coding gene, the DNA mismatch-repair *mtMutS* that is not present in any other metazoan mitochondrial genome (McFadden et al., 2010; Brockman and McFadden, 2012). The presence of the gene in the mitogenome can be explained by a horizontal gene transfer event due to the slow rates of mitochondrial evolution in octocoral and also because the *mtMutS* codes for a homolog functional DNA repair mechanism, which requires double-stranded breakage, followed by repair of the DNA molecule (Bilewitch and Degnan, 2011; Brockman and McFadden, 2012). The functionality of the gene explains the low polymorphism of the mitogenome in octocorals (compared to hexacorals) because it confers the repair of errors in DNA replication. (Bilewitch and Degnan, 2011).

Initially, scientists assumed that octocorals exhibited no gene rearrangements (Brockman and McFadden, 2012). But several recent studies have proven otherwise by documenting five alternative gene orders: the most common gene order, referred to as gene order A (Beagley et al., 1995; Brockman and McFadden, 2012), is considered as the "ancestral gene order"; gene order B is found in the bamboo corals in the Isididae Family (Brugler and France 2008); gene order C and D correspond to the precious corals *Paracorallium japonicum* and *Corallium konojoi* respectively from the Coralliidae Family (Uda et al., 2011); gene order E belongs to *Paraminabea aldersladei* (Alcyoniidae Family) (Brockman and McFadden 2012); and gene order F belongs to *Isidoides armata* (Isididae Family) (Pante et al., 2013) (Figure 1.1). A block of five or more genes of these alternative gene arrangements has been inverted, resulting in a reverse coding strand (Brockman

and McFadden, 2012; Uda et al., 2011). Mitochondrial gene inversions are frequently observed in invertebrates (Uda et al., 2013; Dowton et al., 2003). It has been proposed that the gene order inversion in octocorals happens due to intramitochondrial recombination and rejoining of the double-stranded DNA (Uda et al., 2013; Brugler and France, 2008; Dowton and Austin, 1999). But the story is different for their sister clade, Hexacorallia. Their evolving mechanisms are very different from Octocorallia, exhibiting no gene inversions but instead, they exhibit extreme shuffling of genes among orders, and all genes are encoded on the same strand. Furthermore, the gene order in the mt genomes differs radically between the five orders of Hexacorallia (Actiniaria, Antipatharia, Corallimorpharia, Zoantharia, and Scleractinia) (Brockman and McFadden, 2012).

Figure 1.1. Gene order arrangements found in the mitochondrial genome of Octocorallia (linearized view). Gene order A most common gene order in Octocorallia (Octocoral ancestral gene order); gene order B for Isididae Family (bamboo corals); gene order C and D correspond to two precious corals from the Coralliidae Family (*Paracorallium japonicum* and *Corallium konojoi* respectively); gene order E represents *Paraminabea aldersladei* from the Alcyoniidae Family. Different colors correspond to the four octocoral conserved gene blocks (block 1, blue: *cox1-r12S-Nad1-Cytb*; block 2, orange: *Nad6-Nad3-Nad4L*; block 3, green*: MutS-r16S-Nad2-Nad5-Nad4*; block 4, purple*: trnM-cox3-Atp6-Atp8-cox2*). Arrows indicate the direction of replication. Adapted from Hogan et al., 2019.

1.5. Research objectives

This study aims to contribute to the improvement of molecular analysis for species identification, i.e., DNA barcoding, an alternative to traditional taxonomy. Our barcoding approach increases the resolution to identify the species, therefore there is a greater genetic resolution of the library because we go from using only a small fragment of the COI or any other gene (traditional barcoding) to the complete mitogenome (3rd generation sequencing). 3rd generation NGS platforms (i.e., long-read sequencing like Nanopore) have a huge potential because while the sequencing error is much smaller, we are working on a genomic scale (instead of at the gene level) to identify species.

The overall goal of this research was to develop a pipeline to barcode the mitochondrial genomes of octocorals by coupling long-range PCR with a 3rd generation sequencing platform (i.e., long-read sequencing). The study focused on octocoral species occurring on the coast of Portugal, for which genetic resources are lacking, particularly circalittoral and deep-sea species. However, the methodology developed here aims to be transversal to all octocorals. The specific research objectives were to:

1. Develop coral-specific primers to amplify the mitogenome using reference sequence data available in public databases.
2. Expand the available reference libraries by sequencing, assembling, and annotating the mitogenomes of species that occur in Portugal.
3. Confirm the gene order arrangements of the target species and see if they are in conformance with what has been described for the families to which they belong.

4. Do a phylogenetic reconstruction based on mitogenome sequence data to infer the phylogenetic placement and genetic affinity of the target species within Octocorallia, including species that were not conclusively identified based on morphology.

1.6. References

Altuna, A., Poliseno, A. (2019). 14. Taxonomy, Genetics and Biodiversity of Mediterranean Deep-Sea Corals and Cold-Water Corals. In Mediterranean Cold-Water Corals: Past, Present and Future (pp. 121-156). Springer, Cham.

Andrews, A., Corders, E., Mahoney, M., Munk, K., Coale, K., Cailliet, G., & Heifetz, J. (2002). Age, growth and radiometric age validation of a deep-sea, habitat-forming gorgonian (Primnoa resedaeformis) from the Gulf of Alaska. *Hydrobiologia*, *471*, 101-110. doi:10.1023/A:1016501320206

Appeltans, W. et al. (2012). The magnitude of global marine species diversity. *Current Biology, 22*, 2189–2202. doi:10.1016/j.cub.2012.09.036

Armstrong, C. W., & Hove, S. (2008). The formation of policy for protection of cold-water coral off the coast of Norway. *Marine Policy, 32*, 66-73. doi:10.1016/j.marpol.2007.04.007

Ballesteros, E. (2006). Mediterranean coralligenous assemblages: a synthesis of present knowledge. *Oceanogr. Mar. Biol*. *44*: 123-195.

Barrett, N. J., Hogan, R. I., Allcock, A., Molodtsova, T., Hopkins, K., Wheeler, A. J., & Yesson, C. (2020). Phylogenetics and Mitogenome Organisation in Black Corals (Anthozoa: Hexacorallia: Antipatharia): An Order-Wide Survey Inferred From Complete Mitochondrial Genomes. *Frontiers in Marine Science, 7*, 440. doi:10.3389/fmars.2020.00440

Beagley, C. T., et al. (1995). Mitochondrial genomes of Anthozoa (Cnidaria). In: Palmieri F, editor. Progress in cell research. Vol. 5. Amsterdam: Elsevier Science. p. 149–153.

Bilewitch, J. P., & Degnan, S. M. (2011). A unique horizontal gene transfer event has provided the octocoral mitochondrial genome with an active mismatch repair gene that has potential for an unusual self-contained function. *BMC evolutionary biology*, *11*, 228. doi:10.1186/1471-2148-11-228

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767-1780. doi:10.1093/nar/27.8.1767

Brockman, S. A., & McFadden, C. S. (2012). The mitochondrial genome of Paraminabea aldersladei (Cnidaria: Anthozoa: Octocorallia) supports intramolecular recombination as the primary mechanism of gene rearrangement in octocoral mitochondrial genomes. g*enome biology and evolution, 4*(9), 994-1006. doi:10.1093/gbe/evs074

Brugler, M. R., & France, S. C. (2008). The Mitochondrial Genome of a Deep-Sea Bamboo Coral (Cnidaria, Anthozoa, Octocorallia, Isididae): Genome Structure and Putative Origins of Replication Are Not Conserved Among Octocorals. *Journal of Molecular Evolution, 67*, 125. doi:0.1007/s00239-008-9116-2

Buhl-Mortensen, L., Mortensen, P. B. (2004). Symbiosis in deep-water corals. *Symbiosis, 37(1),* 33-61.

Buhl-Mortensen, L. V.-M. (2010). Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. *Marine Ecology, 31*, 21-50. doi:https://doi.org/10.1111/j.1439-0485.2010.00359.x

Buhl-Mortensen, P., Buhl-Mortensen, L., & Purser, A. (2016). Trophic Ecology and Habitat Provision in Cold-Water Ecosystems. *Springer International Publishing*, 919-944. doi:10.1007/978-3-319-21012-4_20

Cairns, S. D., & Bayer, F. M. (2005). A review of the genus Primnoa (Octocorallia: Gorgonacea: Primnoidae), with the description of two new species. *Bulletin of Marine Science, 77*(2), 225-256.

Calderón, I., Garrabou, J., & Aurelle, D. (2006). Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and Ecology*, 336(2), 184-197. doi:10.1016/j.jembe.2006.05.006

Carreiro-Silva, M., Ocaña, O., Stanković , D., Sampaio, Í., Porteiro, F. M., Marie-Claire, F., & Stefanni, S. (2017). Zoantharians (Hexacorallia: Zoantharia) Associated with Cold-Water Corals in the Azores Region: New Species and Associations in the Deep Sea. *Frontiers in Marine Science, 4*, 88. doi:10.3389/fmars.2017.00088

Cathalot, C., Oevelen, D. V., Cox, T. J., Kutti, T., Lavaleye, M., Duineveld, G., & Meysman, F. J. (2015). Cold-water coral reefs and adjacent sponge grounds: hotspots of benthic respiration and organic carbon cycling in the deep sea. *Frontiers in Marine Science, 2*, 37. doi:10.3389/fmars.2015.00037

Clark, M. R., Althaus, F., Schlacher, T. A., Williams, A., Bowden, D. A., and Rowden, A. A. (2016). The impacts of deep-sea fisheries on benthic communities: a review. *ICES J. Mar. Sci.* 73, i51–i69. doi: 10.1093/icesjms/fsv123

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., . . . Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution, 7*(9), 2041-210X. doi:10.1111/2041-210X.12574

Cordes, E. E., Jones, D. O., Schlacher, T., Amon, D. A., Bernardino, A. F., Brooke, S., Carney, R., DeLeo, D. M., Dunlop, K. M., Escobar-Briones, E. G., Gates, A. R., Genio, L., Gobin, J., Henry, L.-A., Herrera, S., Hoyt, S., Joye, S., Kark, S., Mestre, N. C., Metaxas, A., Pfeifer, S., Sink, K., Sweetman, A. K., & Witte, U. F. (2016). Environmental impacts of the deep-water oil and gas industry: a review to guide management strategies**.** *Front. Environ. Sci., 4*, p. 58.

Costello, M. J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H., et al. (2010) A Census of Marine Biodiversity Knowledge, Resources, and Future Challenges. *Plos One*, *5(8)*, e12110. doi:10.1371/journal.pone.0012110

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., & Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*, 1-11. doi:10.1111/2041-210X.12836

DeLeo, D. M., Herrera, S., Lengyel, S. D., Quattrini, A. M., Kulathinal, R. J., Cordes, E. E. (2018). Gene expression profiling reveals deep-sea coral response to the Deepwater Horizon oil spill. *Molecular Ecology*. 27, 4066– 4077. doi:10.1111/mec.14847

Dover, C., Ardron, J., Escobar, E., Gianni, M., Gjerde, K., Jaeckel, A., Jones, D., Levin, L., Niner, H., Pendleton, L., Smith, C., Thiele, T., Turner, P., Watling, L., & Weaver, P. (2017).

Biodiversity loss from deep-sea mining. *Nature Geoscienc*e. 10. 464–465. doi:10.1038/ngeo2983.

Dowton, M., & Austin, A. D. (1999). Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Molecular Biology and Evolution, 16*(2), 298-309. doi:10.1093/oxfordjournals.molbev.a026111

Dubansky, B., Whitehead, A., Miller, J. T., Rice, C. D., & Galvez, F. (2013). Multitissue molecular, genomic, and developmental effects of the deepwater horizon oil spill on resident Gulf Killifish (*Fundulus grandis*). *Environmental Science and Technology, 47*, 5074-5082. doi:10.1021/es400458p

Etnoyer P., Morgan L.E. (2005) Habitat-forming deep-sea corals in the Northeast Pacific Ocean. In: Freiwald A., Roberts J.M. (eds) Cold-Water Corals and Ecosystems. Erlangen Earth Conference Series. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-. (s.d.).

Figueroa, D. F., & Baco, A. R. (2014). Complete mitochondrial genomes elucidate phylogenetic relationships of the deep-sea octocoral families Coralliidae and Paragorgiidae. *Deep Sea Research Part II: Topical Studies in Oceanography, 99*, 83-91. doi:10.1016/j.dsr2.2013.06.001

Grasshoff, M. (1981). Die Gorgonaria, Pennatularia und Antipatharia des Tiefwassers der Biskaya (Cnidaria, Anthozoa). Ergebnisse der franzosischen Expeditionen Biogas, Polygas, Géomanche, Incal, Noratlante und Fahrten der Thalassa 1. Allgemeiner Teil. *Bulletin du Museum National d'Histoire Naturelle, 4 (3(A4))*,732–766.

Gori, A., Bavestrello, G., Grinyó, J., Dominguez-Carrió, C., Ambroso, S., Bo, M. (2017). Animal forests in deep coastal bottoms and continental shelf of the Mediterranean Sea. *Marine animal forests: The ecology of benthic biodiversity hotspots*, 1-27.

Grehan, A., Unnithan, V., Wheeler, A., Monteys, X., Beck, T., Wilson, M., Guinan, J., Foubert, A., Klages, M., & Thiede, J. (2004). Evidence of major fisheries impact on cold-water corals in the deep waters off the Porcupine bank west coast of Ireland: are interim management measures required? *ICES Annual Report CM*, 1-9.

Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Biological Sciences, 270*, S96–S99. doi:10.1098/rsbl.2003.0

Henry L. A. & Roberts J.M. (2017) Global Biodiversity in Cold-Water Coral Reef Ecosystems. In: Rossi S., Bramanti L., Gori A., Orejas C. (eds) Marine Animal Forests. Springer, Cham. doi:10.1007/978-3-319-21012-4_6

Hogan, R. I., Hopkins, K., Wheeler, A. J., Allcock, L. A., & Yesson, C. (2019). Novel diversity in mitochondrial genomes of deep-sea Pennatulacea (Cnidaria: Anthozoa: Octocorallia). *Mitochondrial DNA Part A, 30(6)*, 764-777. doi: 10.1080/24701394.2019.1634699

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2019). Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. *Scientific Reports, 9*, 1– 11. doi:10.1038/s41598-019-47899-7

Hopkins, G. W., & Freckleton, R. P. (2002). Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation*, 5(3), 245-249. doi:10.1017/S1367943002002299

IPCC, 2019: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press.

Iwasaki, N., & Suzuki, T. (2010). Biology of precious coral. Em N. Iwasaki, & T. Suzuki, *A Biohistory of Precious Coral* (pp. 3-25). Tokyo: Tokai University Press.

Jones, C. G., Lawton, J. H., Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos*, 69: 373-386.

Kayal, E., Roure, B., Philippe, H., Collins, A. G., & Lavrov, D. (2013). Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC evolutionary biology, 13*(5), 1-18. doi:10.1186/1471-2148-13-5

Lacharité, M., & Metaxas, A. (2013). Early Life History of Deep-Water Gorgonian Corals May Limit Their Abundance. *PLoS ONE, 8*(6), e65394. doi: 10.1371/journal.pone.0065394

McFadden, C. S., Benayahu, Y., Pante, E., Thoma, J. N., Andrew, P., Arez, N., & France, S. C. (2011). Limitations of mitochondrial gene barcoding in Octocorallia. *DNA Barcoding, 11*, 19-31. doi: 10.1111/j.1755-0998.2010.02875.x

McFadden, C. S., Brown, A. S., Brayton, C., Hunt, C. B., & van Ofwegen, L. P. (2014). Application of DNA barcoding in biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs, 33*, 275-286. doi:10.1007/s00338-013-1123-0

Oliveira, L. M. C., Lim, A., Conti, L. A., & Wheeler, A. J. (2021). 3D Classification of Cold-Water Coral Reefs: A Comparison of Classification Techniques for 3D Reconstructions of Cold-Water Coral Reefs and Seabed. *Frontiers in Marine Science,8, 2296-7745*. doi:10.3389/fmars.2021.640713

Oury, N., Gélin, P., & Magalon, H. (2020). Cryptic species and genetic connectivity among populations of the coral Pocillopora damicornis (Scleractinia) in the tropical southwestern Pacific. *Marine Biology, 142*, 167. doi:10.1007/s00227-020-03757-z

Palma, M., Casado, M., Pantaleo, U., Pavoni, G., Pica, D., and Cerrano, C. (2018). SfM-based method to assess gorgonian forests (*Paramuricea clavata* (Cnidaria. Octocorallia). *Remote Sens. 10*:1154. doi:10.3390/rs10071154

Pante, E., Heestand Saucier, E., & France, S.C. (2013) Molecular and morphological data support reclassification of the octocoral genus Isidoides. *Invertebrate Systematics, 27,* 365–378. doi:10.1071/IS12053

Patterson, M. R. (1984). Patterns of whole colony prey capture in the octocoral, *Alcyonium siderium*. *Biol. Bull, 167*: 613-629.

Quattrini, A. M., Wu, T., Soong, K., Jeng, M., Benayahu, Y., & McFadden, C. S. (2019). A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evol Biol, 19*(1), 116. doi:10.1186/s12862-019-1427-y

Ragnarsson, S. Á., Burgos, J. M., Kutti, T., Beld, I., Egilsdóttir, H., Arnaud-Haond, S., & Grehan, A. (2016). The Impact of Anthropogenic Activity on Cold-Water Coral. *Springer International Publishing AG*, 989-1023. doi:10.1007/978-3-319-21012-4_27

Roberts, J. M., Wheeler, A. J., & Freiwald, A. (2006). Reefs of the Deep: The Biology and Geology of Cold-Water Coral Ecosystems. *Science, 312*, 543-547. doi:10.1126/science.1119861

Roberts, J. M., Wheeler, A., Freiwald, A., & Cairns, S. D. (2009). Cold-Water Corals: The Biology and Geology of Deep-Sea Coral Habitats. Cambridge: Cambridge University Press.

Roberts, S., & Hirshfield, M. (2004). Deep-sea corals: out of sight, but no longer out of mind. *Frontiers in Ecology and the Environment, 2*(3), 123-130. doi:10.1890/1540-9295(2004)002[0123:DCOOSB]2.0.CO;2

Rogers, A. D. (1999). The Biology of Lophelia pertusa (Linnaeus 1758) and Other Deep Water Reef Forming Corals and Impacts from Human Activities. *International review of hydrobiology, 84(4)*, 315-406.

Rossi, S. &.-M. (2009). Near bottom phytoplankton and seston: Importance in the pelagic-benthic coupling processes. *Marine Phytoplankton*, 45-85.

Rossi, S. B. (2017). An overview of the animal forests of the world. *Marine Animal Forest, Rossi, S., Springer*, 1-25. doi:10.1007/978-3-319-17001-5_1-11007/978-3-319-17001-5_1-1

Roth, M. S. (2014). The engine of the reef: photobiology of the coral-algal symbiosis. *Frontiers in Microbiology, 5*, 422. doi:10.3389/fmicb.2014.00422

Schroeter, J. C., Maloy, A. P., Rees, C. R., & Bartron, M. L. (2020). Fish mitochondrial genome sequencing: expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA. *Conservation Genetics Resources, 12*, 433-446. doi:10.1007/s12686-019-01111-0

Sponaugle, S., LaBarbera, M. (1991). Drag-induced deformation: a functional feeding strategy in two species of gorgonians. J *Exp Mar Biol Ecol 148*: 121-134.

Sundahl, H., Buhl-Mortensen, P., & Buhl-Mortensen, L. (2020). Distribution and Suitable Habitat of the Cold-Water Corals Lophelia pertusa, Paragorgia arborea, and Primnoa resedaeformis on the Norwegian Continental Shelf. *Frontiers in Marine Science, 7*, 213. doi:10.3389/fmars.2020.00213

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology, 21*(8), 2045-2050. doi:10.1111/j.1365-294X.2012.05470.x

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183*, 4-18. doi:10.1016/j.biocon.2014.11.019

Uda, K., Komeda, Y., Fujita, T., Iwasaki, N., Bavestrello, G., Giovine, M., . . . Suzuki, T. (2013). Complete mitochondrial genomes of the Japanese pink coral (Corallium elatius) and the Mediterranean red coral (Corallium rubrum): a reevaluation of the phylogeny of the family Coralliidae based on molecular data. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 8*(3), 209-219. doi:10.1016/j.cbd.2013.05.003

Uda, K., Komeda, Y., Koyama, H., Koda, K., Fujita, T., Iwasaki, N., & Suzuki, T. (2011). Complete mitochondrial genomes of two Japanese precious corals, Paracorallium japonicum and Corallium konojoi (Cnidaria, Octocorallia, Coralliidae): Notable differences in gene arrangement. *Gene, 476*(1-2), 27-37. doi:10.1016/j.gene.2011.01.019

Weinnig, A. M., Gómez, C. E., Hallaj, A., & Cordes, E. E. (2020). Cold-water coral (Lophelia pertusa) response to multiple stressors: High temperature affects recovery from short-term pollution exposure. *Scientific Report, 10*, 1768. doi:10.1038/s41598-020-58556-9

Wilkinson, C. (2000). Status of Coral Reefs of the World: 2000. Global Coral Reef Monitoring *(eds) Network and Austr Inst Mar Sci*, Townsville, Australia: 363 pp

Wolstenholme, D. R. (1992). Animal Mitochondrial DNA: Structure and Evolution. *International Review of Cytology, 141*, 173-216. doi:10.1016/S0074-7696(08)62066-5

Woodhead, A. J., Hicks, C. C., Norström, A. V., Williams, G. J.,& Graham, N. A. J. (2019). Coral reef ecosystem services in the Anthropocene. *Funct Ecol*, 33: 1023– 1034. doi:10.1111/1365-2435.13331

Ying, H., Cooke, I., Sprungala, S., Wang, W., Hayward, D. C., Tang, Y., . . . Miller, D. J. (2018). Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome biology, 19*(1), 175. doi:10.1186/s13059-018-1552-8

2. Chapter 2: Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to assemble Mitochondrial Genomes

2.1. Abstract

The mitochondrial genome in Anthozoans, in particular the Octocorallia subclass, exhibits a very slow rate of molecular evolution compared to other taxonomic groups. As such, DNA barcoding using sequencing of just one mitochondrial gene (typically the COI) is inadequate to distinguish many coral species. We developed an approach for sequencing, assembling and annotating complete mitogenomes from octocorals with large-scale, universal application. We designed universal primers that allowed the amplification of the mitogenome by long-range PCR (PCR products > 8000 bp), followed by sequencing with third-generation technology, namely long-read sequencing (Oxford Nanopore Technologies). A total of 17 octocoral mitogenomes were successfully sequenced and assembled using *de novo*. The circular mitogenomes encode 14 protein-coding genes (*Nad1-6, Nad4L, cox1-3, Cytb, MutS, Atp6,* and *Atp8*), two ribosomal RNA genes (r12S and r16S), and a transfer RNA (*trnM*). Of the five existing gene orders in octocorals, we have sequenced species representing three gene orders, *Isidella elongata* representing gene order B, *Corallium rubrum* representing gene order C, and the remaining species representing gene order A. We successfully placed all 17 reconstructed mitogenomes in the Octocorallia phylogenetic tree and identified all the unknown species. at least at the genus level. We also recovered three well-supported major clades that correspond to the ones in previous phylogenetic studies and one clade that appears to be basal to the *Holaxonia-Alcyoniina* clade. Not only can this methodology be used to barcode species across all species within the Octocorallia subclass, but it also proves to be extremely useful in species identification and to sequencer complete

mitogenomes of environmental samples (eDNA) through metabarcoding, therefore it is essential for coral biodiversity assessments based on eDNA metabarcoding and a way of cataloging biodiversity.

2.2. Introduction

DNA barcoding is a valuable alternative tool for species identification that complements alpha taxonomy in many animal groups. Traditionally, DNA barcoding targets short DNA regions, usually fragments of mitochondrial genes, which are amplified by PCR, sequenced, and subsequently used as barcodes to identify species (Thomsen and Willerslev, 2015). The challenge of DNA barcoding is the lack of genetic resources for many key species, including species from the Octocorallia subclass, that also cascades down to other fields such as biodiversity assessment, from inferring evolutionary relationships between species in phylogenetics to eDNA metabarcoding. That is, the success of biodiversity assessment using DNA barcoding depends on a high-quality updated taxonomic library (Taberlet et al., 2012). The cytochrome *c* oxidase I (COI) is the most commonly used barcode region in species identification in many animal groups (McFadden et al., 2010) but it contains some limitations. The first indicates that it is limited to a specific small region of the mitochondrial genome, and some regions may be unavailable in public databases (Schroeter et al., 2019). Secondly, for some invertebrate groups, there is a lack of mitochondrial gene variation that could limit the use of certain gene markers, such as COI, as a barcode for species recognition. Even though COI in most animal groups is taxonomically informative for Cnidarians, it requires more sequence data (i.e., other genes or the complete mitogenome) (Shearer et al., 2002; Herbert et al., 2003; McFadden et al., 2011).

The advent of next-generation sequencing (NGS) technologies has revolutionized the fields of molecular ecology and evolutionary genetics. NGS allows the generation of a high number of sequence reads in parallel, making it possible to sequence entire mitogenomes at reduced coasts (Bleidorn, 2016; Lischer and Shimizu, 2017). Sequencing mitochondrial genomes, in particular, can be quite helpful in distinguishing species because it covers multiple genes, thus enhancing genetic coverage and increasing the chances of sampling unique, species-specific molecular characters (e.g., mutation, gene inversions). While past mitogenome sequencing methods such as *primer walking* (i.e., Sanger sequencing) divides longer sequences into consecutive fragmented sequences, it only sequences a single DNA fragment at a time at a higher cost, NGS can sequence millions of fragments at once (Baeza, 2020). Oxford Nanopore Techniques (ONT), a third-generation sequencing technique, is able to produce long (i.e., average read lengths between 6-8 Kbp), easy, and fast consensus reads (Bleidorn, 2015). Lately, this technology has been reducing its raw read error rate with increasing raw read accuracy. This technology is becoming the go-to technology for genome assemblies due to its ability to generate long-reads that have much higher contiguity (compared to short-read sequencing platforms such as Illumina), ensuring completeness of the assembly compared to short reads (Nordström et al. 2013; Chang et al., 2020).

In general, mitochondrial genomes consists of one double-stranded circular DNA molecule encoding 14 protein-coding genes, two rRNAs encoding the mitochondrial ribosomes (r12S and r16S), and one tRNAs used for translation (Niu et al., 2020). Compared to the nuclear genome, the mitochondrial genome has an evolutionary rate of 5-10 times faster. Therefore, it is common to use mitochondrial genes to examine phylogenetic relationships between species. The gene arrangement of mitochondrial genomes usually remains unchanged over long periods of evolutionary times, which has helped in the interpretation of phylogenies in and across major

Metazoan groups. (Boore, 1999; Uda et al., 2013). This has been observed in successful Scleractinian phylogeny reconstruction (Niu et al., 2020; Fukami & Knowlton, 2005) and in Octocorallia phylogenetic studies, where they were able to identify at least three gene order rearrangement events that occurred during evolution (Uda et al., 2013, 2011; Brockman & McFadden, 2012).

In this study, our main goal was to develop a pipeline to assess the biodiversity of corals by sequencing the entire mitogenomes of species from the Octocorallia and Hexacorallia subclasses using long-range PCR and long-read sequencing (Oxford Nanopore Technologies). The study focused on cold-water species, particularly circalittoral and deep-sea species occurring along the southern coast of Portugal for which genetic resources are lacking. The pipeline development included four main objectives: 1) to develop coral-specific primers to amplify the mitogenomes using sequences from available databases; 2) expand the available library by sequencing, assembling, and annotating the mitogenomes of available samples from the Portuguese coast; 3) confirm the mitochondrial gene order arrangement of the target species and see if they are in conformance with what was described in previous literature; and 4) do a phylogenetic reconstruction based on mitogenome sequence data to infer the phylogenetic placement and genetic affinity of the target species within Octocorallia, including species that were not conclusively identified based on morphology.

2.3. Materials and methods

2.3.1. Primer design and in silico testing

While the main focus of this study are Octocorals with a minor extension to Scleractinians (although we sequenced some scleractinian samples, we will not present the results for this thesis), we designed and tested (*in silico*) primers for the following orders of the Hexacorallia subclass: Antipatharia, Scleractinia, and Zoantharia. Therefore, we describe the methodology of both subclasses, including these orders, only in this subsection.

The general approach consisted in designing universal primers for the amplification of coral mitogenomes by long-range PCR (PCR products > 8000 bp), followed by sequencing with third-generation technology (i.e., long-read sequencing), namely Oxford Nanopore. Since the mitochondrial genome is circular, we designed primer sets in two opposite and (tentatively) equidistant regions to amplify the mitogenome in two PCR reactions, one for each half of the molecule (Figure 2.1).

The search for conserved regions to anchor the primers was performed hierarchically with the analysis of the mitogenomes of all coral species from the Octocorallia and Hexacorallia subclasses available in NCBI GenBank. The diversity of mitochondrial gene arrangements and high sequence divergence across the different coral groups analyzed hindered our ability to identify regions of universal applicability for the two subclasses. Therefore, we opted to design separate primer sets for the subclass Octocorallia and for the orders Scleractinia, Antipatharia, and Zoantharia (Hexacorallia).

All available complete mitochondrial genomes of the Octocorallia and Hexacorallia Subclasses were downloaded from NCBI GenBank using the Geneious Prime v2020.2.4 software. The mitogenomes of all species available for each group were aligned with the MAUVE plugin of

the Geneious software, which gives an alignment of all genomes with conserved gene regions linked together as blocks. Visual inspection of the alignments made for each group revealed that the regions coding for the two mitochondrial ribosomes (12S rRNA and 16S rRNA) corresponded to highly preserved regions for the vast majority of mitogenomes of each group, with the two genes located at approximately equidistant poles within the circular mitogenome. Their positions in the mitogenome allowed to design two sets of primers to amplify the entire mitogenome in two segments, one to amplify one half of the mitogenome, amplicon 1(set 1), and the other to amplify the other half of the mitogenome, amplicon 2 (set 2) (Figure 2.1). To facilitate genome assembly and circularization (see below) the primer sets were designed so that the resulting amplicons had overlapping regions at both terminal ends. For each group, the r12S and r16S gene blocks were extracted and aligned using the MAUVE plugin from Geneious software. The extracted regions were manually evaluated for conserved regions transversally to all species analyzed. Once these regions were identified, the primer binding sites for forward and reverse primers were selected according to the general rules for primer design using Geneious Prime. The primer set 1 had the forward primer anchored to the r12S and the reverse primer anchored to the r16S, whereas primer set 2 had the forward and reverse primers anchored reversely (Figure 2.1). This allowed the amplification of each amplicon for each sample in two PCR reactions. To account for variability in the sequences of some species (due to mutations) at the primer binding regions, we used the universal base inosine that binds to all-natural DNA nucleotides (Geller et al., 2013). Finally, the primers were tested *in silico* with Geneious Prime for several of the coral species used to design the primers. The sequence information and amplification conditions of the primers are shown in Table 2.1.

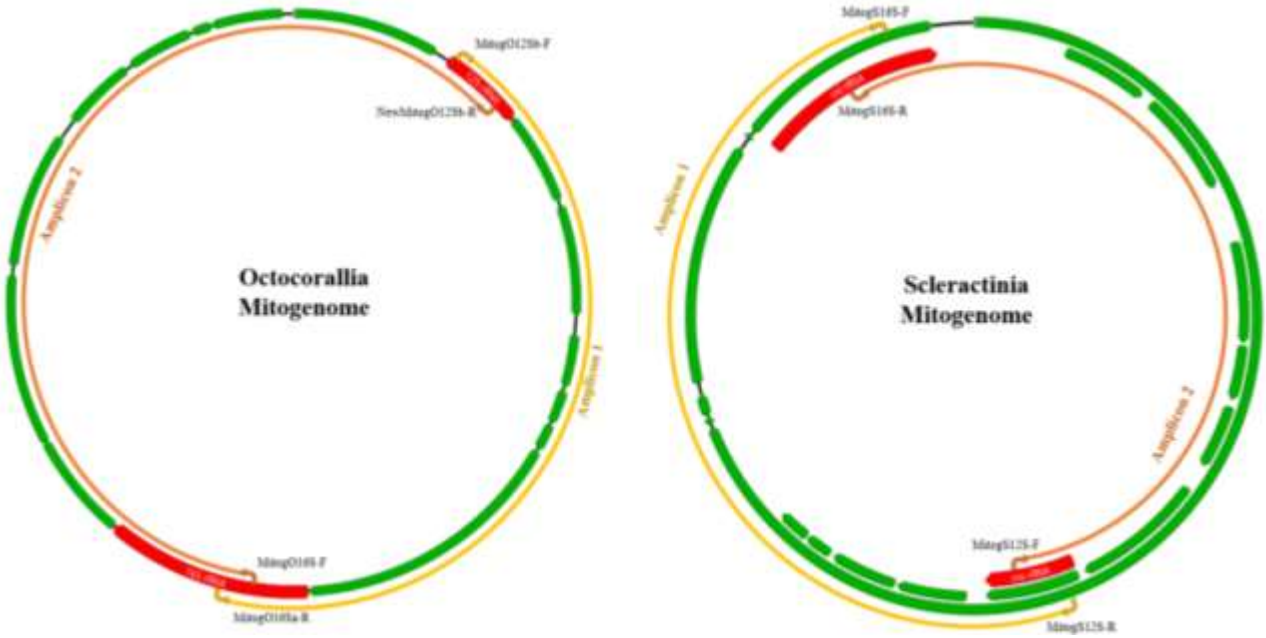

Figure 2.1. Schematic representation of the complete circular mitogenome from representative species of the subclass Octocorallia (left) and order Scleractinia (right). The primer binding regions of each primer set and resulting amplicons (Amplicons 1 and 2) are shown.

Table 2.1. Primer sequence information for the primer sets designed for each coral group to amplify the two regions of the mitogenome (amplicon 1 and amplicon 2) using long-range PCR. The 2'-Deoxyinosine [I] oligo modifications are shown in bold.

Amplified region	Primer Name	Primer sequence	Subclass/Order	Primer Size (bp)	Tm (C°)	GC%
Amplicon 1	MitogO12Sb-F	GGCAGCAGTAGAGAAT**I**TTGTGC	Octocorallia	23	61	48
	MitogO16Sa-R	AGAACGCTCTACTA**I**CAAGCCA**I**T	Octocorallia	24	59	42
Amplicon 2	NewMitogO12Sb-R	GTCTGCTGGCACTTAGTTAGACAG	Octocorallia	24	63	50
	MitogO16S-F	CTAGACTAAACCCC**I**ATAGACACC	Octocorallia	24	61	46
Amplicon 1	MitogS12S-F	AATTCGATA**I**TCCGCGAG**II**ACC	Scleractinia	23	59	44
	MitogS16S-R	CAGTAAAG**I**TCCATGGGG**I**CTTC	Scleractinia	23	61	48
Amplicon 2	MitogS12S-R	ACA**I**AAATTGACGACGGCCATGC	Scleractinia	23	61	48
	MitogS16S-F	TAAATGGCCGCGGTAACACT**I**AC	Scleractinia	23	61	48
Amplicon 1	MitogA12S-F	TTAGAGACCCTGGTAGTC**I**ACAC	Antipatharia	23	61	48
	MitogA16S-R	CCCCAACCAAACTGTC**II**ACTTAC	Antipatharia	24	61	46
Amplicon 2	MitogA12S-R	GTTACGACTTGCT**I**AACCTCGTAG	Antipatharia	24	61	46
	MitogA16S-F	GGTCAATTGTCAAAAGGGCAA**I**CC	Antipatharia	24	61	46
Amplicon 1	NewMitogZ12S-F	CAGGATTAGAGACCCTGGTAGTCC	Zoantharia	24	64	54
	NewMitogZ16S-R	GTAAAGCTCAACGGGGTCTTTTCG	Zoantharia	24	63	50
Amplicon 2	NewMitogZ12S-R	CTTGCGATCGTACTACTCAGGCG	Zoantharia	23	64	57
	NewMitogZ16S-F	CCGAAACCAAGTGATCTAGCCATG	Zoantharia	24	63	50

2.3.2. Validation of the primers in vitro

The primer sets designed for Octocorallia and Scleractinia were tested on multiple samples of cold-water corals occurring in Portugal, as well as a few species of tropical scleractinians. Most samples were collected from incidental coral catches by bottom fisheries in Sagres (Portugal) (Dias et al. 2020) or during sampling campaigns of the project HABMAR, except for one sample collected in Granada (Spain) and four scleractinian samples collected in Cape Verde. (Table S2.1 in Supplementary Materials). The majority of the samples were preserved in a 20% salt-saturated DMSO solution, which has been shown to be a superior preservative of high molecular weight DNA in corals compared to EtOH (Gaither et al. 2010). The samples from Cape Verde were preserved in Zymo DNA buffer. The DNA extractions of preserved coral tissue were made using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer´s protocol.

The mitogenome of each sample was amplified in two long-range PCR using the LongAmp® Taq 2X Master Mix (New England BioLabs). The PCRs were performed for a total reaction volume of 25µL using 5µL of DNA at 10 ng/µL or the stock solution when DNA yield was lower than 10 ng/µL. DNA concentration was measured with a NanoDrop 2000/2000c spectrophotometer. The following cycling conditions were used: initial denaturation at 94℃ for 30 sec, 30 cycles of denaturation at 94℃ for 30 sec, annealing (56℃- 59℃) for 1 min, and extension at 65℃ for (7min-11min), and a final extension step at 65℃ for 10 min (see Table 2.2 for detailed cycling conditions for each species).

In total, we tested the primers *in vitro* on 20 putative species of Octocorallia and 8 specimens of Scleractinia (Table 2.2). The first attempts to amplify the red coral mitogenome (*Corallium rubrum*) with primers designed for Octocorallia failed due to the distinct gene

organization in this species, which contains an inversion of a gene block that includes the 16S rRNA gene, a primer binding region (Figueroa & Baco 2014) (Figure 1.1). To amplify the mitogenome of this species it was necessary to use a different combination of the primers designed, that is, with the primer originally designed as Reverse in the r16S functioning as Forward and vice versa (Table 2.2). The mitogenome of the bamboo coral *Isidella elongata* has also undergone a gene inversion for a block of genes containing the r16S. But since we did not make this change during amplification and did not alter the primer combinations as we did for species with different gene order (i.e., *C. rubrum*) we failed to amplify this species (Figure 1.1) and for that reason do not report the sequencing data for this species. The PCR products were screened on an agarose gel at 0.8%.

Table 2.2. Primer sets and PCR profiles used for amplification of the complete mitogenome of each target species. The annealing temperature was calculated based on the primer with the lowest Tm minus 2℃). Check table 2.1 for detailed information about primers used for each amplicon.

					Amplicon 1			Amplicon 2		
Species	Subclass	Order	Family	[DNA] (ng/µL)	Range of expected/predicted PCR product size(≈bp)	Extension time (min)	Annealing Temperature (°C)	Range of expected/predicted PCR product size(≈bp)	Extension time (min)	Annealing Temperature (°C)
Eunicella verrucosa	Octocorallia	Alcyonacea	Gordoniidae	10						
Paramuricea grayi	Octocorallia	Alcyonacea	Plexauridae	stock						
Alcyonium acaule	Octocorallia	Alcyonacea	Alcyoniidae	stock						
Eunicella gazella	Octocorallia	Alcyonacea	Gorgoniidae	stock						
Unknown sp.1	Octocorallia	Alcyonacea	Unknown	10						
Unknown sp.2	Octocorallia	Alcyonacea	Unknown	10						
Unknown sp.3	Octocorallia	Alcyonacea	Unknown	10						
Unknown sp.4	Octocorallia	Alcyonacea	Unknown	10						
Unknown sp.5	Octocorallia	Alcyonacea	Unknown	10	(8000-9000)	7	57	(11200-12100)	10	59
Spinimuricea atlantica	Octocorallia	Alcyonacea	Plexauridae	10						
Callogorgia verticillata	Octocorallia	Alcyonacea	Primnoidae	10						
Pennatula rubra	Octocorallia	Pennatulacea	Pennatulidae	10						
Leptogorgia sarmentosa	Octocorallia	Alcyonacea	Gorgoniidae	10						
Unknown sp.6	Octocorallia	Alcyonacea	Gorgoniidae	10						
Ellisella paraplexauroides	Octocorallia	Alcyonacea	Ellisellidae	10						
Paramuricea sp.	Octocorallia	Alcyonacea	Plexauridae	10						
Unknown sp.7	Octocorallia	Alcyonacea	Plexauridae	10						
Paramuricea hirsuta	Octocorallia	Alcyonacea	Plexauridae	10	12300	10	57	7100	6	59
Corallium rubrum	Octocorallia	Alcyonacea	Coralliidae	10	12300	10	59	7300	6	57
Isidella elongata	Octocorallia	Alcyonacea	Keratoisididae	stock	11000	9	57	8400	7	59
Dendrophyllia ramea	Hexacorallia	Scleractinia	Dendrophylliidae	10						
Dendrophyllia cornigera	Hexacorallia	Scleractinia	Dendrophylliidae	10	9000			10400		
Dendrophyllia cornigera	Hexacorallia	Scleractinia	Dendrophylliidae	10						
Dendrophyllia ramea	Hexacorallia	Scleractinia	Dendrophylliidae	10		7	57		9	59
Porites porites	Hexacorallia	Scleractinia	Poritidae	stock						
Porites astreoides	Hexacorallia	Scleractinia	Poritidae	stock	(9000-9200)			(8400-10500)		
Siderastrea radians	Hexacorallia	Scleractinia	Rhizangiidae	stock						
Favia fragum	Hexacorallia	Scleractinia	Faviidae	stock						

2.3.3. Library preparation and sequencing

Sequencing libraries were prepared with PCR products obtained under optimized amplification conditions for all target species (Table 2.2). This time, the reactions were carried out to a final volume of 50 µL to obtain higher amounts of DNA (a prerequisite from the Nanopore platform). The PCR products were screened on an agarose gel at 0.8% and subsequently purified with AMPure XP magnetic beads and quantified with Qubit® dsDNA BR Assay Kit (Life Technologies) following the manufacturer's protocols for library preparation and sequencing.

In total, we performed three separate sequencing runs using the Oxford Nanopore's MinION Mk1c portable sequencer. In a first run, we sequenced one sample (*Eunicella verrucosa*) using a single-use flow cell (flongle) and the sequencing kit "Ligation Sequencing Kit" (SQK-LSK109) following Nanopore's protocol and recommendations. The PCR products for each of the two amplicons were pooled in equimolar concentrations. In the second and third runs, we multiplexed multiple samples in the same sequencing run using a reusable flow cell for the MinION Mk1c and the "Native barcoding amplicons" kit (with EXP-NBD104, EXP-NBD114, and SQK-LSK109). We sequenced a total of 10 and 19 samples in the second and third sequencing run, respectively (Table S2.2). Unlike the first sequencing run, for the second run, the PCR products for amplicons 1 and 2 were combined at a ratio of 100:200 fmol because a preliminary analysis of the sequence data obtained for *Eunicella verrucosa* suggested a bias in coverage favoring amplicon 1 (the smallest amplicon: ~8 kb). This bias in coverage was subsequently invalidated and as such for the last sequencing run the amplicons were pooled in equimolar amounts as recommended by Nanopore's protocol. For a few samples, the final DNA amount of the PCR products was lower than that recommended by Nanopore but was still sequenced (Table

S2.2). Basecalling and read demultiplexing were performed in real-time by Guppy v. 5.01.3 (Oxford Nanopore Technologies) in all sequencing runs, with basecalling set to high accuracy mode.

2.3.4. Quality control and mitogenome assembly

Barcode sequences and sequencing adapters of demultiplexed reads were trimmed with Guppy v. 6.0.1. Although Guppy performs real-time filtering of low-quality reads (Phred quality score cutoff of 7; ~80% accuracy) during sequencing, the reads were further filtered on minimum average read Q-score (10; ~90%), read length (depending on the sample; 1000-8000) and trimmed off remaining primer sequences using NanoFilt v 2.8.0 (De Coster et al., 2018).

The mitogenomes were assembled *de novo* with Flye v2.9 (Kolmogorov et al., 2019, 2020). Due to the lack of ideal reference mitogenomes to perform reference-based assemblies in many species, this approach is suitable to our goal of developing a tool for sequencing and assembling the mitogenome of any coral species, including species unknown to science or of uncertain phylogenetic affinity. Flye was run in "metagenome" mode to account for non-uniform coverage and to prevent failure in producing the initial disjointing assembly. We used a total of ten iterations of assembly polishing. The minimum overlap parameter, which sets the minimum overlap length for two reads to be considered overlapping, was set manually and individually adjusted from sample to sample (1kb-8kb) until the assembly obtained contained at least two high coverage contigs that roughly corresponded to amplicons 1 and 2. The resulting assemblies were subject to a final step of polishing to improve the accuracy of the assembly using Medaka v1.4.4 (https://github.com/nanoporetech/medaka). The two high coverage contigs of the final, polished assemblies, corresponding to amplicons 1 and 2, were then merged manually in Geneious Prime by aligning both contigs to each other and generating a consensus sequence to resolve the

overlapping region. The mitogenomes were circularized with AWA (Machado et al., 2018; https://gitlab.com/MachadoDJ/awa), which finds and trims putative overlapping sequences at the ends of a contig thereby validating the circularity of the mitogenome. In a few cases where the sequence overlap was too small to use AWA, the circularization was performed manually in Geneious Prime. Finally, the *fixstart* task of Circlator v1.5.5 (Hunt et al., 2015) was used to change the starting position of the circular mitogenome sequence at the COI gene. The COI sequence of the top coral hit for the putative COI sequence assembled identified with a BLASTn search was passed to *circlator fixstart* to use as a starting point.

2.3.5. Mitogenome annotation

The boundaries of protein-coding genes and non-coding RNAs were annotated in Geneious Prime in three sequential steps. First, open reading frames (ORFs) were identified using the invertebrate genetic code. Second, the mitogenomes were annotated by running BLAST on the ORF-annotated sequences using the Refseq_protein database. Lastly, the annotations obtained by BLAST were refined by comparison to a coral database, thus using multiple coral reference mitogenomes to import annotations based on similarity.

2.3.6. Phylogenetic analysis

The phylogenetic relationships between the species sequenced in this study and other octocorals with complete mitogenome sequences publicly available were examined with a maximum likelihood (ML) analysis in IQ-TREE 2 (Minh et al. 2020). The analysis was based on sequence data for 16 mitochondrial genes, 14 protein-coding genes, and the two mitochondrial ribosomal

RNA subunits (r12S and r16S), for a total of 93 putative species covering the main clades of Octocorallia (Table S2.3-Supplementary materials).

For protein-coding genes, the sequences were aligned with mafft and trimmed with Gblocks in TranslatorX using a custom python script (https://github.com/cymon). For the r12S and r16S genes, individual species sequences were first extracted from mitogenome alignments performed with the MAUVE plugin and subsequently aligned and trimmed in Geneious. The gaps in the alignments of the r12S and r16S genes were removed manually. The final alignments of each gene contained no indels. A partitioned analysis was performed by fitting a separate evolutionary model of sequence evolution for each gene using ModelFinder and based on the *Bayesian information criterion* (BIC) score. Branch support was assessed with ultrafast bootstrap approximation using 1000 replicates (Minh et al 2013; Hoang et al. 2018).

2.4. Results

We designed and tested (*in silico)* primers for the following orders of the Hexacorallia subclass: Antipatharia, Scleractinia, and Zoantharia. Therefore, we present results of both subclasses, including these orders, up until the *in silico* primer analysis subsection, beyond that, we focus the results and discussion only on octocorals.

2.4.1. In silico primer performance in corals

A total of 229 coral mitogenomes were used to design the primers, including 87 species from 24 families of the Octocorallia subclass and 142 species from 42 families of the Hexacorallia subclass (19 species from 4 families of order Antipatharia; 75 species from 16 families of order Scleractinia; and 17 species from 8 families of order Zoantharia) (Table S2.3 & S2.4 in supplementary

materials). Primer amplification of all species from each Subclass was tested *in silico* using the reference library used to design the primers for each target coral group in Geneious Prime (Table 2.1). From the 79 available octocoral mitogenomes, 13 exhibit different gene arrangements (Table S2.3-Supplementary materials). The virtual PCRs worked in all 66 species with gene order A using the original primer combination (Figure 2.1). For the remaining 13 species with different gene arrangements (gene orders B, C, D, and E; see Figure 1.1) it was necessary to use an alternative combination of primers. For example, amplicon 1 (as represented in Figure 2.1) of *Corallium rubrum*, which has gene order C, was amplified with both "forward" primers due to the inversion of r16S. The primers successfully attached to all available Octocorals *in silico* exhibiting two overlaps, one in the r12S regions with an average size of 85 bp, and the other in the r16S regions with an average size of 635bp.

2.4.2. *Mitogenome sequencing and assembly*

Overall, we were able to reconstruct the mitogenomes of most species and were able to determine, at least at the genus-level, the identity of the 7 unknown samples (Table 2.4).

Although we sequenced 8 species of hexacorals but did not reconstruct their mitogenomes, as the focus of this study are the octocorals. Of the 22 octocoral species sequenced we successfully assembled 17 mitogenomes using *de novo* assembly that ranged from 18252 bp (*Corallium rubrum*, barcode 5B) to 19656 bp (*Leptogorgia sp.*, barcode06) (Table S2.5-Supplementary materials). The 22 octocorals represent 6 families from the Alcyonacea Order and 1 family from the Pennatulacea Order.

A total amount of 7,604,451Kb (Kilobases) sequence data was generated after sequencing, with an average (± SE) of 447,320.6Kb ± 60,9513.6 filtered reads per sample (min-max: 63,338Kb

to 2,722,631Kb; Table S2.5-Supplementary materials). The average number of contigs recovered per assembly was 10.8 (± 7.1) with a maximum of 21 contigs and a minimum of 2 contigs. The average length of the contigs corresponding to amplicon 1 was 8,370 bp (± 497.3) ranging between 7,522 bp (*Pennatula rubra*) to 9,091 bp (Unknown sp.5). The average length of the contigs corresponding to amplicon 2 was 11,317 bp (± 353.4) ranging between 10,724 bp (*Callogorgia verticillata*) to 12,096 bp (*Pennatula rubra*). Refer to Table 2.3 for further statistical analysis including the coverage and GC content percentage. These average lengths are based on all species except for *Corallium rubrum* that generated only one contig with very high coverage to reconstruct the mitogenome, with 18,897 bp with a coverage of 19,529, that covers the entire mitogenome, and except barcode 17 (*Paramuricea sp.*) and 18 (Unknown sp.7) that generated 3 contigs instead of 2 to reconstruct the entire mitogenome. See Table S2.5 for detailed contig information. The average sizes of the mitogenomes were 19,053bp (± 406.1) which ranged between 18,669 bp to 19,656bp (Table 2.3).

In each sample, when both contigs were linearized and aligned in Geneious Prime software, they had two overlapping regions, one that was in the region where both contigs joined with an average length of 453bp (±220) that ranged between 43bp to 580bp and the other was located at the end of one of the contigs with an average length of 24bp (±12) that ranged between 1bp to 50bp. Except for *Corallium rubrum* which only had one contig covering the entire mitogenome, therefore no overlaps were present. Also, barcodes 17 and 18 had three contigs representing the entire mitogenome, therefore there were three overlapping regions for each barcode. The great majority of the assembled contigs were slightly longer after the extra polishing with Medaka with an average size difference of 4.8bp (±3.5) that ranged between 0bp to 18bp. After the polishing and circulation of the mitogenome, 17 complete mitochondrial genomes were recovered.

Table 2.3. Assembly information regarding the average number of contigs assembled with Flye.

		Amplicon 1		Amplicon 2			
	Number of contigs recovered	**Contig Length (bp)**	**Contig coverage**	**Contig Length (bp)**	**Contig coverage**	**Final MtDNA (bp)**	**%GC content**
Average (±SE)	10.8 (±7.1)	8370 (±497.3)	22825 (±53190.5)	11318 (±353.4)	9471 (±7154.4)	19053 (±406.1)	37.66 (±2.5)
Max	21	9091	205390	12096	22432	19656	46.50
Min	2	7522	6	10724	3	18669	34.10

Table 2.4. Sample identification at species or genus-level of each barcode based on the *blastn* hit of available MutS gene in Genbank database.

Barcode	**Sample ID**	***blastn* identical%**	***blastn* reference**
03	*Alcyonium sp.*	99.9	*Alcyonium acaule*
04	*Eunicella gazella*	100	*Eunicella spp.*
5A	Unknown sp.1	100	*Eunicella spp.*
06	Unknown sp.2	100	*Leptogorgia piccola*
07	Unknown sp.3	100	*Leptogorgia piccola*
08	Unknown sp.4	100	*Leptogorgia piccola*
09	Unknown sp.5	100	*Leptogorgia piccola*
10A	*Spinimuricea atlantica*	100	*Spinimuricea klavereni*
5B	*Corallium rubrum*	100	*Corallium rubrum*
10B	*Callogorgia verticillata*	100	*Callogorgia verticillata*
11	*Pennatula rubra*	99.8	*Pennatula rubra*
12	*Leptogorgia sarmentosa*	99.9	*Leptogorgia sarmentosa*
13	Unknown sp.6	100	*Spinimuricea klavereni*
14	*Ellisella paraplexauroides*	99.5	*Ellisella ceratophyta*
15	*Isidella elongata*	99.2	*Acanella arbuscula*
16	*Paramuricea hirsuta*	98.5	*Euplexaura crassa*
17	*Paramuricea sp.*	99	*Paramuricea grayi*
18	Unknown sp.7	100	*Paramuricea clavata*
19	*Eunicella verrucosa*	100	*Eunicella verrucosa*

2.4.3. *Mitochondrial gene annotation and genome organization*

The gene order of all the octocoral mitogenomes sequenced here agrees with one of the five gene arrangements described for Octocorallia to date (Figure 1.1). Of the five existing gene orders in octocorals, we have sequenced species representing three gene orders, *Isidella elongata* for gene order B (Brugler and France, 2008), *C. rubrum* for gene order C (Uda et al., 2011), and the remaining species for gene order A (Beagley et al., 1995) (Figure 1.1; Tables 2.5 and 2.6).

All 17 mitogenomes contain the fourteen protein-coding genes (*Atp6*, *Atp8*, *cox1-3*, *Cytb*, *Nad1-6*, *Nad4L*, and *MutS*), two ribosomal RNA subunits (r12S and r16S), and one transfer RNA (*trnM*) described for octocorals (Tables 2.5 and 2.6). Most of the genes are separated by intergenic regions (IGR). Across all species, the sum lengths of the IGRs have an avenge length of 1,061bp (± 461.3), with the largest IGRs varying between species. The largest IGR was found in species with the largest mitogenomes and is located between *Cytb* and *Nad6* with an approximate size of 1,006 bp. The shortest IGR region is between *MutS-Nad4L* extends for only 4bp and is the same for all species. Most genes are directly connected to the adjacent gene (i.e., lack an IGR), with only one pair of overlapping genes present in all mitogenomes, including the species with different gene order (*C. rubrum*), which is located between the *Nad2-Nad5 genes* (13 bp overlap). There was only one mitogenome (*P. rubra*) that exhibited two overlaps between genes, one that is shared with all the other species and the other is located between *Nad1-Cytb* (95 bp) (Table 2.5 and 2.6).

The gene lengths were similar between all species (Table 2.6). There were 6 protein-coding genes where the sizes were the same between all species (*Nad5, Nad4, cox3, Atp6, Atp8,* and *cox2*), as well as for the trnM. The greatest differences in gene length between species were observed for the ribosomal RNA subunit r16S (2,132 bp ± 114.5 bp) and Nad2 (1,199 ± 87.9 bp). Overall base

composition (GC) ranged from 34.1% (*Corallium rubrum*) to 46.5% (*Eunicella gazella*) (Table 2.5 and 2.6).

Table 2.5. Classification and gene order arrangements of the 17 octocorals mitochondrial genomes assembled. Genes that are underlined are located in the heavy strand (H-strand) and the remaining genes are located in the light strand (L-strand).

Barcode	Order	Family	Species	Gene order arrangement																
3	Alcyonacea	Alcyoniidae	*Alcyonium acaule*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
4	Alcyonacea	Gorgoniidae	*Eunicella gazella*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
5A	Alcyonacea	Gorgoniidae	Eunicella sp.	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
6	Alcyonacea	Gorgoniidae	Leptogorgia sp.	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
7	Alcyonacea	Gorgoniidae	Leptogorgia sp	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
8	Alcyonacea	Gorgoniidae	Leptogorgia sp	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
9	Alcyonacea	Gorgoniidae	Leptogorgia sp	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
10A	Alcyonacea	Plexauridae	*Spinimuricea atlantica*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
5B	Alcyonacea	Coralliidae	*Corallium rubrum*	cox1	r12S	Nad1	Cytb	cox2	Atp8	Atp6	cox3	trnM	Nad4	Nad5	Nad2	r16S	MutS	Nad4L	Nad3	Nad6
10B	Alcyonacea	Primnoidae	*Callogorgia verticillata*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
11	Pennatulacea	Pennatulidae	*Pennatula rubra*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
12	Alcyonacea	Gorgoniidae	*Leptogorgia sarmentosa*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
13	Alcyonacea	Plexauridae	*Spinimuricea sp.*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
14	Alcyonacea	Ellisellidae	*Ellisella paraplexauroides*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
17	Alcyonacea	Plexauridae	*Paramuricea sp.*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
18	Alcyonacea	Plexauridae	*Paramuricea sp.*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2
19	Alcyonacea	Gorgoniidae	*Eunicella verrucosa*	cox1	r12S	Nad1	Cytb	Nad6	Nad3	Nad4L	MutS	r16S	Nad2	Nad5	Nad4	trnM	cox3	Atp6	Atp8	cox2

Table 2.6. Mitochondrial gene lengths (bp) and gene organization of species with octocoral ancestral gene order (gene order A) and *Corallium rubrum** with gene order C, for further details about different gene arrangement refer to table 2.5. In bold are the genes that are located in the H-strand and the remaining genes are located in the L-strand. The pairs of overlapping genes are underlined.

Genes	*Alcyonium acaule*	*Eunicella gazella*	Unknown sp.1	Unknown sp.2	Unknown sp.3	Unknown sp.4	Unknown sp.5	*Spinimuricea atlantica*	*Callogorgia verticillata*	*Pennatula rubra*	*Leptogorgia sarmentosa*	Unknown sp.6	*Ellisella paraplexauroides*	*Paramuricea clavata*	Unknown sp.7	*Eunicella verrucosa*	*Corallium rubrum**
cox1	1597	1596	1596	1597	1599	1597	1597	1597	1597	1566	1704	1683	1569	1597	1597	1597	1597
r12S	923	926	926	926	926	926	926	927	1117	1057	926	927	928	925	925	926	1096
Nad1	972	972	972	973	972	972	972	972	981	1074	972	972	981	972	971	972	972
Cytb	1167	1160	1160	1169	1169	1169	1169	1167	1161	1161	1167	1167	1161	1155	1155	888	1194
Nad6	558	558	558	558	558	558	558	558	564	555	558	558	555	558	558	558	**555**
Nad3	354	360	360	366	366	366	366	354	354	354	366	354	354	360	360	360	**354**
Nad4L	294	294	294	294	294	294	294	294	294	294	294	294	294	294	294	294	**294**
MutS	2970	2957	2958	2958	2958	2957	2958	2957	2967	3002	2958	2958	3000	2964	2964	2958	**2991**
r16S	1948	2183	2183	2180	2180	2180	2180	2183	2019	1788	2183	2183	2180	2179	2179	2183	**2249**
Nad2	1374	1158	1158	1158	1158	1158	1158	1158	1371	1383	1158	1158	1158	1158	1158	1158	**1140**
Nad5	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	1818	**1818**
Nad4	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	1449	**1449**
trnM	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	**71**	71
cox3	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	**786**	786
Atp6	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	**708**	708
Atp8	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	**216**	216
cox2	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	**762**	762
MtDNA (bp)	18677	19317	19246	19656	19655	19654	19655	18671	18919	18727	18722	18580	18814	18669	18668	19266	18915
Total IGR (bp)	710	1343	1271	1667	1665	1667	1667	694	684	683	626	516	824	697	697	1562	663
GC content %	36.6	46.5	36.5	37.1	37.1	37.1	37.8	37.7	37.5	37.3	37.3	38.5	37.5	36.2	37.4	38.1	34.1

2.4.4. *Phylogenetic analysis using mitogenome sequence data*

The sequence data to construct the phylogeny contained 93 species of octocorals and 16 gene fragments for a total of 1,488 aligned base pairs (bp). Alignment positions/columns for each gene are described in Table S.2.3-Supplemental materials. All mitochondrial gene sequences were used except for trnM. Numbers of variable and parsimony-informative characters observed within each gene are given in Table S.2.3-Supplemental materials.

The ML analysis recovered three well-supported major clades with 100% bootstrap support that correspond to the same clades previously found in phylogenetic studies of Octocorallia and one clade with 92% bootstrap support that includes the *Sinularia spp*. that appears to be basal to the *Holaxonia-Alcyoniina* clade: the Calcaxonia- Pennatulacea clade comprised of species from the Alcyonacea and Pennatulacea orders and including 12 families (Isididae, Anthoptilidae, Primnoidae, Pennatulidae, Protoptilidae, Ellisellidae, Helioporidae, Kophopotidae, Veretillidae, Renillidae, Virgulariidae, and Umbellulidae); the *Anthomastus-Corallium* clade comprised of species from 3 families (Alcyoniidae, Coralliidae, and Paragorgiidae) of the Alcyonacea order; and the *Holaxonia-Alcyoniina* clade comprised of species only from the Alcyonacea Order including 7 families (Acanthogorgiidae, Clavulariidae, Nephtheidae, Plexauriidae, Gordoniidae, Paramuriceidae. The tree is rooted with *Alcyonium acaule* (outgroup) which falls outside any of the clades. There are no available studies suggesting the placement of *A. acaule* which makes it more challenging to understand their position in the tree. The *Holaxonia -Alcyoniina* clade has two sister clades, one includes species of the Gordoniidae family and the other includes some species of the Alcyoniidae family (i.e., *Sinularia spp.)*. The great majority of the shallow nodes have 100% support value of their placement on the tree, refer to Figure 2.2 for node support values. We were able to place and identify, at least at the genus-level, all 17 reconstructed mitogenomes in the

Octocorallia phylogenetic tree (Figure 2.2; Table S2.3). We identified all unknown taxa (Unknown sp. 1- sp7). Unknown sp.1 was characterized as *Eunicella sp.* BC5A. Unknown sp.2-sp.5 were characterized as *Leptogorgia sp.* as they belong to the *Leptogorgia* genus, Unknown sp.7 (*Leptogorgia sp.* BC07) and Unknown sp.5 (*Leptogorgia sp.* BC09) are of the same species as they are positioned equally in the same branch, while Unknown sp.2 (*Leptogorgia sp.* BC06) and Unknown sp.4 (*Leptogorgia sp.* BC08) are of different species as they are placed in different positions on the tree. Unknown sp. 6 was characterized as *Spinimuricea sp.* (BC13), and Unknown sp. 7 was characterized as *Paramuricea sp.* (BC18). For species in which there are no references available at the genus or species level (*Spinimuricea atlantica and Alcyonium acaule)* were also placed in the tree. *Spinimuricea atlantica* was placed in the same branch as *Spinimuricea sp.* BC13 confirms their position at the genus level. The placement of *Alcyonium acaule* is questionable therefore the sequencing and assembly should be repeated. All remaining taxa, named at species-level, for which reference species of the same genus or species already exist were placed as expected (*Callogorgia verticillata, Pennatula rubra, Ellisella paraplexauroides, Corallium rubrum, Eunicella verrucosa, and Leptogorgia sarmentosa).*

The species assignment based on the *blastn* of the contigs with high coverage and ML analysis differed in some cases (Table 2.4). *Spinimuricea atlantica* was 100% identical to *Spinimuricea klavereni, Ellisella paraplexauroides* was 99.5% identical to *Ellisella ceratophyta, Isidella elongata* was 99.2% identical to *Acanella arbuscula*, and *Paramuricea hirsuta* was 98.5% identical to *Euplexaura crassa.* Since there are no available reference mitogenomes for any of these species, the *blastn* assigns the closest reference. But it does not mean that the assigned identity of the species is correct, even in cases where they are 100% identical, because the *blastn*

is based on only a short fragment (850bp) while the ML analyses were based on the 16 genes, that is, almost the complete mitogenome.

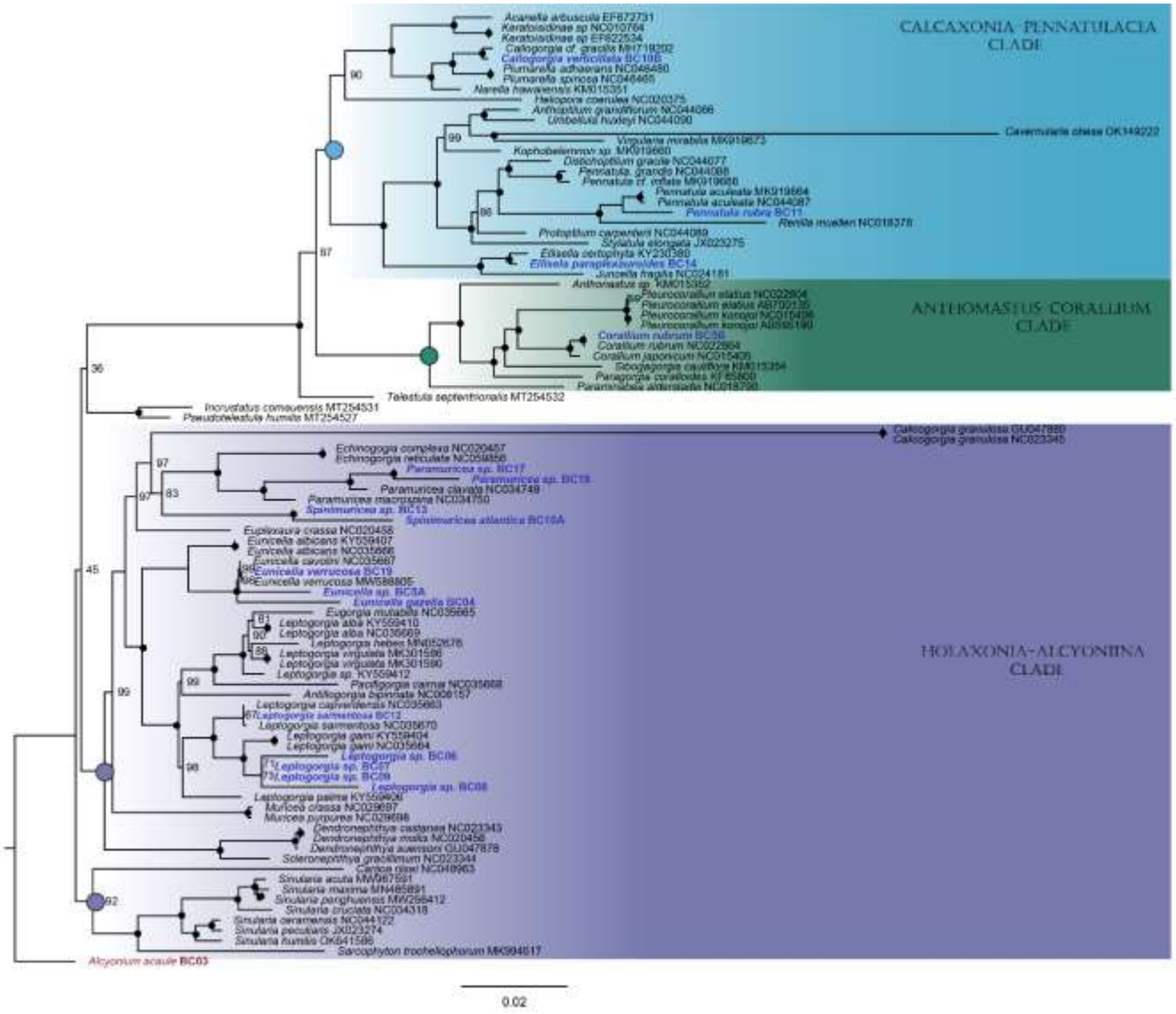

Figure 2.2. Phylogenetic tree of Octocorallia based on maximum likelihood (ML) analysis of combined sequence data. Scale bar indicates substitutions per site. Taxa from available libraries are represented in black and the mitogenomes assembled *de novo* here are represented in purple followed by their barcode number (BC#). Numbers near the nodes indicate bootstrap support (ML), with black circles indicating 100% support. The three main clades described in McFadden et al., 2006 are represented with a colored background: blue - Calcaxonia-Pennatulacea clade; green - *Anthomastus-Corallium* clade; and purple - Holaxonia-Alcyoniina clade. Large colored circles denote the major clades with 100% support except the clade representing *Sinularia spp.* with 92% support.

2.5. Discussion

In this study, we demonstrated that is possible to sequence and assemble the complete mitogenomes of octocorals by coupling long-range PCR amplification and long-read sequencing (Oxford Nanopore Technologies). The barcoding pipeline developed here is transversal to all octocorals tested, including species with different mitogenome organizations and sizes. Importantly, barcoding mitochondrial genomes using long-read sequencing improves the accuracy of the assemblies obtained (due to larger contig contiguity) while making it possible to examine genetic variation at unprecedented scales, thereby increasing the resolution of genetic data used to identify species. In addition, it provides valuable information about gene organization. Overall, we have expanded the available reference library of sequenced mitogenomes for Octocorallia with 17 additional species that occur on the Portuguese coast.

2.5.1. Phylogenetic placement in the Octocorallia tree and taxa identification

It was possible to identify species, at least at the genus-level, through phylogenetic placement in the Octocorallia tree. The *blastn* with the *mtMutS* gene helps identify species for which the mitogenome is available in the reference database but it is based only on small fragments of approximately 850bp. The phylogenetic analysis allows the placement of unknown species based on available libraries with a much higher resolution because it incorporates, in this case, 16 genes. This phylogeny has three main clades that agree with the phylogeny presented by McFaden et al., (2006). One large clade, the *Holaxonia-Alcyoniina* clade, includes members of the sub-order Holaxonia with the great majority of the taxa belonging to the Alcyoniina group. This clade corresponds to the *Alcyoniina-Holaxonia* clade of Sánchez et al. (2003) and Holaxonia-Alcyoniina clade of McFadden et al. (2006). A second large clade includes all sea pens (Pennatulacea), blue

corals (Helioporacea), and members of the sub-order Calcaxonia, the Calcaxonia-Pennatulacea clade, the same presented by McFadden et al. (2006) and corresponds to the "Calcaxonia" clade of Sánchez et al. (2003). A third smaller clade includes members of the *Coralliidae* and Alcyoniidae families with an agreement with the *Anthomastus-Corallium* clade by McFadden et al. (2006). The placement of all species for which had a species name agree with previous literature and are positioned very close to species with the same genus: *Callogorgia verticillata* BC10B (Cairns and Wirshing, 2018), *Pennatula rubra* BC 11 (McFadden 2006), *Corallium rubrum* BC5B (Uda et al, 2013), *Ellisella paraplexauroides* BC14 (Vohsen et al., 2020), *Eunicella verrucosa* BC19 (Hooper, 2021), and *Leptogorgia sarmentosa* BC12 (Poliseno et al., 2017). Except for *Spinimuricea atlantica* BC10A because there are no representatives of this genus or species available to compare and for *Eunicella gazella* BC04 for which there are no taxa from the same species available, but they are positioned closer to the taxa from the *Eunicella* genus. The placement of *Alcyonium acaule* BC03 is questionable, therefore the assembly should be repeated and further analyzed. The Alcyoniidae genus is highly polyphyletic, with genera distributed across Octocorallia but since there are no representatives of this specific species available in the reference database makes it hard to compare the placement of this species (McFadden and Ofwegen, 2013). We were able to place all the unknown taxa (Unknown sp.1-sp.7) and see if some of these taxa correspond to species already available or see what their relationship with the closest species is. Unknown sp.1 was characterized as *Eunicella sp.* BC5A because it was placed closest to the *Eunicella* genus, but this species does not correspond to any of the existing species. Unknown sp.2-sp.5 were characterized as *Leptogorgia sp.* as they belong to the *Leptogorgia* genus, Unknown sp.7 (*Leptogorgia sp.* BC07) and Unknown sp.5 (*Leptogorgia sp.* BC09) are of the same species as they are positioned equally in the same branch, while Unknown sp.2 (*Leptogorgia sp.* BC06)

and Unknown sp.4 (*Leptogorgia sp.* BC08) are of different species as they are placed in different positions on the tree. Unknown sp. 6 was characterized as *Spinimuricea sp.* (BC13) since it is placed in the same branch as *Spinimuricea atlantica* BC10A, and finally, Unknown sp. 7 was characterized as *Paramuricea sp.* (BC18) since it was placed in the same branch of other species from the *Paramuricea* genus but does not correspond to any of the existing species. This approach has proven to be a valuable tool in species identification, even if they are identified only at the genus-level they can later guide us in the morphological identifications. The goal of the phylogenetic analysis was not to resolve phylogenetic relationships between different species nor to resolve deeper nodes but to see if the barcoding approach works and see the placement of our data set.

2.5.2. Gene order arrangements

Our results show that 16 out of 17 taxa share the same gene arrangement as the octocoral ancestral gene order (gene order A), in agreement with what was previously described by McFadden et al. 2006 and Figueroa and Baco, 2015. The last taxa (*Corallium rubrum)* has a different gene order arrangement which is in agreement with species with gene order C, currently described for *Corallium japonicum* (previously *Paracorallium japonicum*) (Uda et al., 2011). We also had another sample with a different gene arrangement, *Isidella elongata* (gene order B) although we were unable to reconstruct the mitogenome of this species, we were able to identify the closest identical reference species, *Acanella arbuscula* (99.2%), based on the *blastn* analysis. Both species belong to the Isididae family (bamboo corals, which have gene order B (Brugler and France, 2008). These different gene arrangements can be explained by the gene inversions and translocation in specific gene conserved segments (Brockman and McFadden 2012). For instance, the gene order B (Isididae family) has undergone an inversion of the gene block that corresponds to block 3 of

the octocoral ancestral gene order (*mtMutS-r16S-Nad2-Nad5-Nad4*) (Brockman and McFadden, 2012). The same inversion happens for the same gene block in gene order C (*Corallium rubrum*) but in addition, there is also an inversion of the gene block that corresponds to gene block 4 of the octocoral ancestral gene order (*trnM-cox3-Atp6-Atp8-cox2*) and gene block 2 of the octocoral ancestral gene order (*Nad6-Nad3-Nad4L*) (Brockman and McFadden, 2012). Like *C. rubrum* (gene order C), the block of genes in which the r16S priming sites are located has also undergone an inversion in bamboo corals (gene order B) (Brugler and France, 2008), which explains the reason why we were unable to reconstruct the mitogenome of *I. elongata*. Adjustment of the combination of primers designed here (as done for *C. rubrum*) is likely to resolve this in future work.

2.5.3. Mitogenome reconstruction and designed primers applicability

Sequencing using third-generation platforms (Nanopore technologies) was fundamental in the reconstruction of the mitogenomes. Unlike past mitogenome sequencing methods such as the first (Sanger) and second (Illumina) generation platforms that are based on short reads (maximum read length of 300bp- 1000bp) it only sequences a single DNA fragment at a time at a higher cost, the third-generation sequencing is based on long-read producing high quality reads up to 150kbp (Bleidorn, 2016). Long sequencing reads reduce alignment and mapping errors which help to fix sequencing errors, therefore it provides substantial improvement of genome assembly (Koren & Phillippy, 2015; Bleidorn, 2016). The ultimate goal was to obtain a single contig that covers the entire mitogenome. It worked for *C. rubrum*, but not for the others. In most cases, we obtained two contigs that carry the complete mitogenome information obtained through long-range PCR amplification and long-read sequencing. There were also cases where three contigs were needed to reconstruct the mitogenome (*Paramuricea spp.* barcodes 17 and 18). The fact that it worked for

one sample and not for the others may have been a limitation of the *de novo* assembler (Flye v2.9) as most assemblers are developed to use data with random, but even coverage of the genome, although Flye was run in "metagenome" mode to account for non-uniform coverage and to prevent failure in producing the initial disjointing assembly using ten iterations of assembly polishing (Baeza, 2020). Another reason could have been that the minimum overlap parameter was adjusted manually and individually from sample to sample until the assembly obtained contained at least two high coverage contigs that roughly corresponded to amplicons 1 and 2. But essentially the complete genetic information was present in all cases through long-read sequencing, therefore we managed to reconstruct their mitogenomes. This long read platform produces a much more contiguous reconstruction of the mitochondrial genome, the amount of genetic variation in the sequence data is far much greater (e.g., 850 bp from the MutS gene versus 18,000 bp from the entire mitogenome) which makes it easier to identify the species (Lee et al., 2016).

The primers designed here for Octocorallia have the potential to amplify any coral from the Octocorallia subclass, including the ones with inverted sequences, as they have universal bases (inosines) at the primer binding sites that are polymorphic across species, including with all types of mitogenome organizations described to date (Geller et al., 2013). Although we were unable to test our primers for all possible gene order arrangements *in vitro*, the *in silico* PCRs conducted in Geneious Prime software suggest that by altering the combination of primers between r12S and r16S for species in which the anchoring region of the primers was inverted (species with gene order B and C) we are able to amplify their complete mitogenome, as it was in the case of *C. rubrum* (gene order C) and possibly the case of *Isidella elongata* (gene order B). For species with gene order D and E, we are able to amplify their complete mitogenomes by altering the extension time of both primer sets during PCR amplification since the distance between r12S with r16S is

much shorter compared with species with gene order A. Therefore, we were able to create primers that are transversal to all octocorals including the ones with different gene orders. We also believe that primers designed specifically for each order of the Hexacorallia subclass (Antipatharia, Scleractinia, and Zoantharia) have the potential to detect any species within that order including some species from the remaining orders of the Hexacorallia , Corallimorpharia and Actiniaria, as all species were successfully tested *in silico*, and some of them were tested *in vitro* (*Dendrophyllia* spp. from the Scleractinia Order) but we did not assemble them since the main focus of this study are the Octocorals.

2.5.4. Approach applicability and final remarks

Using the complete mitogenome provides a much better resolution to identify species due to the use of long-range PCR amplification and long-read sequencing than using a short specific gene such as the universal barcode COI (Nordström et al. 2013; Chang et al., 2020). Examining the polymorphism in 16 complete genes (or complete mitogenome) in contrast to a 650 bp used in COI or the 850 bp used for *mtMutS* barcode increased the ability to distinguish species. It has been recognized that for some invertebrate groups, such as the octocorals, the use of COI as a barcode for species identification is limited due to the lack of mitochondrial gene variation (Herbert et al., 2003; McFadden et al 2010). It was previously demonstrated that the use of the complete mitochondrial genome has proven to be effective in phylogenetic analysis within Octocorallia (Figueroa and Baco, 2015). This barcoding pipeline based on complete mitogenomes usinlong-readad sequencing provided much higher contiguity which allowed the placement in the Octocorallia tree and allowed us to identify species, confirming its effectiveness in phylogenetic analysis.

This method can be improved by using one primer set to amplify the entire mitogenome instead of two, this would reduce the number of steps required during PCR and reduce the number of reagents used. We have also designed a set of adjacent primers anchored on the r12S region to amplify the mitogenome of octocorals in a single reaction (as done by Deiner et al., 2017 for fish), these primers were tested *in silico* in Geneious Prime software for several octocorals. However, this approach will not capture sequence variability at primer bind regions across species nor will it produce sequence overlap to reliably circularize the mitochondrial genomes but to barcode is still useful as it still generates quite a lot of sequence data. Another way to improve the method would be a reduction in the amplification steps to reduce any amplification errors. The use of rolling circle amplification (RCA) has proven to be very beneficial for the amplification of complete metazoan mitogenomes making a product that is acceptable to high throughput sequencing techniques, by reducing and eliminating several PCR steps (Simison et al., 2006). Hooper et al. (2021) were able to successfully reconstruct the complete mitogenome of *Eunicella verrucosa* using phi29-induced rolling circle amplification. The RCA employs several random short primers that anneal to the DNA template at several sites. The extension of the annealed primers is done with the phi29 DNA polymerase, producing long concatemers of the circular DNA in a single reaction. This process is quite advantageous over the traditional PCR amplification method, the main advantages include the removal of time-consuming primer design and testing, removal of multiple PCR products to build a complete genome, and a decrease in cost, for more detailed advantages refer to Simison et al. 2006.

This approach with NGS has the advantage that allows high-throughput data and parallel multiplexed analysis on a massive scale of DNA sequences allowing the large-scale sequencing of multiple samples per run at a reduced cost (Verma and Gazara, 2021). If we want to sequence a

genome, a single run with the normal flow cell may not be enough to reliably reconstruct the genome, depending on the size of the genome. The mitochondrial genome is very small compared to the nuclear genome, therefore, to rebuild a mitogenome in a single run can be a "waste" of genetic resources, because the output generates enough data (in Gigabases) to build dozens of mitogenomes (Shearer et al., 2008; Baeza, 2020). Nanopore provides a wide range of library preparation kits, including kits that are compatible to sequence the whole genome amplification-free which reduces any potential PCR errors and that can demultiplex up to 24 native barcodes. Some studies managed to use this method which does not require the amplification step during sequencing library preparation (Bleidorn, 2016; Baeza, 2020). The approach used in Baeza, 2020 has the advantage of not having a PCR step, but the sequence reads obtained contain both mitochondrial genome and nuclear genome. As such, if the goal is just to rebuild the mitogenome, there is a waste of genetic resources. Although the approach developed in this thesis, that couples long-range PCR and long-read sequencing, might have the disadvantage of introducing PCR errors, with high sequencing coverage we hope to get around this limitation. The PCR amplification increases the copies of the mitogenome of the DNA extract that is purified, in multiplex along with many barcoded and pooled samples. This process reduces the proportion of genomic (i.e., non-mitochondrial) DNA in the extracts, the sequencing creates sequence reads that are mostly from the mitogenome. Therefore, this approach allows to sequence in a much more "targeted" way, allowing to pool many samples in the same sequencing run.

Importantly, this approach allowed us to sequence mitogenomes of multiple species, therefore improving the reference libraries available for coral biodiversity assessments based on eDNA metabarcoding (sequence data available for many more species and many more genes could potentially be used). With increasing improvement in the accuracy of basecalling algorithms for

Nanopore sequence data, this approach has the potential to be highly applicable in studies of eDNA metabarcoding based on entire (or a single of the two large amplicons produced here) mitogenomes. For instance, Deiner et al. (2017) were able to successfully sequence fish mitogenomes from eDNA samples (seawater), which contradicts the previous assumption that eDNA is highly degraded in the environment (Thomsen &Willerslev, 2015). This way, this approach, along with eDNA metabarcoding, is valuable to catalog biodiversity and aid in biodiversity assessment.

2.6. References

Baeza, J. A. (2020). Yes, we can use it: a formal test on the accuracy of low-pass nanopore long-read sequencing for mitophylogenomics and barcoding research using the Caribbean spiny lobster Panulirus argus. *BMC Genomics, 21*, 882. doi:10.1186/s12864-020-07292-5

Bayer, F. M. (2003). A new genus of the scleraxonian family Coralliidae (Octocorallia: Gorgonacea). *Proceedings of the Biological Society of Washington, 116*, 222-228.

Bleidorn, C. (2016). Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity, 14*(1), 1-8. doi:10.1080/14772000.2015.1099575

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research, 27*(8), 1767-1780. doi:10.1093/nar/27.8.1767

Brockman, S. A., & McFadden, C. S. (2012). The mitochondrial genome of Paraminabea aldersladei (Cnidaria: Anthozoa: Octocorallia) supports intramolecular recombination as the primary mechanism of gene rearrangement in octocoral mitochondrial genomes. G*enome Biology and Evolution, 4*(9), 994-1006. doi:10.1093/gbe/evs074

Brugler, M. R., & France, S. C. (2008). The Mitochondrial Genome of a Deep-Sea Bamboo Coral (Cnidaria, Anthozoa, Octocorallia, Isididae): Genome Structure and Putative Origins of Replication Are Not Conserved Among Octocorals. *Journal of Molecular Evolution, 67*, 125. doi:0.1007/s00239-008-9116-2

Cairns, S. D., & Bayer, F. M. (2005). A review of the genus Primnoa (Octocorallia: Gorgonacea: Primnoidae), with the description of two new species. *Bulletin of Marine Science, 77*(2), 225-256.

Cairns, S. D., & Wirshing, H. H. (2018). A phylogenetic analysis of the Primnoidae (Anthozoa: Octocorallia: Calcaxonia) with analyses of character evolution and a key to the genera and subgenera. *BMC Evolutionary Biology, 18*. doi:10.1186/s12862-018-1182-5

Chang, J. J., Ip, Y. C., Bauman, A. G., & Huang, D. (2020). MinION-in-ARMS: Nanopore Sequencing to Expedite Barcoding of Specimen-Rich Macrofaunal Samples From Autonomous Reef Monitoring Structures. *frontiers in Marince Science, 7*, 2296-7745. doi:10.3389/fmars.2020.00448

De Coster, W. D. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics, 34*(15), 2666–2669. doi:10.1093/bioinformatics/bty149

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., & Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*, 1-11. doi:10.1111/2041-210X.12836

Dias, V., Oliveira, F., Boavida, J., Serão, E. A., Gonçalves, J. M., & Coelho, M. A. (2020). High Coral Bycatch in Bottom-Set Gillnet Coastal Fisheries Reveals Rich Coral Habitats in Southern Portugal. *Frontiers in Marine Science, 7*, 2296-7745. doi:10.3389/fmars.2020.603438

Donath, A., Jühling, F., Al-Arab, M., Bernhart, S. H., Reinhardt, F., Stadler, P. F., Middendorf, M. Bernt, M. (2019). Improved annotation of protein-coding genes boundaries in

metazoan mitochondrial genomes. *Nucleic Acids Research, 47*(20), 10543–10552. doi:10.1093/nar/gkz833

Figueroa, D. F., & Baco, A. R. (2014). Complete mitochondrial genomes elucidate phylogenetic relationships of the deep-sea octocoral families Coralliidae and Paragorgiidae. *Deep Sea Research Part II: Topical Studies in Oceanography, 99*, 83-91. doi:10.1016/j.dsr2.2013.06.001

Fukami, H., & Knowlton, N. (2005). Analysis of complete mitochondrial DNA sequences of three members of the Montastraea annularis coral species complex (Cnidaria, Anthozoa, Scleractinia). *Coral Reefs, 24*,410–417. doi:10.1007/s00338-005-0023-3.

Gaither, M. R. (2010). Preservation of corals in salt-saturated DMSO buffer is superior to ethanol for PCR experiments. *Coral Reefs, 30*, 329–333.

Geller, J. M. (2013). Redesign of PCR primers for mitochondrial Cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular ecology resources, 13*(5), 851–861. doi:10.1111/1755-0998.12138

Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Biological Sciences, 270*, S96–S99. doi:10.1098/rsbl.2003.0

Hoang, D. T. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution, 35*(2), 518–522. doi:10.1093/molbev/msx281

Hooper, L. J. (2021). The complete mitochondrial genome of the pink sea fan, Eunicella verrucosa (Pallas, 1766). *Mitochondrial DNA Part B, 6*(11), 3309-3311. doi:10.1080/23802359.2021.1994899

Hunt, M., De Silva, N., Otto, T. D., Parkhill, J., Keane, J. A., & Harris, S. R. (2015). Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biology, 16*, 294. doi:10.1186/s13059-015-0849-0

Kolmogorov, M. B. (2020). metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods, 17*, 1103–1110. doi:10.1038/s41592-020-00971-x

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology, 37*, 540-546. doi:10.1038/s41587-019-0072-8

Koren, S. &. (2015). One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Current Opinion in Microbiology, 23*, 110-120.

Lee, H. G. (2016). Third-generation sequencing and the future of genomics. *bioRxiv*, 048603. doi:10.1101/048603

Lischer, H. E., & Shimizu, K. (2017). Reference-guided de novo assembly approach improves genome reconstruction for related species. *BMC Bioinformatics, 18*, 474. doi: 10.1186/s12859-017-1911-6

Machado, D. J., Janies, D., Brouwer, C., & Grant, T. (2018). A new strategy to infer circularity applied to four new complete frog mitogenomes. *Ecology and Evolution, 8*(8), 4011-4018. doi:10.1002/ece3.3918

McFadden, C. S. (2006). A molecular phylogenetic analysis of the Octocorallia (Cnidaria: Anthozoa) based on mitochondrial protein-coding sequences. *Molecular phylogenetics and evolution, 41*(3), 513-527. doi:10.1016/j.ympev.2006.06.010

McFadden, C. S., Benayahu, Y., Pante, E., Thoma, J. N., Nevarez, A. P., & France, S. C. (2011). Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources, 11*, 19-31. doi:10.1111/j.1755-0998.2010.02875.x

McFadden, C. S., Brown, A. S., Brayton, C., Hunt, C. B., & van Ofwegen, L. P. (2014). Application of DNA barcoding in biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs, 33*, 275-286. doi:10.1007/s00338-013-1123-0

McFadden, C. S., Haverkort-Yeh, R., Reynolds, A. M., Halàsz, A., Quattrini, A. M., Forsman, Z. H., Toonen, R. J. (2017). Species boundaries in the absence of morphological, ecological or geographical differentiation in the Red Sea octocoral genus Ovabunda (Alcyonacea: Xeniidae). *Molecular Phylogenetics and Evolution, 112*, 174-184. doi:10.1016/j.ympev.2017.04.025

McFadden, C. S., & van Ofwegen, L. P. (2013). Molecular phylogenetic evidence supports a new family of octocorals and a new genus of Alcyoniidae (Octocorallia, Alcyonacea). *ZooKeys, 346*, 59–83. doi:10.3897/zookeys.346.6270

McFadden, C. S., Sánchez, J. A., & France, S. C. (2010). Molecular Phylogenetic Insights into the Evolution of Octocorallia: A Review. *Integrative and Comparative Biology, 50*(3), 389–410. doi:10.1093/icb/icq056

Minh, B. Q. (2013). Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution, 30*(5), 1188-1195. doi:10.1093/molbev/mst024

Minh, B. Q. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular biology and evolution, 37*(5), 1530–1534. doi:10.1093/molbev/msaa015

Niu, W. X. (2020). Characterization of the complete mitochondrial genome sequences of three Merulinidae corals and novel insights into the phylogenetics. *PeerJ, 8*, e8455. doi:0.7717/peerj.8455

Nordström, K. J. (2013). Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. *National Biotechnology, 31*(4), 325-30. doi:10.1038/nbt.2515

Poliseno, A., Feregrino, C., Sartoretto, S., Aurelle, D., Wörheide, G., McFadden, C., & Vargas, S. (2017). Comparative mitogenomics, phylogeny and evolutionary history of Leptogorgia (Gorgoniidae). *Molecular Phylogenetics and Evolution*. 115. doi:10.1016/j.ympev.2017.08.001.

Ragnarsson, S. Á., Burgos, J. M., Kutti, T., Beld, I., Egilsdóttir, H., Arnaud-Haond, S., & Grehan, A. (2016). The Impact of Anthropogenic Activity on Cold-Water Coral. *Springer International Publishing AG*, 989-1023. doi:10.1007/978-3-319-21012-4_27

Roberts, S., & Hirshfield, M. (2004). Deep-sea corals: out of sight, but no longer out of mind. *Frontiers in Ecology and the Environment, 2*(3), 123-130. doi:10.1890/1540-9295(2004)002[0123:DCOOSB]2.0.CO;2

Sánchez, J. A. (2003). Phylogenetic analyses among octocorals (Cnidaria): Mitochondrial and nuclear DNA sequences (lsu-rRNA, 16S and ssu-rRNA, 18S) support two convergent clades of branching gorgonians. *Mol. Phylogenet. Evol, 29*(1), 31-42.

Schroeter, J. C., Maloy, A. P., Rees, C. R., & Bartron, M. L. (2020). Fish mitochondrial genome sequencing: expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA. *Conservation Genetics Resources, 12*, 433-446. doi:10.1007/s12686-019-01111-0

Shearer, T.L., van Oppen, M.J.H., Romano, S.L. & Wörheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology, 11*, 2475-2487. doi:10.1046/j.1365-294X.2002.01652.x

Simison, B. W., Lindberg, R. D., & Boore, L. J. (2006). Rolling circle amplification of metazoan mitochondrial genomes. *Molecular Phylogenetics and Evolution, 39*, 562–567.

Simison, W. B. (2006). olling circle Amplification of Metazoan Mitochondrial Genomes. *Molecular Phylogenetic Evolution, 39*, 562-567. doi:10.1016/j.ympev.2005.11.006

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology, 21*(8), 2045-2050. doi:10.1111/j.1365-294X.2012.05470.x

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183*, 4-18. doi:10.1016/j.biocon.2014.11.019

Uda, K., Komeda, Y., Fujita, T., Iwasaki, N., Bavestrello, G., Giovine, M., Cattano-Vietti, R., & Suzuki, T. (2013). Complete mitochondrial genomes of the Japanese pink coral (Corallium elatius) and the Mediterranean red coral (Corallium rubrum): a reevaluation of the phylogeny of the family Coralliidae based on molecular data. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 8*(3), 209-219. doi:10.1016/j.cbd.2013.05.003

Uda, K., Komeda, Y., Koyama, H., Koda, K., Fujita, T., Iwasaki, N., & Suzuki, T. (2011). Complete mitochondrial genomes of two Japanese precious corals, Paracorallium japonicum and Corallium konojoi (Cnidaria, Octocorallia, Coralliidae): Notable differences in gene arrangement. *Gene, 476*(1-2), 27-37. doi:10.1016/j.gene.2011.01.019

Verma, S., & Gazara, R. K. (2021). Chapter 3 - Next-generation sequencing: an expedition from workstation to clinical applications. *Academic Press, 13*, 29-47. doi: 10.1016/B978-0-323-89824-9.00003-3

Vohsen, S. A., Anderson, K. E., Gade, A. M. *et al.* (2020). Deep-sea corals provide new insight into the ecology, evolution, and the role of plastids in widespread apicomplexan symbionts of anthozoans. *Microbiome 8,* 34. doi:10.1186/s40168-020-00798-w

Wang, Y. Z. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology , 39*, 1348-1368. doi:10.1038/s41587-021-01108-x

3. Supplementary materials

Chapter 2: Octocoral Biodiversity in Portugal: A Barcoding Approach Coupling Long-range PCR and Long-read Sequencing to Assemble Mitochondrial Genomes

Table S2.1. Location, coordinates and collection depth of the samples used in this study.

Species	Subclass/ Order	Depth (m)	Location	Site	Latitude	Longitude	Color morph
Alcyonium acaule	*Octocorallia*	88	Sagres	Coral bycatch	36.96443	-8.9308	Normal
Callogorgia verticillata	*Octocorallia*	292.6	Sagres	Coral bycatch	37.25441	-9.20236	Normal
Corallium rubrum	*Octocorallia*	103.7	Sagres	Coral bycatch	36.93482	-9.01378	Red
Dendrophyllia cornigera	*Scleractinia*	-	Sagres	Coral bycatch	-	-	Normal
Dendrophyllia cornigera	*Scleractinia*	-	Sagres	Coral bycatch	-	-	Orange
Dendrophyllia ramea	*Scleractinia*	-	Sagres	Coral bycatch	-	-	Normal
Dendrophyllia ramea	*Scleractinia*	-	Sagres	Coral bycatch	-	-	Orange & Yellow
Ellisella paraplexauroides	*Octocorallia*	103.7	Sagres	Coral bycatch	36.93482	-9.01378	Normal
Eunicella gazella	*Octocorallia*	97	Sagres	Coral bycatch	37.02367	-9.07925	Normal
Eunicella verrucosa	*Octocorallia*	97.2	Sagres	Coral bycatch	37.03142	-9.051267	Salmon
Favia fragum	*Scleractinia*	0	Cabo Verde	Intertidal platform	14.9	-23.53	Normal
Isidella elongata	*Octocorallia*	448.1	Sagres	Coral bycatch	36.95546	-9.11384	Normal
Leptogorgia sarmentosa	*Octocorallia*	-	Sagres	Coral bycatch	37.11768	-9.07211	Purple & Yellow
Paramuricea sp.	*Octocorallia*	-	Granada	Ponta de la Mona	-	-	yellow
Paramuricea sp.	*Octocorallia*	75-82	Portimão	Offshore - Red coral area	36.94407	-8.63417	Yellow
Paramuricea hirsuta	*Octocorallia*	132.6	Sagres	Coral bycatch	36.93157	-9.08092	Normal
Pennatula rubra	*Octocorallia*	106.6	Sagres	Coral bycatch	36.95440	-9.03265	Normal
Porites astreoides	*Scleractinia*	0	Cabo Verde	Intertidal platform	14.9	-23.53	Normal
Porites porites	*Scleractinia*	0	Cabo Verde	Intertidal platform	14.9	-23.53	Normal
Siderastrea radians	*Scleractinia*	0	Cabo Verde	Intertidal platform	14.9	-23.53	Normal
Spinimuricea atlantica	*Octocorallia*	97.8	Sagres	Coral bycatch	37.04549	-9.05947	Normal
Unknown sp.1	*Octocorallia*	95.1	Sagres	Coral bycatch	36.93880	-9.01217	Normal
Unknown sp.2	*Octocorallia*	95.1	Sagres	Coral bycatch	36.93880	-9.01217	Normal
Unknown sp.3	*Octocorallia*	81.9	Portimão	Offshore - Red coral area	36.94064	-8.63614	White
Unknown sp.4	*Octocorallia*	88	Sagres	Coral bycatch	36.96443	-8.9308	White
Unknown sp.5	*Octocorallia*	85	Sagres	Coral bycatch	36.90428	-9.02567	Normal
Unknown sp.6	*Octocorallia*	107.5	Sagres	Coral bycatch	36.95284	-9.01007	Normal
Unknown sp.7	*Octocorallia*	85	Sagres	Coral bycatch	36.90428	-9.02567	Purple

Table S2.2. Specimens of Octocorallia and Hexacorallia used for the three sequencing runs in this study with the fmol genomic DNA from purified long-range PCR products. Samples that did not have enough genomic DNA (fmol) for the Nanopore protocol recommendations are represented in bold.

Sequencing run	*Species*	Subclass	Order	Family	Amplicon 1 (fmol)	Amplicon 2 (fmol)
1	***Eunicella verrucosa****	Octocorallia	Alcyonacea	Gordoniidae	520	100
2	***Isidella elongata****	Octocorallia	Alcyonacea	Keratoisidinae	**66**	**66**
2	***Paramuricea sp.***	Octocorallia	Alcyonacea	Paramuriceidae	100	200
2	***Alcyonium acaule***	Octocorallia	Alcyonacea	Alcyoniidae	100	200
2	***Eunicella gazella***	Octocorallia	Alcyonacea	Gordoniidae	100	200
2	**Unknown sp.1**	Octocorallia	Unknown	Unknown	100	200
2	**Unknown sp.2**	Octocorallia	Unknown	Unknown	100	200
2	**Unknown sp.3**	Octocorallia	Unknown	Unknown	100	200
2	**Unknown sp.4**	Octocorallia	Unknown	Unknown	100	102
2	**Unknown sp.5**	Octocorallia	Alcyonacea	Ellisellidae	100	200
2	***Spinimuricea atlantica***	Octocorallia	Alcyonacea	Plexauridae	100	200
3	***Corallium rubrum***	Octocorallia	Alcyonacea	Coralliidae	100	100
3	***Callogorgia verticillata***	Octocorallia	Alcyonacea	Primnoidae	100	100
3	***Pennatula rubra***	Octocorallia	Pennatulacea	Pennatulidae	100	100
3	***Leptogorgia sarmentosa***	Octocorallia	Alcyonacea	Alcyoniidae	100	100
3	**Unknown sp.6**	Octocorallia	Alcyonacea	Gorgoniidae	100	100
3	***Ellisella paraplexauroides***	Octocorallia	Alcyonacea	Ellisellidae	100	100
3	***Paramuricea hirsuta***	Octocorallia	Alcyonacea	Acanthogorgiidae	100	100
3	***Isidella elongata****	Octocorallia	Alcyonacea	Keratoisididae	100	100
3	***Paramuricea sp.***	Octocorallia	Alcyonacea	Plexauridae	100	100
3	**Unknown sp.7**	Octocorallia	Alcyonacea	Plexauridae	100	100
3	***Eunicella verrucosa****	Octocorallia	Alcyonacea	Gorgoniidae	100	100
3	***Dendrophyllia ramea***	Hexacorallia	Scleractinia	Dendrophylliidae	**85**	**85**
3	***Dendrophyllia cornigera***	Hexacorallia	Scleractinia	Dendrophylliidae	100	100
3	***Dendrophyllia cornigera***	Hexacorallia	Scleractinia	Dendrophylliidae	100	100
3	***Dendrophyllia ramea***	Hexacorallia	Scleractinia	Dendrophylliidae	**94**	**94**
3	***Porites porites***	Hexacorallia	Scleractinia	Poritidae	100	100
3	***Porites astreoides***	Hexacorallia	Scleractinia	Poritidae	100	100
3	***Siderastrea radians***	Hexacorallia	Scleractinia	Siderastreidae	**20**	**20**
3	***Favia fragum***	Hexacorallia	Scleractinia	Faviidae	**48**	**48**

*Species that were sequenced twice because we were unable to assemble the mitogenomes using the data from the previous sequencing run.

Table S2.3. All available complete mitogenomes that were downloaded from GenBank from the Octocorallia Subclass including accession number and corresponding mitogenome size (bp). Species in bold exhibit different gene arrangements than the most encountered gene order. Underlined species were used to reconstruct the phylogeny tree of octocorals.

Subclass	Order	Family	Species	Accession No.	Mitogenome size (bp)
Octocorallia	**Alcyonacea**	**Isididae**	***Acanella arbuscula***	**EF_72731**	**18,616**
Octocorallia	**Alcyonacea**	**Alcyoniidae**	***Anthomastus sp.***	**KM_015352**	**18,715**
Octocorallia	**Pennatulacea**	**Anthoptilidae**	***Anthoptilum grandiflorum***	**NC_044086**	**18,583**
Octocorallia	**Pennatulacea**	**Anthoptilidae**	***Anthoptilum sp.***	**MK_919656**	**18,850**
Octocorallia	Alcyonacea	Briareidae	*Briareum asbestinum*	DQ_640649	18,632
Octocorallia	Alcyonacea	Acanthogorgiidae	*Calicogorgia granulosa*	NC_023345	20,246
Octocorallia	Alcyonacea	Primnoidae	*Callogorgia cf. gracilis*	MH_719202	18,937
Octocorallia	Alcyonacea	Clavulariidae	*Carijoa riisei*	NC_048963	18,714
Octocorallia	Pennatulacea	Pennatulacea	*Cavernularia obesa*	OK_149222	18641
Octocorallia	Alcyonacea	Coralliidae	*Corallium elatius*	NC_022804	18,969
Octocorallia	Alcyonacea	Coralliidae	*Corallium konojoi*	NC_015406	18,969
Octocorallia	**Alcyonacea**	**Coralliidae**	***Corallium rubrum***	**NC_022864**	**18,915**
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya castanea*	NC_023343	18,907
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya gigantea*	NC_013573	18,842
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya mollis*	NC_020456	18,844
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya putteri*	NC_036022	18,853
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya suensoni*	NC_022809	18,851
Octocorallia	Alcyonacea	Nephtheidae	*Dendronephthya suensoni*	GU_047878	18,885
Octocorallia	Pennatulacea	Protoptilidae	*Distichoptilum gracile*	NC_044077	19,171
Octocorallia	Alcyonacea	Plexauridae	*Echinogorgia complexa*	NC_020457	19,445
Octocorallia	**Alcyonacea**	**Plexauridae**	***Echinogorgia reticulata***	**NC_059856**	**19,445**
Octocorallia	**Alcyonacea**	**Ellisellidae**	***Ellisella ceratophyta***	**KY_230380**	**18814**
Octocorallia	Alcyonacea	Gordoniidae	*Eugorgia mutabilis*	NC_035665	19,157
Octocorallia	Alcyonacea	Gordoniidae	*Eunicella albicans*	NC_035666	19,175
Octocorallia	Alcyonacea	Gordoniidae	*Eunicella albicans*	KY_556407	19,175
Octocorallia	Alcyonacea	Gordoniidae	*Eunicella cavolinii*	NC_035667	19,316
Octocorallia	Alcyonacea	Gordoniidae	*Eunicella verrucosa*	MW_588805	19,267
Octocorallia	Alcyonacea	Plexauridae	*Euplexaura crassa*	NC_020458	18,674
Octocorallia	Pennatulacea	Anthoptilidae	*Funiculina quadrangularis*	NC_044078	18906
Octocorallia	Pennatulacea	Halipteridae	*Halipteris cf. finmarchica*	MK919659	18,513
Octocorallia	**Helioporacea**	**Helioporidae**	***Heliopora coerulea***	**NC_020375**	**18,957**
Octocorallia	Alcyonacea	Clavuriidae	*Incrustatus comauensis*	MT_254531	18977
Octocorallia	**Alcyonacea**	**Ellisellidae**	***Junceella fragilis***	**NC_024181**	**18,724**
Octocorallia	**Alcyonacea**	**Isididae**	***Keratoisidinae sp.***	**EF_622534**	**18,923**
Octocorallia	Alcyonacea	Isididae	*Keratoisidinae sp.*	NC_010764	18923

Octocorallia	Pennatulacea	Kophobelemnidae	*Kophobelemnon sp. 1*	MK919660	18,883
Octocorallia	Pennatulacea	Kophobelemnidae	*Kophobelemnon sp. 3*	MK919661	19,109
Octocorallia	Pennatulacea	Kophobelemnidae	*Kophobelemnon sp. 4*	MK919662	19,130
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia alba*	NC_035669	18,848
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia alba*	KY_559410	18,848
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia capverdensis*	NC_035663	18,722
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia cf. palma*	KY_559406	18,731
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia gaini*	KY_559404	19,682
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia hebes*	MN_052676	19,247
Octocorallia	Alcyonacea	Alcyoniidae	*Leptogorgia sarmentosa*	NC_035670	18722
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia sp.*	KY_559412	18,849
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia virgulata*	MK_301586	18,845
Octocorallia	Alcyonacea	Gordoniidae	*Leptogorgia virgulata*	MK_301590	18,845
Octocorallia	Alcyonacea	Plexauridae	*Muricea crassa*	NC_029697	19,586
Octocorallia	Alcyonacea	Plexauridae	*Muricea purpurea*	NC_029698	19,358
Octocorallia	Alcyonacea	Primnoidae	*Narella hawaiinensis*	KM_015351	18,838
Octocorallia	Alcyonacea	Gordoniidae	*Pacifigorgia cairnsi*	NC_035668	19,156
Octocorallia	**Alcyonacea**	**Coralliidae**	***Paracorallium japonicum***	**NC_015405**	**18,913**
Octocorallia	Alcyonacea	Paragorgiidae	*Paragorgia coralloides*	KF_785800	19,016
Octocorallia	Alcyonacea	Alcyoniidae	*Paraminabea aldersladei*	NC_018790	19,886
Octocorallia	Alcyonacea	Paramuriceidae	*Paramuricea clavata*	NC_034749	18,669
Octocorallia	Alcyonacea	Paramuriceidae	*Paramuricea macrospina*	NC_034750	18,921
Octocorallia	Pennatulacea	Pennatulidae	*Pennatula aculeata*	NC_044087	18,715
Octocorallia	Pennatulacea	Pennatulidae	*Pennatula aculeata*	MK_919664	18,715
Octocorallia	Pennatulacea	Pennatulidae	*Pennatula cf. aculeata*	MK919664	18,715
Octocorallia	Pennatulacea	Pennatulidae	*Pennatula cf. Inflata*	MK919666	19,127
Octocorallia	Pennatulacea	Pennatulidae	*Pennatula grandis*	NC_044088	18,973
Octocorallia	Alcyonacea	Coralliidae	*Pleurocorallium elatius*	AB_700135	18,970
Octocorallia	Alcyonacea	Coralliidae	*Pleurocorallium konojoi*	AB_595190	18,969
Octocorallia	Alcyonacea	Primnoidae	*Plumarella adhaerans*	NC_046480	19036
Octocorallia	Alcyonacea	Primnoidae	*Plumarella adhaerans*	NC_046480	19,036
Octocorallia	Alcyonacea	Primnoidae	*Plumarella spinosa*	NC_046465	19,037
Octocorallia	Pennatulacea	Protoptilidae	*Protoptilum carpenteri*	NC_044089	18,729
Octocorallia	Alcyonacea	Gordoniidae	*Pseudopterogorgia bipinnata*	NC_008157	18,733
Octocorallia	**Pennatulacea**	**Renillidae**	***Renilla muelleri***	**NC_018378**	**18,643**
Octocorallia	Alcyonacea	Alcyoniidae	*Sarcophyton trocheliophorum*	MK994517	18,508
Octocorallia	Alcyonacea	Nephtheidae	*Scleronephthya gracillimum*	NC_023344	18,950
Octocorallia	Alcyonacea	Paragorgiidae	*Sibogagorgia cauliflora*	KM_015354	19,030
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia acuta*	MW_987591	18,730
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia ceramensis*	NC_044122	18,740
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia cf. cruciata*	NC_034318	18,730

Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia humilis*	OK_641586	18,743
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia maxima*	MN485891	18,730
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia peculiaris*	JX_023274	18,742
Octocorallia	Alcyonacea	Alcyoniidae	*Sinularia penghuensis*	MW_256412	18,730
Octocorallia	Pennatulacea	Virgulariidae	*Stylatula elongata*	JX_023275	18,733
Octocorallia	Alcyonacea	Clavulariidae	*Telestula humilis*	MT_254527	18,740
Octocorallia	Alcyonacea	Clavulariidae	*Telestula septentrionalis*	MT_254532	18,751
Octocorallia	Pennatulacea	Umbellulidae	*Umbellula huxleyi*	NC_044090	18,927
Octocorallia	**Pennatulacea**	**Umbellulidae**	***Umbellula sp. 1***	**MK919669**	**18,714**
Octocorallia	Pennatulacea	Umbellulidae	*Umbellula sp. 2*	MK919670	18,766
Octocorallia	Pennatulacea	Anthoptilidae	*Virgularia mirabilis*	MK_919673	18,770

Table S2.4. All available complete mitogenomes that were downloaded from GenBank from the Hexacorallia Subclass including accession number and corresponding mitogenome size (bp).

Subclass	Order	Family	Species	Accession No.	Mitogenome size (bp)
Hexacorallia	Scleractinia	Lobophylliidae	*Acanthastrea maxima*	FO_904931	18,168
Hexacorallia	Scleractinia	Acroporidae	*Acropora aculeus*	NC_029251	18,528
Hexacorallia	Scleractinia	Acroporidae	*Acropora aspera*	NC_022827	18,479
Hexacorallia	Scleractinia	Acroporidae	*Acropora digitifera*	NC_022830	18,479
Hexacorallia	Scleractinia	Acroporidae	*Acropora florida*	NC_022828	18,365
Hexacorallia	Scleractinia	Acroporidae	*Acropora horrida*	NC_022825	18,480
Hexacorallia	Scleractinia	Acroporidae	*Acropora humilis*	NC_022823	18,479
Hexacorallia	Scleractinia	Acroporidae	*Acropora hyacinthus*	NC_022826	18,566
Hexacorallia	Scleractinia	Acroporidae	*Acropora muricata*	NC_022824	18,481
Hexacorallia	Scleractinia	Acroporidae	*Acropora nasuta*	NC_022831	18,481
Hexacorallia	Scleractinia	Acroporidae	*Acropora robusta*	NC_022833	18,480
Hexacorallia	Scleractinia	Acroporidae	*Acropora tenuis*	AF_338425	18,338
Hexacorallia	Scleractinia	Acroporidae	*Acropora valida*	MH_141598	18,385
Hexacorallia	Scleractinia	Acroporidae	*Acropora yongei*	NC_022829	18,342
Hexacorallia	Actiniaria	Actiniidae	*Actinia equina*	MH_545699	20,690
Hexacorallia	Actiniaria	Actiniidae	*Actinia tenebrosa*	NC_044902	20,691
Hexacorallia	Scleractinia	Agariciidae	*Agaricia humilis*	DQ_643831	18,735
Hexacorallia	Actiniaria	Aiptasiidae	*Aiptasia pulchella*	NC_022265	19,791
Hexacorallia	Actiniaria	Aliciidae	*Alicia sansibarensis*	KR_051001	19,575
Hexacorallia	Scleractinia	Poritidae	*Alveopora japonica*	NC_040136	18,144
Hexacorallia	Scleractinia	Poritidae	*Alveopora sp.*	KJ_634271	18,146
Hexacorallia	Corallimorpharia	Discosomatidae	*Amplexidiscus fenestrafer*	NC_027101	20,188
Hexacorallia	Scleractinia	Acroporidae	*Anacropora matthai*	AY_903295	17,888
Hexacorallia	Actiniaria	Actiniidae	*Anemonia majano*	KY_860670	19,545

Hexacorallia	Actiniaria	Actiniidae	*Anemonia viridis*	KY_860669	20,108
Hexacorallia	Actiniaria	Actinostolidae	*Antholoba achates*	KR_051002	17,816
Hexacorallia	Actiniaria	Actiniidae	*Anthopleura midori*	KT_989511	20,039
Hexacorallia	Antipatharia	Antipathidae	*Antipathes cf. dichotoma*	MT_318841	19,969
Hexacorallia	Scleractinia	Rhizangiidae	*Astrangia sp.*	DQ_643832	14,853
Hexacorallia	Scleractinia	Acroporidae	*Astreopora explanata*	NC_024090	18,146
Hexacorallia	Scleractinia	Acroporidae	*Astreopora myriophthalma*	NC_024092	18,106
Hexacorallia	Antipatharia	Schizopathidae	*Bathypathes sp.*	MT_318842	17,700
Hexacorallia	Actiniaria	Actiniidae	*Bolocera sp.*	KU_507297	19,463
Hexacorallia	Actiniaria	Actiniidae	*Bolocera tuediae*	NC_022470	19,143
Hexacorallia	Antipatharia	Antipathidae	*Cirrhipathes lutkeni*	JX_023266	20,448
Hexacorallia	Scleractinia	Faviidae	*Colpophyllia natans*	DQ_643833	16,906
Hexacorallia	Corallimorpharia	Corallimorphidae	*Corallimorphus profundus*	KP_938440	20,488
Hexacorallia	Corallimorpharia	Corallimorphidae	*Corynactis californica*	NC_027102	20,715
Hexacorallia	Antipatharia	Schizopathidae	*Dendrobathypathes sp.*	MT_318845	17,687
Hexacorallia	Scleractinia	Dendrophylliidae	*Dendrophyllia arbuscula*	NC_027590	19,069
Hexacorallia	Scleractinia	Dendrophylliidae	*Dendrophyllia cribrosa*	NC_026026	19,072
Hexacorallia	Scleractinia	Caryophylliidae	*Desmophyllum dianthus*	NC_034275	16,310
Hexacorallia	Scleractinia	Caryophylliidae	*Desmophyllum pertusum*	NC_015143	16,150
Hexacorallia	Scleractinia	Faviidae	*Dipsastraea favus*	MK_516277	17,054
Hexacorallia	Scleractinia	Faviidae	*Dipsastraea rotumana*	MH_119077	16,466
Hexacorallia	Corallimorpharia	Discosomatidae	*Discosoma nummiforme*	NC_027100	20,925
Hexacorallia	Corallimorpharia	Discosomatidae	*Discosoma sp.*	NC_008071	20,908
Hexacorallia	Scleractinia	Lobophylliidae	*Echinophyllia aspera*	NC_040169	17,697
Hexacorallia	Actiniaria	Actiniidae	*Entacmaea quadricolor*	NC_049066	20,960
Hexacorallia	Zoantharia	Epizoanthidae	*Epizoanthus illoricatus*	MN_873588	20,447
Hexacorallia	Scleractinia	Caryophylliidae	*Euphyllia ancora*	NC_024672	18,875
Hexacorallia	Scleractinia	Merulinidae	*Favites abdita*	NC_035879	17,825
Hexacorallia	Scleractinia	Merulinidae	*Favites pentagona*	NC_034916	18,006
Hexacorallia	Scleractinia	Fungiacyathidae	*Fungiacyathus stephanus*	NC_015640	19,381
Hexacorallia	Scleractinia	Euphylliidae	*Galaxea fascicularis*	NC_029696	18,751
Hexacorallia	Scleractinia	Poritidae	*Goniopora columna*	NC_015643	18,766
Hexacorallia	Scleractinia	Poritidae	*Goniopora djiboutiensis*	MH_746816	18,765
Hexacorallia	Scleractinia	Poritidae	*Goniopora lobata*	MN_795054	18,770
Hexacorallia	Actiniaria	Halcampoididae	*Halcampoides purpurea*	NC_027612	18,038
Hexacorallia	Actiniaria	Hormathiidae	*Hormathia digitata*	NC_022471	18,754
Hexacorallia	Scleractinia	Merulinidae	*Hydnophora exesa*	MH_086217	17,790
Hexacorallia	Zoantharia	Hydrozoanthidae	*Hydrozoanthus gracilis*	MN_873589	20,689

Hexacorallia	Scleractinia	Acroporidae	*Isopora palifera*	NC_024091	18,725
Hexacorallia	Scleractinia	Acroporidae	*Isopora togianensis*	NC_024089	18,637
Hexacorallia	Actiniaria	Actiniidae	*Isosicyonis striata*	KR_051006	19,001
Hexacorallia	Antipatharia	Antipathidae	*Leiopathes cf. glaberrima*	MT_318846	21,669
Hexacorallia	Antipatharia	Antipathidae	*Leiopathes expansa*	MT_318847	21,653
Hexacorallia	Scleractinia	Pocilloporidae	*Madracis mirabilis*	EU_400212	16,951
Hexacorallia	Scleractinia	Oculinidae	*Madrepora oculata*	NC_018364	15,841
Hexacorallia	Actiniaria	Metridiidae	*Metridium senile*	NC_000933	17,444
Hexacorallia	Zoantharia	Microzoanthidae	*Microzoanthus occultus*	MN_873590	21,656
Hexacorallia	Scleractinia	Merulinidae	*Montastraea faveolata*	NC_007226	16,138
Hexacorallia	Scleractinia	Acroporidae	*Montipora aequituberculata*	NC_037359	17,886
Hexacorallia	Scleractinia	Acroporidae	*Montipora cactus*	AY_903296	17,887
Hexacorallia	Scleractinia	Acroporidae	*Montipora efflorescens*	NC_040137	17,886
Hexacorallia	Scleractinia	Mussidae	*Mussa angulosa*	NC_008163	17,245
Hexacorallia	Antipatharia	Antipathidae	*Myriopathes japonica*	NC_027667	17,733
Hexacorallia	Zoantharia	Nanozoanthidae	*Nanozoanthus harenaceus*	NC_046402	18,577
Hexacorallia	Actiniaria	Edwardsiidae	*Nematostella sp.*	DQ_643835	16,389
Hexacorallia	Zoantharia	Neozoanthidae	*Neozoanthus aff. uchina*	MN_873592	21,804
Hexacorallia	Scleractinia	Merulinidae	*Orbicella annularis*	NC_007224	16,138
Hexacorallia	Scleractinia	Merulinidae	*Orbicella franksi*	NC_007225	16,138
Hexacorallia	Zoantharia	Sphenopidae	*Palythoa heliodiscus*	NC_035579	20,841
Hexacorallia	Zoantharia	Sphenopidae	*Palythoa mizigama*	MN_873594	21,104
Hexacorallia	Zoantharia	Sphenopidae	*Palythoa mutuki*	MN_873595	21,278
Hexacorallia	Antipatharia	Schizopathidae	*Parantipathes hirondelle*	MT_318850	17,734
Hexacorallia	Antipatharia	Schizopathidae	*Parantipathes sp.*	MT_318851	17,734
Hexacorallia	Zoantharia	Parazoanthidae	*Parazoanthus elongatus*	NC_046405	21,148
Hexacorallia	Zoantharia	Parazoanthidae	*Parazoanthus swiftii*	NC_046475	21,499
Hexacorallia	Scleractinia	Agariciidae	*Pavona clavus*	NC_008165	18,315
Hexacorallia	Scleractinia	Agariciidae	*Pavona decussata*	NC_026527	18,378
Hexacorallia	Actiniaria	Phymanthidae	*Phymanthus crucifer*	KR_051007	19,727
Hexacorallia	Scleractinia	Merulinidae	*Platygyra carnosa*	NC_020049	16,463
Hexacorallia	Scleractinia	Plesiastreidae	*Plesiastrea versipora*	MH_025639	15,320
Hexacorallia	Scleractinia	Pocilloporidae	*Pocillopora damicornis*	NC_009797	17,415
Hexacorallia	Scleractinia	Pocilloporidae	*Pocillopora damicornis*	EU_400213	17,425
Hexacorallia	Scleractinia	Pocilloporidae	*Pocillopora eydouxi*	NC_009798	17,422
Hexacorallia	Scleractinia	Caryophylliidae	*Polycyathus sp.*	NC_015642	15,345
Hexacorallia	Scleractinia	Poritidae	*Porites fontanesii*	NC_037434	18,658
Hexacorallia	Scleractinia	Poritidae	*Porites harrisoni*	NC_037435	18,630
Hexacorallia	Scleractinia	Poritidae	*Porites lobata*	NC_030186	18,647

Hexacorallia	Scleractinia	Poritidae	*Porites lutea*	NC_029695	18,646
Hexacorallia	Scleractinia	Poritidae	*Porites okinawensis*	NC_015644	18,647
Hexacorallia	Scleractinia	Poritidae	*Porites panamensis*	NC_024182	18,628
Hexacorallia	Scleractinia	Poritidae	*Porites porites*	DQ_643837	18,648
Hexacorallia	Actiniaria	Gonactiniidae	*Protanthea simplex*	MH_500774	17,134
Hexacorallia	Corallimorpharia	Corallimorphidae	*Pseudocorynactis sp.*	KP_938437	21,239
Hexacorallia	Scleractinia	Siderastreidae	*Pseudosiderastrea formosa*	NC_026530	19,475
Hexacorallia	Scleractinia	Siderastreidae	*Pseudosiderastrea tayami*	NC_026531	19,475
Hexacorallia	Corallimorpharia	Discosomatidae	*Rhodactis indosinensis*	NC_027103	20,100
Hexacorallia	Corallimorpharia	Discosomatidae	*Rhodactis mussoides*	NC_027104	20,826
Hexacorallia	Corallimorpharia	Discosomatidae	*Rhodactis sp.*	DQ_640647	20,093
Hexacorallia	Corallimorpharia	Ricordeidae	*Ricordea florida*	NC_008159	21,376
Hexacorallia	Corallimorpharia	Ricordeidae	*Ricordea yuma*	KP_938441	22,015
Hexacorallia	Actiniaria	Sagartiidae	*Sagartia ornata*	KR_051008	17,446
Hexacorallia	Zoantharia	Parazoanthidae	*Savalia savaglia*	DQ_825686	20,764
Hexacorallia	Scleractinia	Pocilloporidae	*Seriatopora caliendrum*	EF_633601	17,010
Hexacorallia	Scleractinia	Pocilloporidae	*Seriatopora hystrix*	EF_633600	17,059
Hexacorallia	Antipatharia	Cladopathidae	*Sibopathes cf. macrospina*	MT_318853	17,734
Hexacorallia	Scleractinia	Siderastreidae	*Siderastrea radians*	DQ_643838	19,387
Hexacorallia	Scleractinia	Caryophylliidae	*Solenosmilia variabilis*	NC_025472	15968
Hexacorallia	Zoantharia	Sphenopidae	*Sphenopus marsupialis*	NC_046406	21,199
Hexacorallia	Antipatharia	Schizopathidae	*Stauropathes arctica*	MT_318854	17,700
Hexacorallia	Antipatharia	Schizopathidae	*Stauropathes cf punctata*	MT_318855	17,690
Hexacorallia	Antipatharia	Antipathidae	*Stichopathes abyssicola*	MT_318856	19968
Hexacorallia	Antipatharia	Antipathidae	*Stichopathes lutkeni*	NC_018377	20,448
Hexacorallia	Antipatharia	Antipathidae	*Stichopathes sp.*	MT_318857	19,839
Hexacorallia	Scleractinia	Pocilloporidae	*Stylophora pistillata*	EU_400214	17,177
Hexacorallia	Antipatharia	Myriopathidae	*Tanacetipathes thamnea*	MN_265369	17,712
Hexacorallia	Antipatharia	Schizopathidae	*Telopathes sp.*	MT_318858	17,681
Hexacorallia	Antipatharia	Cladopathidae	*Trissopathes cf. tetracada*	MT_318840	18,468
Hexacorallia	Scleractinia	Dendrophylliidae	*Tubastraea coccinea*	JQ_290078	19,070
Hexacorallia	Scleractinia	Dendrophylliidae	*Tubastraea coccinea*	KX_024566	19,094
Hexacorallia	Scleractinia	Dendrophylliidae	*Tubastraea tagusensis*	NC_030352	19,094
Hexacorallia	Scleractinia	Dendrophylliidae	*Turbinaria peltata*	NC_024671	18,966
Hexacorallia	Antipatharia	Antipathidae	*Tylopathes sp.*	MT_318859	17,679
Hexacorallia	Actiniaria	Actiniidae	*Urticina eques*	NC_022469	20,458
Hexacorallia	Zoantharia	Zoanthidae	*Zoanthus cf. pulchellus*	MN_873599	21,139
Hexacorallia	Zoantharia	Zoanthidae	*Zoanthus cf. sociatus*	MN_873600	21,140

Hexacorallia	Zoantharia	Zoanthidae	*Zoanthus sansibaricus*	NC_035578	20,972
Hexacorallia	Zoantharia	Zoanthidae	*Zoanthus sociatus*	NC_046476	21,202
Hexacorallia	Zoantharia	Zoanthidae	*Zoanthus sp.*	MN_873602	21,140

Table S2.5. Descriptive statistics of the assemblies for each species containing the sizes of clean reads used before assembly, the read length minimum overlap, and output information of the assembly.

Barcode	Species	Clean reads (Kb)	Flye minimum overlap	Total length after assembly (bp)	No. Contigs	N50	Contigs representing amplicon 1&2	Contig length (bp)	Contig coverage	Overlap length (bp)	Size after polishing(bp)	Final MtDNA (bp)
3	*Alcyonium sp.*	186460	7000	48526	5	8781	contig 2	8129	3742	1	8133	18677
							contig 5	11154	15139	565	11159	
4	*Eunicella gazella*	164214	1000	19882	2	11744	contig 1	8138	4280	24	8146	19317
							contig 2	11744	12221	555	11751	
5A	Unknown sp.1	185408	1000	22163	3	11698	contig 1	11698	12867	32	11707	19246
							contig 3	8145	6154	580	8152	
6	Unknown sp.2	2722631	7000	70196	8	9084	contig 5	11158	10	572	11161	19656
							contig 1	9084	7704	24	9091	
7	Unknown sp.3	63338	8000	20236	2	11154	contig 1	11154	4794	16	11160	19655
							contig 2	9082	1683	579	9090	
8	Unknown sp.4	96404	3000	20231	2	11154	contig 1	11154	4174	26	11161	19654
							contig 2	9077	6377	565	9084	
9	Unknown sp.5	183358	3000	47069	7	8130	contig 1	11145	14002	30	11150	19655
							contig 2	9091	4468	564	9099	
10A	*Spinimuricea atlantica*	236878	6000	51535	6	8597	contig 1	11133	16222	16	11139	18671
							contig 5	8116	8285	571	8119	
5B	*Corallium rubrum*	481628	5000	165430	21	8850	contig 2	18897	19520	0	18915	18915
10B	*Callogorgia verticillata*	377389	4000	113375	19	6336	contig 5	10724	24166	16	10728	18919
							contig 13	8248	15713	43	8250	
11	*Pennatula rubra*	397337	4000	126142	17	8127	contig 16	12096	205390	17	11268	18727
							contig 7	7522	21197	50	7526	
12	*Leptogorgia sarmentosa*	209166	2000	114928	17	8935	contig 6	11228	5728	50	11224	18722
							contig 9	8137	10660	50	8141	
13	Unknown sp.6	414592	4000	116806	18	8113	contig 17	8113	6	31	8117	18672
							contig 22	11137	22432	552	11138	
14	*Ellisella paraplexauroides*	450910	7000	93678	12	8148	contig 10	11254	23959	17	11257	18814
							contig 15	8148	3	575	8149	
17	*Paramuricea sp.*	315177	4000a	60126	8	8891	contig 7	11157	19311	3647	11158	18669
							contig 8	8267	8709	1902	6957	
							contig 9	3077	8796	573	3078	
18	Unknown sp.7	391548	3000a	127254	19	8217	contig 1	8125	37	560	8128	18670
							contig 2	6388	20756	4426	6388	
							contig 13	9194	21418	58	9196	
19	*Eunicella verrucosa*	728013	5000	166743	18	10247	contig 1	11670	205	38	11672	19266
							contig 20	8151	559	521	8153	

Table S2.6. IQ-TREE gene alignment analysis information includes the number of distinct patterns, singleton sites (un-informative variable), and parsimony-informative sites for each gene region.

Genes	Alignment sequences	Alignment columns	Distinct patterns	Singleton sites	Parsimony-informative
Atp6	93	690	269	53	238
Atp8	93	207	85	24	65
Cytb	93	861	355	57	292
MutS	93	2268	1243	223	1258
Nad1	93	663	283	232	214
Nad2	93	1059	405	97	343
Nad3	93	330	116	27	92
Nad4L	93	291	93	19	81
Nad4	93	1407	507	106	464
Nad5	93	1779	655	150	571
Nad6	93	423	171	29	145
cox1	93	1428	491	183	435
cox2	93	729	271	58	237
cox3	93	780	250	49	210
r12S	93	563	143	43	119
r16S	93	902	247	77	212