GENOME-RESOLVED METAGENOMIC ANALYSIS OF A CONSERVED SPIROCHETE SYMBIONT IN PYRROLOIMINOQUINONE PRODUCING SPONGES

A THESIS SUBMITTED IN FULFILMENT OF THE DEGREE OF

MASTER OF SCIENCE

at

RHODES UNIVERSITY

RHODES UNIVERSITY
Where leaders learn

By

Gabriella Malaika Solomons

Orchid ID: 0000-0002-3278-6293

November 2024

ACKNOWLEDGEMENTS

To Professor Dorrington, my supervisor since 2021, there aren't enough words to thank you and describe how grateful I am for you as a mentor and supervisor. For the opportunities you have granted me, the financial assistance you have graciously provided and the ways you have moulded me as a researcher as well as a woman in such a male dominated field. Thank you for believing in me, encouraging me and motivating me to strive for excellence and to produce research of a globally recognized standard. To Dr Samantha Waterworth, my (unofficial) co-supervisor since 2021, I am eternally grateful for all the afternoons and nights you've sacrificed to help me get here. To work under two incredible women at the top of their respective fields has been no small feat and I am eternally in awe of both of you. Thank you for transforming this shy, anxious, imposter-syndrome ridden honour's student into the confident, hard-working researcher I am today. Thank you for believing in me, constantly encouraging me and seeing me through my first publication as a scientist.

To the current and former members of Lab417, thank you for your support and expertise in helping me get here. Luthando and Wakisa, who (under the guidance of Dr Samantha Waterworth) thank you for getting me through the first year of my Master's as someone with zero background knowledge in the field of genomics and for continuing to be the core of our genomics unit. Idris who has been a constant source of wisdom, kindness and expertise, thank you for your input as a researcher and your support as a friend. To Asi, Tumi, Lesedi and Dele thank you for providing laughter and encouragement during the challenging times. To Michelle, Tarryn Swart, Tarryn Potts, Sis B and Roxanne thank you for working incredibly hard behind the scenes and being the backbone of the lab. Each and every one of you for being such wonderful colleagues and pleasures to work with.

To my dear family, Mommy, Kimmy, Gift, Aunty Sus, thank you for carrying me through this journey. Thank you for your prayers, your financial assistance and unconditional love since the very beginning of my career. Mommy and Kimmy, thank you for encouraging me to strive for the very top of this academic journey and believing that I could, and I should make it to PhD level. A special thank you to my sister Kimmy who was always cheering me on willing to let me explain my work despite my inability to successfully translate it into layman's terms. To my church and hockey friends, thank you for keeping me sane, providing safe spaces and

healthy ways of juggling my academic, social and spiritual lives. To Sasha and Nhlakanipho, thank you for your constant prayers, encouragement, comfort and love that helped me through the toughest of months where the pressures of this degree felt too difficult to share. I am truly blessed to have met you both so early in my life.

I would like to acknowledge the National Research Facility for funding my three years of studies (Honour’s as well as Master’s) as well as Professor’s Quarraisha and Salim Abdool Karim for graciously awarding me with a prize in my Honour’s year that provided additional financial support to fund my Master’s degree. A special thank you to the South African Institute for Aquatic Biodiversity (SAIAB), the Central Analytics Facility (CAF) in Stellenbosch and National Integrated Cyberinfrastructure System Centre for High Performance Computing (NICIS-CHPC) for allowing the use of their equipment towards the success of my project. Moreover, I would like to thank Dr Gwynneth Matcher and Dr Shirley Parker-Nance, Ms Alvera Vorster and Mr Carel van Heerden who provided access to the platforms provided by these above-mentioned facilities.

Last and certainly not least, I would like to thank the Almighty Lord Jesus Christ for His grace and provision and for constantly reminding me that everything is possible according to His will and at His own time.

Table of Contents

List of figures

List of tables

List of abbreviations

AAI – Average amino acid identity

ABC transporter – ATP-synthase binding cassette transporter

ANI – average nucleotide identity

ANOSIM – Analysis of similarities

AntiSMASH – Antibiotics and Secondary Metabolite Analysis Shell

ATP – Adenosine triphosphate

Barrnap – Bacterial ribosomal RNA predictor

BLAST – Basic Local Alignment Search Tool

BGC – Biosynthetic gene cluster

Bp – Base pairs

CAF – Central Analytics Facility

Contig – Contiguous sequence

dNTP – deoxyribonucleotide triphosphate

gDNA – genomic DNA

GTDB-Tk – Genome Taxonomy Database Toolkit

HMA – High microbial abundance

iTOL – Interactive Tree of Life

KEGG – Kyoto Encyclopaedia for Genes and Genomes

KO – Kegg Orthology

LMA – Low microbial abundance

MAG – Metagenome-assembled genome

MIMAG – Minimum information about a MAG

NCBI – National Centre for Biotechnology Information

NGS – Next Generation Sequencing

NMDS – Non-metric multidimensional scaling

OMA – Orthologous Matrix

ORF – Open Reading Frame

OTU – Operational taxonomic unit

PCR – Polymerase chain reaction

QUAST – Quality assessment tool

RiPP – Ribosomally synthesized and post-translationally modified peptides

RNA – Ribonucleotide triphosphate

rRNA – ribosomal RNA

SAUL – Sponge-associated unclassified lineage

SPAdes – St. Petersburg Genome Assembler

UPGMA – Unweighted pair group method with arithmetic mean

Abstract

Symbiotic relationships between sponges and their associated bacteria represent one of the oldest examples of symbiosis on Earth. Bacterial symbionts are acquired by their sponge hosts for a number of reasons including the uptake of nutrients, acquisition of carbon, nitrogen and phosphorous, the recycling of waste and other metabolic byproducts, the production of amino acids, stabilizing of the host skeletal system and facilitating chemical defensive strategies against pathogens and predators through the production of bioactive secondary metabolites. Marine sponges of the *Latrunculiidae* family are the most prolific producers of the bioactive secondary metabolites known as pyrroloiminoquinones, a class of alkaloids with a broad spectrum of biological activities and considerable potential as future drug leads. The microbial communities associated with several genera within this family of sponges are highly conserved and dominated by two bacterial symbionts – a broad-host range Tethybacterales symbiont and a Spirochete symbiont, exclusive to sponges of the *Tsitsikamma* genus and *Cyclacanthia bellae* species. Dominant spirochete communities have only been reported in the microbiomes of *Clathrina clathrus* sponges, where they play a role in calcification of the sponge skeletal tissue. This study was aimed at investigating the nature of the symbiotic relationship between latrunculid sponges and their conserved spirochete symbiont with respect to the production of pyrroloiminoquinones. These spirochetes were shown to be newly acquired symbionts that are currently undergoing genome reduction and are distinct from other sponge-associated spirochetes. Assessment of the functional potential of these spirochetes suggested that these symbionts were selected by their sponge hosts due to their ability to produce terpenoids and may provide antioxidizing effects to the sponge.

Data from this thesis has been included in a manuscript approved for publication and presented at two national conferences (Appendix C – Publications and Conferences).

Waterworth, S.C., Solomons, G.M., Kalinksi, J-C.J., Madonsela, L.S., Parker-Nance, S., Dorrington, R.A. (2024). The unique and enigmatic spirochete symbiont of latrunculid sponges. (Accepted at *mSphere*). Doi: 10.1128/msphere.00845-24

Chapter 1: Review of literature

1.1 Introduction

The history of symbiotic relationships is one that is almost as old as time itself and could be considered the very core of all evolutionary biology. The term ‘symbiosis’ is described as the cohabitation of different organisms within a single system (De Bary, 1978; Raval *et al*., 2022) with the purpose of undergoing a range of complex and intermingled relationships (Tirichine and Piganeau, 2023). Although numerous theories have been proposed on their origins there has been widespread acceptance that both metabolic symbiosis and successive endosymbiosis led to the evolution and diversification of modern eukaryotes (Hartman and Fedorov, 2002; Tirichine and Piganaeu, 2023). Consequently, symbiosis is considered one of the most important evolutionary processes that has brought forth thousands of new genomes and species and plays a key role in the development of new life throughout the Earth (Kierse and West, 2015; O’Malley, 2015).

The earliest example of symbiotic associations between two organisms are those between eukaryotic cells and mitochondria and chloroplasts dating back over two billion years ago. The role of both organelles in their eukaryotic hosts was to provide the cell with energy through adenosine triphosphate (ATP) production via respiration in the mitochondria and the conversion of solar energy into chemical energy and fixed carbon through photosynthesis in the chloroplasts (Dijiman, 2000). Over the years, studies have shown that the genetic make-up of these two cellular structures leans more towards those of prokaryotes, suggesting that these were once free-living prokaryotic cells that were acquired by predecessors of their eukaryotic cellular hosts who then coevolved to form a single unit consisting of the host cell and its endosymbiont, a process known as endosymbiosis (Apprill, 2002; Dijiman, 2000). Over time the genes associated with these once free-living organisms migrated into their eukaryotic cellular hosts, leading to the expansion of these eukaryote taxa through the diversification of both hosts and their symbiotic organelles while retaining traces of genes that show evidence of what were once free-living organisms (Apprill, 2002; Dijiman, 2000; Martin, 2000). One other example of a symbiotic association is one that occurs between a free-living fungus and a species of algae or cyanobacteria, or both, to form a lichen. The fungus provides over 90% of the biomass and structure to encapsulate the cyanobacteria and algae, which in turn provide the products of photosynthesis and carbon-fixation to the structure (Dijiman, 2000). Despite the genes for the development of this structure being present within the genome of the free-living fungus, the expression of these genes and subsequent formation of the structure is entirely dependent on the successful association of both or all three organisms, thus, providing evidence to support the “function-based” acquisition of microorganisms in symbiotic relationships (Dijiman, 2000; Dijiman, 2002).

Within the marine world are some of what can be considered the most genetically, morphologically, metabolically, and behaviourally diverse organisms on Earth (Dijiman, 2000; Margulis, 2010). In environments as harsh and complex as the ocean, these organisms have developed intricate and innovative ways of adapting and surviving through the acquisition of other organisms, often microorganisms including bacteria, algae, archaea, and several others, and the formation of symbiotic relationships (Apprill, 2020). These relationships often provide beneficial processes in both the hosts and acquired symbionts which contribute to the fitness and survival of both participants (Apprill, 2020; Cavanaugh, 2015; De Voogd *et al.*, 2015).

1.2 The role of symbiosis in evolution

The concept of a ‘holobiont’ and ‘hologenome’ was introduced by Rosenberg *et al.* (2017) to describe successful associations, such as these involved in the formation of lichen, that lead to successful functioning of the host and microbial symbiont as a single organism with the combined genetic material existing as an independent level of selection in evolutionary biology. Despite the controversy surrounding the concept of the hologenome, namely, that these do not meet the requirement of being organisms, evolutionary individuals or units of selection, the concept sheds a positive light on the role symbiosis plays in driving evolutionary advancements in biology and ecology. (Gilbert *et al.*, 2010; Kiers and West, 2015; Sudakaran *et al.*, 2017). The time frames during which new functional traits or behaviours develop within the organisms in these so-called “units” tends to be far shorter than they would in standard evolutionary events, further supporting the role symbiosis plays in expediting large-scale evolutionary changes (Margulis, 1991).

The term ‘symbiogenesis’ has often been associated with major evolutionary events like those that gave rise to the formation of the mitochondria and chloroplasts as new cellular structures in ancient eukaryotic cells (Cavalier-Smith, 2013). However, when dealing with small-scale evolutionary changes the terms ‘co-evolution’ and ‘phylosymbiosis’ accurately describe the changes that take place directly within these host-symbiont systems (Ehrlich and Raven, 1964). “Co-evolution” or “co-speciation”, as the names suggest, occurs when two or more partners in a symbiotic relationship evolve together in response to one another (Ehrlich and Raven, 1994). A well-studied model of this association exists between the Hawaiian bobtail squid *Euprymna scolopes,* and the bacterium *Vibrio fischeri*, a free-living bacterium acquired for its bioluminescent capabilities utilized by the squid when attracting prey and evading predators (McFall-Ngai *et al*., 2012). Over time, both participants co-evolved to possess a range of genes for signalling and recognition that allow the continuation of this mutualistic relationship. Molecular studies on the dynamics between hosts and symbiont communities introduced the term ‘phylosymbiosis’ which provided evidence of the correlation between host phylogeny and the structure of these associated microbial communities (O’Brien *et al.,* 2019; Wallin, 1927). The most well-known marine forms of this association can be found in coral reefs, reef fish and marine sponges (Pollock *et al*., 2018; Chiarello *et al*., 2018; Reveillaud *et al*., 2014).

1.3 Symbiotic relationships within marine sponges

Due to their existence as sessile creatures marine sponges (Phylum Porifera) have been solely dependent on the uptake of nutrients from the environment through filter-feeding via phagocytosis (Brain *et al*., 2012; Maloof *et al*., 2010). Once ingested, several microorganisms such as algae, archaea, bacteria are able to avoid digestion by migrating through the mesohyl of the sponge, allowing them to survive and proliferate successfully (Brain *et al*., 2012; Maloof *et al*., 2010). These microorganisms can account for a large portion of the sponge biomass, as evident by bacteria that can account for over 60% of the sponge mass and represent a significant difference in diversity and metabolic activity when compared to those present in the surrounding seawater (Hentschel *et al*., 2002; Hentschel *et al*., 2006). The distribution and properties of these microbial communities differs according to different sponges, however, are often dictated by the characterization of sponges as either ‘low microbial abundance’ (LMA) sponges or ‘high microbial abundance’ (HMA) sponges (Hentschel *et al*., 2006). The microbial communities within LMA sponges exhibit approximately the same microbial densities as the surrounding seawater while HMA sponge microbial communities are two to four times higher than the surrounding seawater (Hentschel *et al*., 2006; Azam, 2007; Gloeckner *et al*., 2014; Moitinho-Silva *et al*., 2017). The independent evolution of these two patterns of microbial distribution has recently been implicated in the fundamental differences observed in the taxonomy composition and function of sponge-associated microbial communities (Gloeckner *et al*., 2014; Moitinho-Silva *et al*., 2017). HMA sponges are enriched with populations of Acidobacteria, Chloroflexi and Poribacteria which carry out autotrophic and heterotrophic metabolisms within the sponges. Conversely, LMA sponges carry out their own autotrophic and heterotrophic metabolisms and are enriched with Cyanobacteria and Proteobacteria (Gloeckner *et al*., 2014; Thomas *et al*., 2016).

Symbiotic relationships tend to develop between the sponge hosts and the associated bacteria within the sponge tissue. The sponge provides the symbiont with the necessary nutrients to survive and proliferate, while the bacteria can perform various roles within the sponge such as the fixation and cycling of sulfur, nitrogen, and phosphorus (Jensen *et al*., 2017; Karimi *et al*., 2018; Zhang *et al*., 2015; Zhang *et al*., 2019), the uptake of carbon, cycling and detoxification (De Voogd *et al*., 2015; Hentschel *et al.,* 2006), providing stability to the host skeletal system (Garate *et al*., 2017; Ruocco et al., 2021) and in some cases, facilitate the production of bioactive secondary metabolites as chemical defensive strategies (Mahon *et al.,* 2003; Paul *et al*., 2006; Mehbub *et al*., 2014). The maintenance of these symbiotic relationships is highly dependent on strict coordination between the sponge host and the bacterial symbionts. The two processes that ensure the maintenance of symbiotic relationships across host-symbiont generations are horizontal and vertical transmission (Bright and Bulgheresi, 2010; McFall-Ngai *et al*., 2012). Horizontal gene transfer involves the non-continuous transmission of genetic elements in non-mating organisms as observed in the symbiotic association between the *V. fischeri* and its Bobtail squid host (McFall-Ngai *et al*., 2012). In contrast, symbiotic associations facilitated and maintained through vertical transmission are transmitted from parent to offspring hosts, with the symbiotic association evident in all life stages across several generations of the hosts (Bright and Bulgheresi, 2010; McFall-Ngai *et al*., 2012). A study conducted by Waterworth *et al.* (2017) on the microbial communities associated with the sponge *Tethya rubra* showed evidence of vertical transmission due to the conservation of the same Alpha- and Betaproteobacteria symbionts in parent and embryo *T. rubra* sponges suggesting the crucial role bacterial symbionts such as these play in the survival and development of marine sponges (Waterworth *et al*., 2017). Although there has been widespread evidence to support use of vertical transmission in the long-term maintenance of symbiosis across several generations of sponge microbiomes, there is no reported evidence that the mechanism is universally observed across sponge microbiomes (Tyler *et al.,* 2022).

1.4 Ubiquitous sponge-symbionts

The mechanism utilized by sponge hosts when distinguishing between bacteria that are either harmful or beneficial are poorly understood. However, intricate morphological, chemical, and molecular mediated interactions between these sponge hosts and structural components of the bacterial symbionts such as the lipopolysaccharide, peptidoglycan and flagella could provide the necessary assistance in making these distinctions (Pita *et al.,* 2018; Hentschel *et al*., 2006; Frost, 1976; Maldonado *et al*., 2010; Schmittmann *et al*., 2020; Marulanda-Gomez *et al.,* 2023). Host sponge-mediated immune responses may also drive the selectivity and specificity of bacterial symbionts in HMA versus LMA sponges (Marulanda-Gomez *et al.,* 2023). Although marine sponges exhibit selectiveness and specificity of their bacterial symbionts, there are some groups of symbionts found to be distributed across sponge hosts that are phylogenetically and geographically distant (Thomas *et al*., 2016; Astudillo-García *et al.,* 2018). To date, three groups of ubiquitous sponge-associated symbionts have been identified across phylogenetically distant host sponges, namely, the *Poribacteria*, the “Sponge-Associated Unclassified Lineage” (SAUL) (Astudillo-García *et al.,* 2018) as well as a newly identified Gammaproteobacteria group (Taylor *et al.,* 2021).

Previously thought to only exist as a sponge-symbiont, the phylum *Poribacteria* has been reclassified to contain two distinct groups of bacteria, namely, the *Entoporibacteria*, a sponge-associated symbiont, and the free-living *Pelagiporibacteria* (Fieseler *et al*., 2004; Tully *et al*., 2018; Podell *et al*., 2019). *Entoporibacteria* symbionts have been characterised in several unrelated marine sponges collected in three geographical regions of the world, therefore, showing the lack of correlation between the phylogenies between the symbionts or their hosts or their geographic locations (Podell *et al*., 2019; Lafi *et al*., 2009). The presence of several genes within the *Entoporibacteria* encoding enzymes responsible for carbon degradation and sulfate and uronic acid metabolism suggest that the symbiont may perform the same role of proteoglycan degradation in the different hosts (Kamke *et al*., 2013; Kamke *et al*., 2014; Hill and Sacristán-Soriano, 2017), with the exception of *Aplysina aerophile* sponges in which a higher expression of the genes responsible for 1,2-propanediol degradation and vitamin B12 importation was revealed, suggesting the symbiont’s role in the host as 1,2-propanediol degradation and energy production (Kamke *et al*., 2014).

The SAUL bacterial symbionts, classified as belonging to the larger taxon of candidate phylum PAUC34f within the *Candidatus* Nemesobacterales order have been identified in several marine sponges, however, at relatively lower abundances (Chen *et al*., 2020; Gavrilidou *et al*., 2023). The isolation of these symbionts in other eukaryotic hosts such as corals and tunicates suggest a high likelihood that they were acquired at different points in evolution from the surrounding environment (Chen *et al*., 2020). Despite being a relatively poorly studied phylum, current knowledge suggests that these symbionts may play a role in the degradation of host and algal carbohydrates and phosphate storage as shown in the sponge *Discodermia calyx* which utilizes its phosphate stores during chemical defence strategies (Astudillo-García *et al*., 2018; Chen *et al.,* 2020).

A newly described group of broad host range sponge-associated symbionts are the Betaproteobacteria (recently reclassified as members of the Class Gammaproteobacteria (Park *et al.,* 2018)) belonging to the proposed new order *Tethybacterales* (Taylor *et al*., 2021). The order comprises the families Persebacteraceae and Tethybacteraceae (Taylor *et al*., 2021) as well as the newly identified order Polydorabacteraceae (Waterworth *et al*., 2021). An assessment of the metagenome-assembled-genomes (MAGs) from different species of families *Tethybacteraceae* and *Persebacteraceae* provided evidence suggesting that the bacteria had been acquired by their host sponges relatively early in their evolutionary history and coevolved with their hosts (Taylor *et al*., 2012).

Symbionts from the *Tethybacteraceae* family have previously been reported in several phylogenetically distant sponges from various geographical locations representative of both low and high microbial abundance sponges (Taylor *et al*., 2021; Waterworth *et al.,* 2021). In addition, these bacterial symbionts have also been identified in other marine invertebrates, the surrounding seawater, and marine sediments (Taylor *et al*., 2021), further, supporting the hypothesized acquisition of these bacteria from the environment. Several sponge microbiome studies have reported Tethybacterales symbionts as the most conserved and numerically dominant bacteria, with a select few existing as cellular endosymbionts (Croué *et al*., 2013; Gautier *et al.,* 2016; Matcher *et al*., 2017; Thiel *et al*., 2007; Waterworth *et al*., 2017; Webster *et al.,* 2001). Within the microbiome of the marine sponge *Amphimedon queenslandica*, are two codominant symbionts, *Amphirhobacter heroislandensis* AqS2 and its sulfur-oxidizing relative AqS1. Interestingly, although both symbionts exhibit functional similarities in energy production, sulfur-uptake and the production of essential amino acids, differences in what metabolites the symbionts could transport have been suggested (Gautier *et al.,* 2016). Furthermore, the association of *A. heroislandensis* AqS2 with all stages of the sponge life cycle as well as the presence of a reduced genome indicate that the bacterium exists as an endosymbiont within the sponge host (Gautier *et al*., 2016; Fieth *et al*., 2016).

1.5 The role of bacterial symbionts in the production of bioactive secondary metabolites in marine sponges

As previously discussed, the sessile state and porous nature of marine sponges results in the ingestion of large quantities of sea water and a diverse range of microorganisms. This sessile state not only leaves sponges susceptible to predators, but the uptake of these microorganisms also leads to the introduction of pathogenic bacteria, viruses, potential parasites, and fungi into the sponge tissues, leaving them susceptible to a range of harmful infections and diseases (Mehbub *et al*., 2014). Marine sponges have, therefore, acquired mechanisms of adapting to these threats through the production of highly potent and cytotoxic bioactive secondary metabolites that demonstrate inhibitory effects against the growth and survival of bacteria, viruses, parasites, and several other microbial infections (Mahon *et al.,* 2003; Paul *et al*., 2006; Mehbub *et al*., 2014). In many cases these compounds have shown significant pharmacological potential against viral, bacterial, parasitic, and fungal infections and have served as potential leads for the development of drugs against these types of infections as well as anti-cancer agents on account of their cytotoxicity (Paul *et al.,* 2006; Blunt *et al*., 2005; Li *et al*., 2023). Although these compounds are produced in very small amounts, they are often highly potent and stable compared to the high salinity levels associated with the ocean (Abad *et al.,* 2012), leading to the isolation of pure compounds as marine natural products that can be directly assessed for biological activity (Abad *et al*., 2012; Kalinski *et al*., 2019; Kalinski *et al.,* 2021).

Since their first discovery in 1950 thousands of new compounds have been isolated from marine sponges, (El-Demerdash *et al.,* 2018; Sagar *et al.,* 2018; Chu *et al*., 2022) many of which include members of numerous sponge orders Agelasida, Astrophorida, Axinellida, Chondrosida, Choristida, Clathrinida, Dendroceratida, Dictyoceratida, Hadromerida, Halichondrida, Haplosclerida, Homosclerophorida, Leucosolenida, Lithistida, Lyssacinosida, Ocilosclerida, Poecilosclerida, Spirophorida and Verongida (Mehbub *et al*., 2014). These compounds have been categorized according to the chemical and structural differences associated with each (Abad *et al.,* 2012; Hu *et al*., 2011; Bian *et al*., 2020) into eighteen classes – acids, alkaloids, esters, fatty acids, glycosides, ketones, lipids, macrolides, alcohols, peptides, peroxides, polyketides, quinones, steroids, sterols, terpenes, terpenoids and unclassified natural products as the distinct classes by which sponge-derived natural products can be categorized (Blunt *et al*., 2012; Mehbub *et al*., 2014; Li *et al.,* 2023). Furthermore, certain pharmacological capabilities were shown to be associated with a particular class – anti-cancer and tumour inhibition activity were associated with quinones, steroids, fatty acids, diketopiperazines, alkaloids, terpenes, terpenoids, polyketides, cyclopeptides and glycoglycerolipids; quinoline derivatives were associated with anti-HIV activity; fatty acid esters and fatty acids were associated with anti-inflammatory activity; and alkaloids and quinoline derivates were associated with anti-malaria activity (Thomas *et al*., 2010; Mehbub *et al*., 2014). Anti-microbial activity was associated with polyketides, glycopeptides, α-pyrone derivatives, peptides, proteins, antimycin, lipopeptides, polybrominated biphenyl ether, cyclic depsipetide, terpenes, pentaketides, furan carboxylic acids, alkaloids, diketopiperazines, anthraquinones, chromones, steroids, lactones, quinolone derivatives, trisindole derivatives, macrolactams, ethers, phenol derivatives, while neuroprotectivity was associated with dihydropiridines (Thomas *et al*., 2010; Mehbub *et al*., 2014). As the most prolific contributor of marine-derived compounds, alkaloids represent the largest and most well-studied class of marine natural products (Elissawy *et al*., 2021). Marine-derived alkaloids are a group of chemically diverse compounds containing nitrogenated cyclic cores (Bian *et al*., 2020; Munekata, *et al*., 2021; Elissawy *et al*., 2021). Alkaloids are widely distributed across a range of marine organisms, including algae, microorganisms such as cyanobacteria and actinobacteria, sponges, cnidarians, and other invertebrates (Elissawy *et al.,* 2021). These compounds are highly potent with major pharmacological applications as anti-cancer, anti-inflammatory, anti-parasitic and anti-microbial drug leads (Bian *et al.,* 2020; Elissawy *et al*., 2021; Souza *et al*., 2020; Tempone *et al*., 2021). Currently the five major classes of marine alkaloids that have been described are the pyrroloiminoquinone alkaloids, guanidine alkaloids, bromotyrosine alkaloids, indole alkaloids and pyridine alkaloids, the largest of which are the pyrroloiminoquinones (Elissawy *et al*., 2021).

1.5.1 Pyrroloiminoquinones as a source of bioactive secondary metabolites in marine sponges

Pyrroloiminoquinones have been isolated predominantly from marine sponges (Urban *et al.,* 2000; Antunes *et al*., 2005; Hu *et al*., 2011; Li *et al*., 2021; reviewed in Kalinski *et al*., 2022) and exhibit extremely high levels of cytotoxicity and potency, making them the focus of research into the development of broad-spectrum anticancer, anti-bacterial, anti-inflammatory and anti-malarial drug leads (reviewed in Kalinski *et al.,* 2022). These compounds have been isolated mostly from marine sponges of the *Latrunculiidae* family (order Poecilosclerida), however, ascidians (Copp and Ireland, 1991; Grckovic *et al*., 2021), cultured myxomycetes (Ishibashi *et al*., 2001; Nakatani *et al.,* 2005), hydroids (Zlotkowski *et al.,* 2017; Yan *et al*., 2018), terrestrial fungi (Peters and Spitella, 2007; Peters *et al.,* 2008; Pulte *et al.,* 2016; Lohmann *et al.,* 2018) and marine actinobacteria (Nagata *et al.,* 1997; Hughes *et al.,* 2007) have also been shown to produce similar alkaloids or derivates of pyrroloiminoquinones. The structure of these compounds begins with a common condensed tricyclic pyrroloquinoline core (Figure 1.1), which is proposed to be responsible for the cytotoxicity and anti-proliferative capabilities associated with these compounds (Radisky *et al*., 1992; Antunes *et al.,* 2005; Kalinski *et al.,* 2022). These compounds can be categorized into three main classes based on the complexity of their structures – makaluvamines as the simplest class containing the shared pyrroloquinoline core, followed by the more complex bispyrroloquinones and discorhabdins (Antunes *et al*., 2004; Antunes *et al.,* 2005; Kalinski *et al*., 2022).

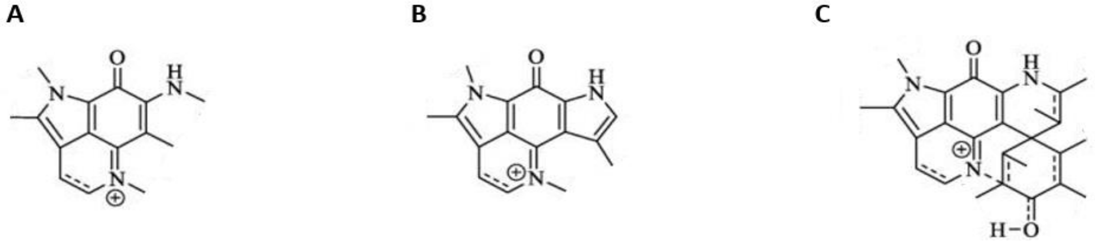

Figure 1.1: Chemical structure of pyrroloiminoquinones showing A) makaluvamine, B) bispyrroloquinone and C) discorhabdin classes. All three classes possess the common tricyclic pyrroloquinoline cores with variant substituents. (Adapted from Kalinski *et al*., 2022).

1.5.2 The biosynthetic origins of pyrroloiminoquinone production

The biosynthetic origins of pyrroloiminoquinones have been a topic of research and debate since their pharmacological potential was first discovered. The isolation of pyrroloiminoquinones and pyrroloiminoquinone-related compounds from organisms unrelated to marine sponges, ascidians, terrestrial bacteria and myxomycetes (Copp, Ireland and Barrows, 2002; Davis *et al.,* 2012; Hughes *et al.,* 2009; Ishibashi *et al.,* 2001; Miyanaga *et al*., 2011; Nagata *et al.,* 1997) suggests the presence of shared biosynthetic gene clusters in both groups of terrestrial and marine pyrroloiminoquinone producers or the presence of a shared microbial symbiont involved in the biosynthesis of pyrroloiminoquinones.

The biosynthetic pathway for pyrroloiminoquinone production has not yet been conclusively proven, however, chemotaxonomic relationships, structural similarities shared among pyrroloiminoquinones, and direct experimental evidence have proposed a rational sequence of biosynthetic reactions (Urban *et al.,* 2000; Antunes *et al.,* 2005; Lill *et al.,* 1995) (Figure 1.2). The pathway is proposed to begin with the decarboxylation of tryptophan and several oxidation and condensation steps to produce a makaluvamine precursor of "proto"-makaluvamine. From this precursor, oxidation, or amination results in the unbranched makaluvamines (Lill *et al.,* 1995), known as pyrrolo-*ortho*-quinones that are speculated to be either degradation products of pyrroloiminoquinones or the byproducts biosynthetically produced during makaluvamine biosynthesis. The uncommon oxazole pyrroloiminoquinones makaluvamine W and citharoxazole have been proposed as derivates of damirones synthesized through condensation with glycine, followed by decarboxylation and oxidation (Taufa *et al.,* 2019; Genta-Jouve *et al.,* 2011). Similar condensation reactions between makaluvamine and alternative reaction partners may give rise to zyzzamines (Kalinski *et al.,* 2022). The conversion from unbranched makaluvamines to branched makaluvamines is proposed to be initiated by the addition of tyramine, likely proceeding the incorporation of its precursor molecule phenylalanine (Lill *et al*., 1995). The more complex pyrroloiminoquinones discorhabdins and bispyrroloiminoquinones may be derived from makaluvamines containing the N-phenylethyl side chain, a process not confined to specific steps within the pathway (Lill *et al.,* 1995). Discorhabdins containing the C5-C8 sulfur bridges, including A- and D-series discorhabdins, are proposed to have been biosynthesised from either makaluvamine F or C-series discorhabdins (Urban *et al*., 2000; Antunes *et al*., 2005; Lill *et al*., 1995). In 2013, Zou and Hamman hypothesized that the pyrroloiminoquinone atkamine was a derivative of a N-phenylethyl makaluvamine such as makaluvamine F while aleutianimine was proposed to be a downstream product of discorhabdin A (Zou *et al*., 2019).

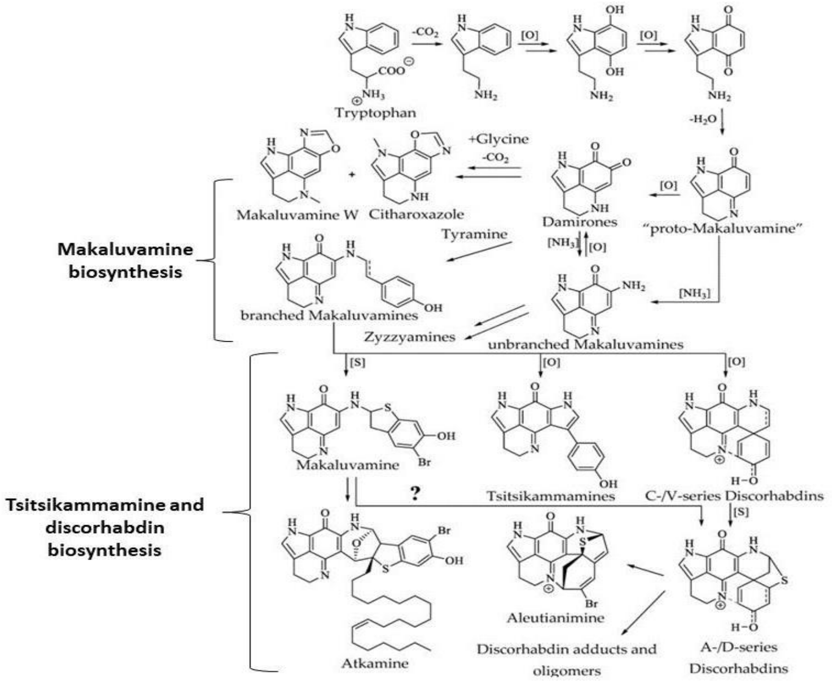

Figure 1.2: Proposed biosynthetic pathway for pyrroloiminoquinone biosynthesis showing the distinction between makaluvamine production and the successive production of the more complex bispyrroloiminoquinones (such as tsitsikammamine) and discorhabdins. (Adapted from Kalinski *et al*., 2019 and Kalinski *et al*., 2022).

The biosynthetic origins of these compounds are yet to be fully understood. A study conducted by Lill *et al*. (1995) in which portions of *Latrunculia sp*. sponge specimens were immersed in a solution of broad-spectrum antibiotics, suggested the biosynthesis of discorhabdins was not orchestrated by the associated bacteria within the sponge, but rather by the sponge itself. However, the broad distribution of pyrroloiminoquinones across unrelated marine and terrestrial organisms including sponges, ascidians and terrestrial myxomycetes suggests that the makaluvamine core shared amongst all three organisms may be produced by a microbe, with the sponge hosts coordinating the conversions to discorhabdins and tsitsikammamines (Kalinski *et al*., 2019; Walmsley *et al*., 2012a). Previous studies into the pyrroloiminoquinone-related compounds, lymphostins and ammosamides, identified the presence of a Ribosomally synthesized and post-translationally modified peptides (RiPP) class gene cluster associated with the genomes from which both compounds were isolated, implicating the gene cluster in the biosynthetic origins of these compounds (Jordan and Moore, 2016). Furthermore, a later study showed that ammosamides are biosynthesized by the attachment of a tryptophan to the C-terminus of ribosomally synthesized peptides, followed by hydroxylations and oxidation to the quinone cores, while the primary amines are introduced from a glycerol residue (Daniels *et al.*, 2022). Although as yet unproven, the same mechanism may be employed within the pyrroloiminoquinone biosynthetic pathway.

1.6 The *Latrunculiidae* family of marine sponges as producers of pyrroloiminoquinones

The *Latrunculiidae* (Class Demospongiae; Order Poecilosclerida) are a diverse family of marine sponges found in temperate and cold waters of both the Southern and Northern Hemispheres off the coasts of South Africa, Australia, New Zealand, Tasmania, and Antarctica, as well as British Columbia, the Aleutian Islands, the Gulf of Alaska, the Russian Sea of Okhotsk, and the Pacifica North-West (Kelly *et al.*, 2016). The family was first described in 1922 and has since undergone several revisions (Topsent, 1922). Morphologically, these are large, encrusting, hemispherical, spherical or pedunculate sponges with smooth surfaces between raised fistular oscules, with dense and uniformed body masses (Kelly and Samaai, 2002). These sponges are leathery in texture, with colours that range from brownish-black, dark brown, forest green, to pale beige or white (Kelly and Samaai, 2002). Additionally, the most distinguishing feature of the family are the presence of discorhabd microscleres with apical and basal spines, generally arranged in compact perpendicular spicules (Kelly and Samaai, 2002). Currently the family has seven recognized genera, namely, *Bomba, Sceptrella, Strongylodesma, Cyclacanthia, Latrunculia, Latrunclava,* and *Tsitsikamma* (WoRMS Editorial Board, 2017; Li *et al*., 2021). Members of the *Strongylodesma, Cyclacanthia, Latrunculia* and *Tsitsikamma* genera are endemic to South Africa.

Latrunculid sponges are the most prolific producers of all three classes of pyrroloiminoquinones (Kelly *et al.,* 2016; Davies-Coleman *et al.,* 2023). The genus *Tsitsikamma* currently comprises seven species, including, *T. favus, T. pedunculata* and recently discovered novel species, *T. nguni, T. michaeli* (Parker-Nance *et al.,* 2019), *T. madiba, T. amatholaensis* and *T. beukesii* (Samaai *et al*., 2020). Species *T. pedunculata* was first collected in 1999 within the Algoa Bay region of the Eastern Cape Province (South Africa) and has reportedly yielded the known pyrroloiminoquinones 14-bromodiscorhabdin C, 14-bromo-3-dihydrodiscorhabdin C, and 3-dihydrodiscorhabdin C. as well as four novel, minor metabolites 3-dihydro-7,8-dehydrodiscorhabdin C, 14-bromo-3-dihydro-7,8-dehydrodiscorhabdin C, discorhabdin V, and 14-bromo-1-hydroxydiscorhabdin V (Antunes *et* *al*., 2004). *T. favus* sponges were first collected in 1994 within the Tsitsikamma marine region and were later morphologically described by Kelly and Samaai in 2002. Hooper *et al.* (1996) isolated three major pyrroloiminoquinones, from *T. favus* sponges, namely, tsitsikammamines A and B, as well as 14-bromo-3-dihydrodiscorhabdin C. Advancements in chromatographic and spectroscopy techniques led to the characterisation of minor secondary metabolites from *T. favus* specimens, leading to the isolation of new pyrroloiminoquinones tsitsikammamine N-oxime, tsitsikammamine B N-oxime, 7,8-dehydro-3-dihydrodiscorhabdin C, 14-bromo-1-hydroxy-discorhabdin S, and 2,4-debromo-3-dihydrodiscorhabdin C (Antunes *et al.,* 2005). The two major compounds, tsitsikammamines A and B, were isolated exclusively from *T. favus* sponges, highlighting their uniqueness to the species (Antunes *et al*., 2004). Species *T. michaeli* and *T. nguni* were only recently reported as pyrroloiminoquinone producers of non- or mono-brominated, hydroxylated discorhabdins (Kalinski *et al.,* 2021; Parker-Nance *et al.,* 2019). No compounds have been isolated from *T. scurra* species. Discorhabdins have also been isolated from members of the *Cyclacanthia, Strongylodesma* and *Latrunculia* genera (Antunes *et al*., 2004; Li *et al*., 2021), with similar chemical profiles identified in *Cyclacanthia bellae* and *Latrunculia apicalis* sponges following the isolation of C-series discorhabdins from both species (Yang *et al.,* 1995; Antunes *et al.,* 2004; Kalinski *et al.,* 2021).

Kalinski and colleagues have reported the presence of two distinct chemical profiles (chemotypes) in *T. favus* and *T. michaeli* sponges collected from Evans Peak and Riy Banks in Algoa Bay, South Africa (Kalinski *et al.,* 2019; Kalinski *et al.,* 2021). Chemotype I *T. favus* sponges were abundant in discorhabdins and tsitsikammamines, while chemotype II *T. favus* sponges had abundant makaluvamines (Kalinski *et al*., 2019). Chemotype I *T. michaeli* sponges contained abundant tribrominated discorhabdins while chemotype II *T. michaeli* sponges produced mostly dibrominated and hydroxydiscorhabdins (Kalinski *et al*., 2021). Both chemotypes were identified in *T. favus* sponges collected from the same location in Evans Peak, while the chemotype II profiles were only observed in *T. michaeli* sponges collected from Evans Peak. The presence of multiple chemical profiles in the same species is not a rare occurrence in marine sponges and intra-species variations in chemical profiles of *T. favus* sponges collected from the same site has not been previously reported (Kalinski *et al.,* 2021). Interestingly an abundance of phenylalanine was observed in chemotype II *T. favus* sponges, likely the result of an interference in the biosynthetic pathway preventing the incorporation of phenylalanine into makaluvamine, resulting in the accumulation of this amino acid in chemotype II *T. favus* sponges (Kalinski *et al.,* 2021).

1.7 The microbial communities associated with latrunculid sponges

Marine sponges of the *Latrunculiidae* family share highly conserved microbiomes that are dominated by two Gammaproteobacteria and Spirochete symbionts (Walmsley *et al.,* 2012a; Matcher *et al*., 2017). These conserved symbionts were initially identified in *T. favus* specimens collected over the course of six years from Algoa Bay, Port Elizabeth, (Walmsley *et al.,* 2012a). Later, through 16S rRNA clones sequence analysis with several closely related symbionts identified in other *Tsitsikamma* species, *Cyclacanthia bellae* and *Latrunculia* species as well as several non-latrunculid species (Matcher *et al*., 2017). Phylogenetic analysis of the relationship of the operational taxonomic units ($OTU_{0.03}$) representative of these Gammaproteobacteria symbionts, relative to other sponge-associated Gammaproteobacteria, identified two monophylogenetic clusters, one containing free-living, and sponge associated Gammaproteobacteria bacteria and the other containing exclusively sponge-associated Gammaproteobacteria (Figure 1.3A, Group II). The latter group, Group II, included the dominant Gammaproteobacteria clone Sp02-1 previously identified in an early *T. favus* specimen (TIC2009-002) (Walmsley *et al.,* 2012a). OTUs within the Group II cluster showed sequence similarity of greater than 99% with Sp02-1 indicating that they represent the same species. Interestingly, the phylogenetic relationship displayed by these Gammaproteobacteria symbionts is strikingly similar to that displayed by their hosts (Figure 1.3B) providing evidence to suggest that these symbionts coevolved with their hosts.

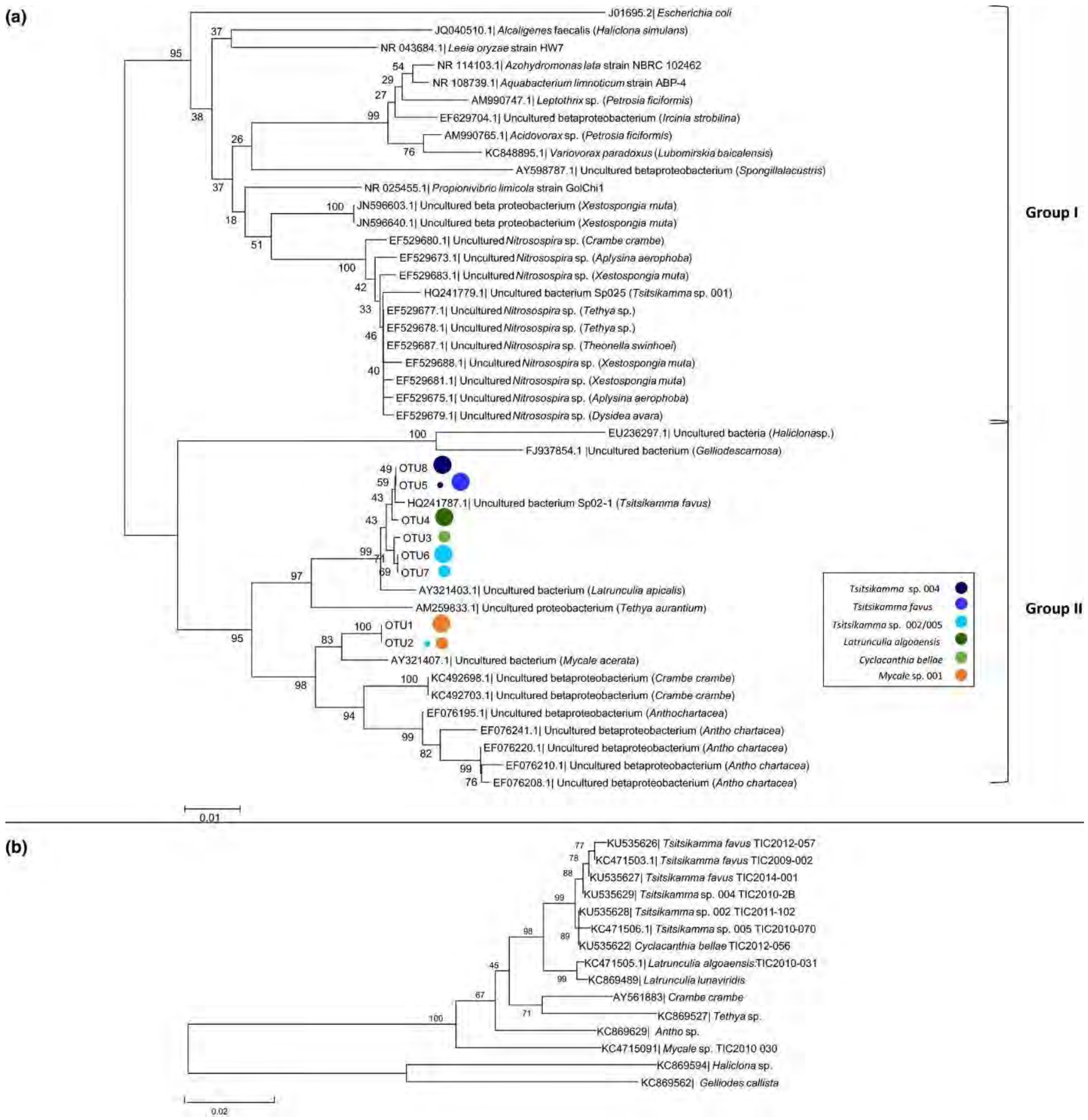

Figure 1.3: Phylogenetic analysis of dominant Gammaproteobacteria species in latrunculid sponges showed that A) the OTUs identified were closely related to the dominant Gammaproteobacteria clone Sp02-1 previously identified in *T. favus* TIC2009-002. B) The divergence patterns displayed by these Gammaproteobacteria closely resembles that displayed by their sponge hosts suggesting that these symbionts coevolved with their hosts. (Taken from Matcher *et al*., 2017)

Within the same study Matcher *et al*., (2017) identified that symbionts of the Spirochete phylum were exclusively found in latrunculid sponges where the Spirochaete reads were also dominated by a single OTU. Interestingly, although only one unique OTU was identified in *C.*

bellae, T. favus sponges revealed four unique *Tsitsikamma*-specific Spirochete OTUs with an apparent species-specific distribution and relative abundance of two species per sponge host. Phylogenetic analysis showed an extremely close relationship between *Tsitsikamma*-spirochetes with all five species clustering with *T. favus* spirochete clone Sp02-3 (Walmsley *et al.,* 2012a) and the free-living Spirochete *Salinispira pacifica* (Figure 1.4). Subsequently, the closest sponge-associated Spirochete relative identified within the genus *Clathrina* showed less than 90% species similarity with the latrunculid Spirochetes, suggesting that these latrunculid Spirochetes could represent members of a new family of specialized sponge spirochetes.

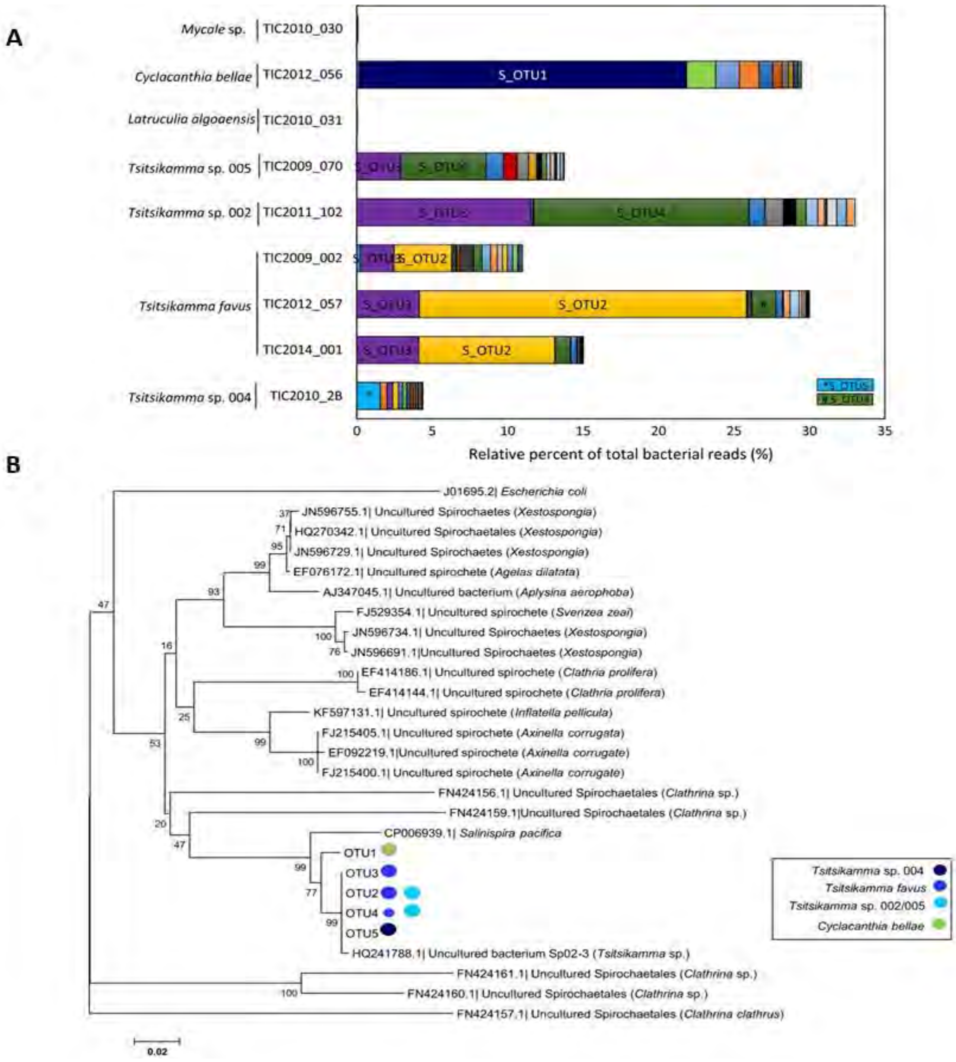

Figure 1.4: A unique lineage of dominant Spirochete symbionts is unique to latrunculid sponges. Phylogenetic analysis showed the latrunculid spirochetes are distantly related to other sponge-associated spirochetes, suggesting that these represent a new family of specialized sponge spirochetes. (Taken from Matcher *et al.,* 2017)

To determine whether these Gammaproteobacteria and Spirochete symbionts displayed a direct correlation between numerical dominance and metabolic activity Matcher *et al*., (2017) utilized a protocol that isolated RNA and DNA from sample specimens simultaneously allowing the identification of a potential correlation between 16S rRNA abundance and 16S rRNA expression. Consequently, this revealed that the Gammaproteobacteria symbionts identified in the study were both numerically dominant and metabolically active in latrunculid sponges while the Spirochete symbionts were numerically dominant but not metabolically active. Further inspection did, however, indicate the presence of a significantly less-dominant yet metabolically active Spirochete symbiont in several *T. favus* sponges. This suggests that the Gammaproteobacteria symbionts were acquired to assist the sponge with metabolic activities such as carbon, sulfur or nitrogen metabolism (Waterworth *et al*., 2021) whereas the Spirochetes may have been acquired by *T. favus* (and possibly *C. bellae*) sponges to perform specialized, non-metabolically active functions, potentially assisting in the production of the tsitsikammamine and discorhabdin pyrroloiminoquinones characteristic of the *Tsitsikamma* and *Cyclacanthia* genera (Walmsley *et al*., 2012a).

1.8 Elucidation of the dominant Gammaproteobacterial symbiont genome associated with the *T. favus* microbiome

A study conducted by Waterworth *et al*. (2021) set out to determine the functional potential of the dominant Gammaproteobacterial symbiont associated with *T. favus* sponges. Herein, the metagenome of a *T. favus* sponge (TIC2018-003B) was sequenced and used to identify and characterize the genome of the dominant Gammaproteobacteria symbiont (Bin 003B_4). Analysis of the genome of this Gammaproteobacterial symbiont identified the presence of a 16S rRNA sequence that was closely related to the previously identified, highly conserved Sp02-1 symbiont (Walmsley *et al*., 2012a; Matcher *et al.,* 2017). Bin 003B_4, thus, served as a representative within the newly proposed *Tethybacterales* order and was used to determine the role of the symbiont across a broad range of sponge hosts (Waterworth *et al*., 2021).

Characterization of the genome according to the minimum information about a genome (MIMAG) standards (Bowers *et al*., 2017) revealed that the genome was of medium quality, approximately 2.95 million base pairs (Mbp) in size and appeared to have undergone genome reduction due to the abundance of pseudogenes (approximately 25% of all genes within the genome) and notably low coding density (65.27%), significantly lower than the characteristic 80% for bacterial genomes (McCutcheon and Morran, 2011). This was consistent with other members within the proposed *Tethybacterales* order (Taylor *et al.,* 2021). Taking into account the low coding density, the genome encoded several genes for glycolysis, Phosphoribosyl diphosphate (PRPP) biosynthesis, as well as the majority of genes required for the citrate cycle and oxidative phosphorylation. Additionally, the genes for the biosynthesis of valine, leucine, isoleucine, tryptophan, phenylalanine, tyrosine, and ornithine amino acids, as well as the genes required for the transport of L-amino acids, proline, and branched amino acids. This suggested that the symbiont may be involved in the exchange of amino acids with the host, as previously shown in insect and sponge-symbiont symbiosis (Moitinho-Silva *et al.,* 2017; Feng *et al*., 2019). A total of 13 unique genes were identified, one of which was predicted to encode an ABC transporter permease subunit that was likely involved in glycine betaine and proline betaine uptake (Waterworth *et al*., 2021). A second gene encoding a 5-oxoprolinase subunit PxpA suggested that the genome was also involved in acquisition of proline and its subsequent conversion to glutamate (Niehaus *et al*., 2017). Other unique genes include restriction endonucleases and site-specific DNA methyltransferase, suggested to aid in the removal of foreign DNA, and seven genes associated with phages, including the anti-restriction protein ArdA, proposed to prevent DNA cleavage and provide anti-restriction energy (Chen *et al*., 2014). Lastly, two unique genes were predicted to encode an ankyrin repeat domain-containing protein and a von Willebrand factor type A VWA protein, both involved in cell-adhesion and protein-protein interactions (Boyd *et al*., 2014; Al-Khodor *et al*., 2011), suggesting its role in facilitating the state of symbiosis within the host.

A comparative analysis of twenty-seven *Tethybacterales* genomes, including that of Sp02-1, was conducted to determine whether the functional potential of each genome differed according to the host in which they were acquired (Waterworth *et al*., 2021). Herein, the phylogeny of these *Tethybacterales* symbionts was determined and established as members of the deep-branching *Persebacteraceae* family. Furthermore, members of a newly proposed family, *Polydoraceae*, were characterised, sharing an average amino acid identity (AAI) of 80% within the family and less than 89% sequence similarity between all three families – suggesting that these may represent novel classes within the *Tethybacterales* order (Waterworth *et al.,* 2021). Consequently, the newly identified species, including Sp02-1 (Bin 003B_4), were proposedly renamed as members of newly identified genera: "*Candidatus* Ukwabelana africanus", "*Candidatus* Regalo mexicanus", "*Candidatus* Dora taiwaniensis", "*Candidatus* singaporensis", and "*Candidatus* Hadiah malaccus". Furthermore, 18 shared genes were identified between all twenty-seven *Tethybacterales* genomes encoding several genes involved in energy production, tryptophan production, and stress response.

Additionally, a comparison of the 17 *Tethybacterales* genomes against genomes of the symbiotic *Entoporibacteria* and free-living *Pelagioporibacteria* clearly illustrated that sponge-symbionts had not converged to perform similar roles within their hosts as previously hypothesized (Waterworth *et al*., 2021) but that each symbiont family had evolved to perform distinct functions within their sponge hosts. Several genes were present across two out of the three *Tethybacterales* families but absent in the *Poribacteria* families, including, those associated with nitrogen reduction, thiosulfate oxidation, and the transport of glycine betaine, proline, glycerol, taurine, tungstate, and lipooligosaccharides. Contrastingly, genes that were present in the two *Poribacteria* families, yet absent in the Tethybacteria, included those encoding trehalose biosynthesis, galactose degradation, phosphate metabolism, assimilatory sulfate reduction, and the transport of phosphonate, urea, iron complexes, molybdate, and hydroxymethylpyrimidine. Furthermore, an analysis of the divergence patterns associated with *Tethybacterales* and *Poribacteria* genomes revealed that the patterns of divergences did not follow that of their sponge hosts, thus, suggesting that although these two groups of symbionts were ubiquitous within their hosts, they were most likely acquired at different points in their evolutionary lineages to fulfil distinct roles within their sponge hosts (Waterworth *et al*., 2021).

1.9 Research rationale and aims

Algoa Bay, South Africa is home to taxonomically diverse groups of marine sponges including those of the *Latrunculiidae* family from which a number of the bioactive compounds known as pyrroloiminoquinones have been isolated. The microbial communities associated with sponges of this family are highly conserved and dominated by Gammaproteobacteria symbionts. Within the well-studied *Tsitsikamma* genus are sponge microbial communities associated with co-dominant Gammaproteobacteria and Spirochete symbionts. Additionally, within the *Tsitsikamma favus* and *Tsitsikamma michaeli* species' two distinct chemical profiles have been identified – *T. favus* sponges of chemotype I are abundant in discorhabdins and tsitsikammamines, and chemotype II abundant in makaluvamines and the amino acid phenylalanine, while *T. michaeli* sponges of chemotype I are abundant in tridiscorhabdins and chemotype II are abundant in didiscorhabdins and hydroxydiscorhabdins. Makaluvamines have previously been identified in organisms unrelated to marine sponges, namely the myxomycete *Didymium iridis* and *Didymium bahiensis*, thus supporting a proposed microbial origin for these and related marine natural products.

The genome of the highly conserved bacterial Gammaproteobacteria symbiont Sp02-1 closely associated with the microbiome of *T. favus* sponges has recently been characterised (Waterworth *et al.,* 2021). The functional potential of this symbiont has since been elucidated and several genes corresponding to denitrification, nitrate reduction, thiosulfur oxidation and phosphate metabolism have been identified, suggesting the reasons for its acquisition by a broad range of hosts including *T. favus* sponges, may be facilitating these metabolic processes within the sponge hosts. The functional potential of the co-dominant Spirochete symbiont in *T. favus* sponges has not been elucidated but is proposed to be linked to the production of pyrroloiminoquinones in *T. favus* sponges (Matcher *et al.,* 2017).

The overall aim of this study was therefore to understand the nature of the symbiotic relationships between marine sponges of the *Latrunculiidae* family with respect to their conserved spirochete symbiont and the biosynthesis of pyrroloiminoquinones. To achieve this, the following objectives were targeted:

1) Characterization of the bacterial communities associated with latrunculid sponges through 16S rRNA amplification
2) Analysis of the genome of the spirochete symbiont in *Tsitsikamma* sponges recovered from metagenomic data
3) Comparative phylogenomic assessment of free-living and symbiotic spirochetes within the Spirochaetaceae family

Chapter 2: Collection of sponge specimens and development of research methodology

2.1 Introduction

Amplicon sequencing libraries generated from 16S rRNA gene sequences have long-since been useful for the characterization of sponge-associated bacterial communities (Fernández *et al.,* 2019), however, the disadvantage thereof is that the method does not shed light on the functional roles these bacterial symbionts play in their host environment (Matchado *et al.,* 2024). Investigation of sponge-associated bacteria via culture-dependent methods is extremely limited, with only 0.1 – 14% of the sponge-associated bacteria predicted to be culturable (Dat *et al*., 2021). This low bacterial viability under laboratory conditions has been attributed to the inability to accurately reproduce the intricate environmental conditions present within the host (Clooney *et al*., 2016; Christensen and Martin, 2017). To overcome these limitations, one can employ culture-independent methods, chief among which is shotgun metagenomics. The method involves the fragmentation of all genomic DNA within a single sample to generate short or long sequence reads (Clooney *et al.,* 2016; Sharpton 2014). Short sequence reads can be generated using either MiSeq (Illumina) or Ion Torrent platforms both of which require DNA fragments to generate reads of 200 to 600 base pairs in length (Ravi *et al*., 2018; Tlili *et al.*, 2018). The sequence of nucleotides of the DNA fragments are detected via a change in pH in the Ion-Torrent platform, and via fluorescence in the Illumina platform. While the fragments are only small, the error rates of the MiSeq and IonTorrent platforms are 0.8% and 1.78% respectively (Sahlin and Medvedev, 2021), resulting in highly reliable libraries. The Oxford Nanopore sequencing platform works by allowing DNA to move through tiny holes (nanopores) in a flow cell, where each hole is equipped with an electrode sensor. DNA moving through the nanopore disrupts the current, which are read and deciphered via the sensor into human readable nucleotide sequence. This can result in reads thousands of base-pairs in lengths, but the error rate is 14% (Sahlin and Medvedev, 2021), leading to long but less reliable read libraries (Wang *et al*., 2021). Similarly, PacBio sequencing platform works by generating longer reads than the Oxford Nanopore averaging reads of 5 to 60 kilobases in length, with an error rate of 15% (Xie *et al*., 2020; Pourmohammadi *et al*., 2021). These sequence reads can then be assembled into contiguous sequences (contigs), or circular consensus sequences (CCS) with respect to PacBio sequencing (Rhoads and Au, 2015) and clustered into bacterial genomic bins or eukaryotic genomes using genome binning software such as Autometa (Miller *et al.*, 2017). The resultant genomic bins or metagenome-assembled genomes (MAGs) can then be annotated and characterized to determine their functional potential.

2.2 Development of research pipeline

The aim of this chapter was to develop a reproducible and robust workflow that would allow us to investigate the biological processes within the bacterial communities associated with marine sponges. The pipeline (Figure 2.1) begins with analysis of 16S rRNA generated amplicon sequencing libraries for profiling of sponge-associated microbial communities (Fernández *et al.,* 2019) and identification of bacterial symbionts of interest. The second stage of the pipeline involves the use of metagenomic-based approaches to elucidate the functional and biosynthetic potential of bacterial symbionts of interest and their role in the overall health and function of the sponge holobiont (Clooney *et al*., 2016; Christensen and Martin, 2017; Waterworth, 2018).

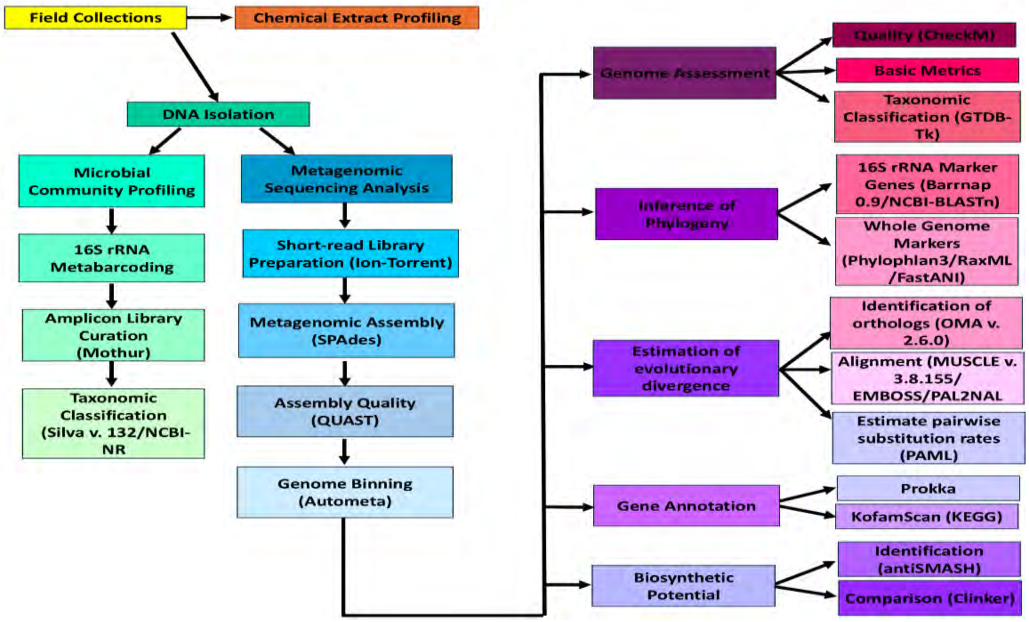

Figure 2.1: Summarized overview of the research pipeline developed for analysis sponge-associated microbial communities and elucidating their functional and biosynthetic potential.

2.2.1 Sponge collection, chemotype and taxonomic identification

Sponge specimens of the *Latrunculiidae* family were collected between 1994 and 2022 by SCUBA or Remotely Operated Vehicle (ROV) from multiple sites within the Tsitsikamma Marine Protected Area, Algoa Bay (Port Elizabeth), the Amathole Marine Protected Area (East London) and the Garden Route National Park. In addition, three *Latrunculia apicalis* specimens were collected by trawl net off the coast of Bouvet Island in the Southern Ocean. Collection permits were acquired prior to collections from the Department of Environment, Forestry and Fisheries (EFF). Sponge specimens were stored on ice and transferred to -20 °C until processing. Subsamples of each sponge, approximately 5 cm^3 in size, were preserved in RNAlater Solution (Invitrogen) and stored at -80°C. Dates and collection site data for each sponge specimen used in this study are listed in collection metadata Appendix A – Table A1. Sponge specimens were identified through morphological inspection and spicule analysis by sponge taxonomist and reef ecologist Dr Shirley Parker-Nance at the South African Environmental Observation Network, Elwandle Coastal Node (SAEON) (Nelson Mandela University, Gqeberha, South Africa).

2.2.2 Preparation of sponge extracts

Crude chemical extracts of the 79 latrunculid sponges were prepared by postdoctoral fellow Dr Jarmo Kalinski (Rhodes University, Makhanda) via extraction with methanol, drying and suspension in methanol at 1-10 mg/mL. Liquid-chromatography-Mass-spectrometry/Mass-spectrometry (LC-MS/MS) data was acquired on a Bruker ESI-Q-TOF Compact (Bruker, Bremen) in positive ionization mode coupled to a Dionex Ultimate3000 Chromatograph (ThermoScientific, Sunnyvale, CA, USA) and using reversed-phase C18 columns and mobile phases consisting of water and acetonitrile with 0.1% formic acid each. The data was converted to .mzXML format and analysed using MZmine3 (Schmid *et al.*, 2023) to assemble an aligned feature list. The feature list was filtered based on comparison of *m/z* values and MS/MS spectra to known or putative pyrroloiminoquinones. Peak area values were normalized to the overall pyrroloiminoquinone signal per sample and aggregated to the pyrroloiminoquinone class to summarise the latrunculid pyrroloiminoquinone profiles.

2.2.3 Extraction of Genomic DNA

Small sections (approximately 1 cm^3 in size) of prepared subsamples were aseptically cut, and total genomic DNA (gDNA) was extracted using the Zymo Research Quick DNA Fecal/Soil Microbe Miniprep Kit (Catalogue number: D6012) according to the manufacturer's specifications with the following modifications: during the cell lysis step, the tubes lysis tubes were processed in the bead beater at maximum speed for 10 to 15 minutes and the gDNA was eluted in 40 µl of the elution buffer provided. The concentration of the extracted gDNA for each sample was measured using the Nanodrop Spectrophotometer (ThermoFisher Scientific, Catalogue number: ND-2000) and stored at - 4 degrees Celsius. The quality and integrity of the gDNA was determined using agarose gel electrophoresis by running 2 µl of the gDNA through a 1% SeaKem® LE agarose (Lonza) gel stained with Invitrogen SYBR^{TM} Safe DNA gel stain (ThermoFisher Scientific) in 1X TAE buffer: 40 mM Tris-base (Sigma-Aldric), 1mM EDTA (Sigma-Aldric) and 20 mM glacial acetic acid (Sigma-Aldric). The resulting gel was visualized on a UV transilluminator (ChemiDocTM XRS+, Bio-Rad).

2.2.4 Illumina amplicon library preparation

To characterize the microbial communities associated with the latrunculid sponge specimens, the bacterial 16S rRNA genes were analyzed by Next Generation Sequencing (Johnson *et al.,* 2019) at the South African Institute for Aquatic Biodiversity (SAIAB) in Makhanda for processing by instrument scientist Dr Gwynneth Matcher. Amplification of the 470 bp spanning the V4 – V5 hypervariable regions of the bacterial 16S rRNA genes was performed using the primer pair MiSeq16Sa-F (Flip) (5' –TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GTA AGG TTC YTC GCG T – 3') and MiSeq16Sa-R (Flip) (5' – GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GCA GCC GCG GTA A – 3'). Polymerase Chain Reaction (PCR) amplification was performed using 0.3 ⌠M primers, 0.3 mM dNTPs, 1X Buffer with MgCl2, and 0.5 U Kapa Hifi Hotstart Taq polymerase (KAPA Biosystems) and 50ng of template DNA in reaction volumes of 50 ⌠l with adjustments made according to the manufacturer's specifications. PCR cycle parameters were as follows: initial denaturation and enzyme activation at 98°C for 5 minutes, five cycles of 98°C for 45 seconds and 45°C for 30 seconds, followed by eighteen cycles of 72°C for 1 minute, 98°C for 45 seconds and 50°C for thirty seconds, five cycles of 72°C for 1 minute, eighteen cycles of 98°C for 45 seconds, 50°C for 30 seconds and 72°C for one minute with a final extension at 72°C for 5 minutes. Resultant amplicons were visualized through agarose gel electrophoresis using

1% agarose gel in 1X TAE buffer (0.1 M Tris-HCl, 2.5 mM EDTA, 0.05 M Glacial acetic acid, Triple distilled water) stained with SYBER SAFE and run at 110V for 1 hour. Resultant PCR products were purified using the FavorPrep Gel/PCR Purification Kit according to the manufacturer's specifications (Catalogue number: FAGCK001). Purified products containing 30 ng of DNA were sequenced at the South African Institute for Aquatic Biodiversity (SAIAB) Aquatic Genomic Research Platform using the Illumina MiSeq technology.

2.2.5 Amplicon library sequence data curation

Amplicon library sequence datasets of the forward reads (214 base pairs) of 155 latrunculid sponges and 8 seawater samples collected between 1994 and 2022 were curated using the Mothur platform (release version 144.3) (Schloss *et al.,* 1999) as previously described by Matcher *et al.* (2017). Low-quality sequence reads shorter than 250 base pairs and containing ambiguous bases and/or homopolymeric reads greater than 7 were discarded. Chimeric reads were identified using the VSEARCH algorithm (Rognes *et al.,* 2016) and discarded. Sequence reads were classified using the Naïve Bayesian classifier via alignment against the SILVA database (release version 132) and any sequences classified as "Chloroplast", "Mitochondria", "Unknown", "Archaea" or "Eukaryota" were removed. A distance matrix (cut-off of 0.05) was generated in Mothur and used to classify the sequence reads into operational taxonomic units (OTUs) at distance values of 0.03 (species level) and read counts were converted to relative abundances (data available in Appendix A – Table A2). Spirochete OTUs were subset out and aligned with reference sequences from the NCBI nucleotide database using MUSCLE (v. 5.8.1). Phylogenetic trees of the spirochete OTU and reference sequence alignments were generated using MEGA11 software (Tamura *et al*., 2021) using ClustalW (Larkin *et al.,* 2007; Goujoun *et al.,* 2010) and Maximum-likelihood algorithms, with 1000 bootstrap replicates. The resultant trees were then visualized in the Interactive Tree of Life (iTol) (Letunic and Bork, 2021). The same analysis was then repeated with only the raw amplicon reads from latrunculid sponges, focusing on OTUs generated and clustered at a distance of 0.01 (strain-level).

Three-dimensional non-metric dimensional scaling (NMDS) plots of the relative abundance data were generated in R using the *vegan* R package based on the Bray-Curtis dissimilarity index (Bray and Curtis, 1957) (custom scripts – Appendix B Custom scripts). Resultant NMDS/ANOSIM data is available in Appendix A – Table A4 and A5. The co-correlation analysis of the 50 most abundant OTUs (found as an average across all samples) was performed using the '*cor'* function (Langfeld and Horvath, 2012) native to R using dataframes of OTU and compound abundances as input (custom scripts – Appendix B Custom scripts). A 16S rRNA gene sequence phylogeny was built from the representative sequences of the top 50 OTUs, aligned with MUSCLE (version 5.1) (Edgar, 2004; Edgar, 2022), using the Neighbor-joining approach with 1000 bootstraps in MEGA11 (Tamura, *et al*., 2021). The final tree was visualized in iTol (Letunic and Bork, 2021) where the correlation matrix and the average OTU abundance per sponge species was visualized alongside the tree as datasets. Compound absorbance data for the latrunculid sponges are available in Appendix A – Table A6.

2.2.6 Metagenomic sequencing and analysis of *Tsitsikamma* specimens

Eight *Tsitsikamma* sponge specimens, collected between 2016 and 2022, were selected for metagenomic sequence analysis to identify the functional and biosynthetic potential of the associated microbiota (Clooney *et al*., 2016; Christensen and Martin, 2017). Shotgun metagenomic sequencing was performed using 600 ng of gDNA per sample with a $A_{260/280}$ ratio between 1.8 and 2.0 (see Appendix A – Table A7 for collection metadata) using Ion Torrent platforms at the Central Analytics Facility (CAF), Stellenbosch. Shotgun metagenomic libraries were prepared on all samples using an Ion P1.1.17 chip. Resultant libraries had an average of 70 million reads per sample, with an average read length of 200 bp. Additional sequence data of 400 bp was generated for sample TIC2016-050A using an Ion S5 – 530 chip as previously described by Waterworth *et al.* (2021). Metagenomic datasets for all eight samples were then assembled into contiguous sequences (contigs) with SPAdes v3.12.0 (Bankevic *et al.,* 2012) using the -iontorrent and -only-assembler options. The quality of the scaffold assemblies such as the total number of contigs, longest contig, length of the minimum number of contigs that cover the length of half the assembly (N50), and number of contigs equal to or longer than the N50 (L50) were determined using QUAST (Table 2.1) (Gurevich *et al.*,2013; Manchanda *et al.,* 2020; Price and Syme, 2024). Scaffold files with N50 values greater than 1000 kilobases were considered good quality assemblies and were prioritized for genomic binning. Scaffold contigs were classified using Autometa (Miller *et al*., 2019) and all "Archaea", "Eukaryota", "Viral" and "Unclassified" contigs were discarded.

Table 2.1: Metagenomics sequencing data generated for *Tsitsikamma* sponge sample metagenomes

Sample	**No. of reads generated**	**Average read length (bp)**	**Scaffold length (bp)**	**N50**	**L50**
TIC2016-050A	*12 710 484	363	-	-	-
	*89 638 524	188	-	-	-
	*14 537 738	329	429936 329	2 183	41 209
TIC2016-050C	29 149 091	183	251 350 367	1 579	3 6914
TIC2018-003B	46 930 766	185	229 633 985	2788	18134
TIC2018-003D	56 303 358	187	194 923 371	1748	18128
TIC2018-003M	85 299 678	182	392 030 523	1 594	50 989
TIC2019-013N	79 444 584	182	307 847 516	1610	41 189
TIC2022-009	90 386 386	182	492 072 875	1 880	62 643
TIC2022-059	88 310 820	183	254 743 103	3 417	17002

*The three TIC2016-050A libraries were combined into a single hybrid assembly.

2.2.7 Acquisition of reference genomes and MAGs

Four spirochete MAGs assembled from the genomes of *Aplysina aerophoba* and *Rhopaloeides odorabile* sponges from a study by Robbins and colleagues (Robbins *et al.,* 2021) were downloaded from the Australian Centre for Ecogenomics online index (University of Queensland, Australia), and five sponge-associated spirochete MAGs were acquired from the China National GeneBank DataBase (CNGBdb) from studies by O'Brien and colleagues (O'Brien *et al.,* 2021; O'Brien *et al.,* 2023). Additionally, twenty-eight other genomes characterised as members of the Spirochaetacea*e* family, three of which were host-associated (termite-associated), were selected and downloaded from the NCBI database.

2.2.8 Characterization and taxonomic identification of spirochete genomes

Quality (estimated completion and contamination) of MAGs was assessed using CheckM v1.0.12 (Parks *et al*., 2015). Of the 35 recovered MAGs, 5 were of high quality, 12 were of medium quality and 18 were of low quality according to MIMAG standards (Bowers *et al.,* 2017). A full list of recovered genome assembly data is available in Appendix A – Table A7. The recovered MAGs were taxonomically classified with GTDB-Tk database release version

95 (Chaumeil *et al.,* 2019) and only the MAGs classified as members of the phylum "Spirochaetota" were selected. Basic metrics such as genome size, number of contigs, N50, and GC content (%) for each spirochete MAG and reference genome were calculated using a custom script bin_summary.py. The number of genes, pseudogenes, and coding density per genome were calculated using all_included_genomics_characteristics.py. Custom scripts were generated by Dr Samantha C. Waterworth and are listed in Appendix B – Custom scripts.

2.2.9 Inference of phylogeny from latrunculid-associated spirochete MAGs

Ribosomal gene sequences (16S and 23S rRNA) were extracted from the spirochete MAGs using Barrnap 0.9 (Seemann and Both, 2018). The 16S rRNA sequences were aligned against the BLAST nucleotide database (BLASTn) to identify the closest matches. The resultant sequences were aligned using MUSCLE (v.5.1.) and phylogeny was inferred using the Neighbour-Joining method with 1000 bootstrap replicates in MEGA11 (Tamura *et al.,* 2021)

Phylogeny of the latrunculid-associated spirochete MAGs was additionally inferred using conserved genome markers from whole genomes using PhylopPhlan3 (Asnicar *et al.,* 2020) and RaxML (Stamatakis, 2014). These tools were selected as they provided a genome-to-genome comparison of entire spirochete genomes and allow a better overview of the evolutionary and functional relationships between the different spirochete genomes (Asnicar *et al.,* 2020; Stamatakis, 2014) The phylogeny of all 8 spirochete MAGs and 37 reference genomes was inferred using PhyloPhlan3: PhyloPhlan3 run with diversity set to medium, with default values in supermatrix_aa configuration. The resultant marker gene protein alignment was used in RaxML to generate a phylogenetic tree with 1000 bootstrap replicates using the PROTGAMMAAUTO model (scripts for all three programmes are available in Appendix B – Custom Scripts). The resultant tree was visualized in iTol (Letunic and Bork, 2021). The genome from *Leptonema illini* (GCA_002009735.1) was selected as an outgroup to root the tree as it was not a member of the Spirochaetace*ae* family. Pairwise average nucleotide identify (ANI) between all sponge-associated genomes was calculated using FastANI (Jain *et al.,* 2018). Only pairwise alignment fractions above 70% were considered as reliable ANI scores (Gosselin *et al.,* 2022). Resultant pairwise ANI values were visualized using ANIclustermap (v.1.2.0) (Shimoya, 2022).

2.2.10 Estimation of relative evolutionary divergence of spirochete genomes

The output phylogeny generated by PhyloPhlan3 (Section 3.2.6) was used to determine the closest relatives of the *Tsitsikamma*-associated spirochetes and the other sponge-associated spirochetes. The Orthologous Matrix (OMA) (v.2.6.0) (Altenhoff *et al*., 2019) was used to identify orthologous (shared) genes between the *Tsitsikamma*-associated spirochetes, other sponge-associated spirochetes genomes, as well as the genomes of their closest relatives. A total of 12 orthologs common to all genomes were found using the count_OGs.py script. The resultant orthologous genes were aligned using MUSCLE (v.3.8.155) (Edgar, 2004). The corresponding nucleotide sequence for each gene was retrieved using the streamlined_seqretriever.py script, all stop codons were removed using remove_stop_codons.py script, and sequences were aligned using MUSCLE (v.3.8.155) (Edgar, 2004). The gene sequences for each ortholog were grouped per genome using merge_fasta_for_dNds.py script. The nucleotide and amino acid sequences for each genome were concatenated, respectively, using the union function from EMBOSS (Rice *et al.,* 2000). The concatenated nucleotide and concatenated amino acid sequences were aligned against one another using PAL2NAL (Suyama *et al.,* 2004). The alignment was used to estimate pairwise synonymous substitution rates (dS) and infer the pattern of divergence between all the genomes using codeml from the PAML package (Yang, 2007). The resultant pairwise substitution rates (dS) were visualized on MEGA11 (Tamura *et al.,* 2021)

2.2.11 Annotation of spirochete genomes

Amino acid sequences and nucleotide sequences for all genes in all MAGs and genomes used in this study were detected and annotated using Prokka (Seemann, 2014). All genes were additionally annotated against the KEGG database using KOfamSCAN with detail-tsv as the output format (Amaraki *et al*., 2020). Reliable annotations were then extracted based on the criteria that the annotation score was greater than the estimated threshold. All reliable annotations per genome were counted and summarized using the kegg_parser.py script. This generated a table of the KO counts per genome (Appendix A – Table A8) that was used for 3-dimensional visualization of the data and analysis-of-similarity assessment (ANOSIM) (Appendix A – Table A9) as well as the construction of a presence/absence heatmap in R (script for generating the 3D NMDS plot, ANOSIM assessment and presence/absence maps available in Appendix B – Custom Scripts).

2.2.12 Identification of putative biosynthetic gene clusters in spirochete genomes

To identify any putative biosynthetic gene clusters and their predicted compounds all spirochete MAGs and reference genomes from *Tsitsikamma* sponges and the reference sponges were manually uploaded onto antiSMASH (online v.7.0) (Blin *et al.,* 2023) with relaxed detection strictness and --cb-general --cb-knownclusters --cb-subclusters --asf --pfam2go --smcog-trees options enabled. The resultant putative biosynthetic gene clusters were identified, and clusters of interest were downloaded in GenBank format. Clusters of interest were visualized and manually annotated using clinker (Gilchrist and Chooi, 2021).

Chapter 3: Elucidating the functional potential of the spirochete symbiont in latrunculid sponges

The bacterial communities associated with *T. favus* sponges (Order Poecilosclerida) were previously shown to be stable, highly conserved (Walmsley *et al.,* 2012a) and dominated by members of the class Betaproteobacteria (recently reclassified as an order within the class Gammaproteobacteria) (Park *et al*., 2018), as well as the Spirochaetes, Gammaproteobacteria, Alphaproteobacteria, Planctomycetes and Verrucomicrobia phyla (Walmsley *et al*., 2012b). The dominant Betaproteobacteria associated with *Tsitsikamma* sponges and other members of the Latrunculiidae family was shown to be a single species with unique strains specific to each sponge species (Matcher *et al.,* 2017). The dominant Betaproteobacterium in *T. favus*, *Candidatus* Ukwabelana africanus, has been characterized as a member of the newly proposed *Tethybacterales* order that includes well-studied betaproteobacterial symbionts such as *Amphirhobacter heroislandensis* AqS2, a symbiont of the Australian coral reef sponge *Amphimedon queenslandica* (Waterworth *et al.,* 2021; Gauthier *et al*., 2016). These symbionts have since loosely been referred to as 'Tethybacteria'. Unlike the Tethybacteria, the dominant Spirochaete population reported in *T. favus* sponges is exclusive to sponges of the *Tsitsikamma* and *Cyclacanthia* genera (Matcher *et al*., 2017). Analysis of 16S rRNA gene amplicon libraries of these two sponge-associated microbial communities at strain level (OTUs clustered at a distance of 0.01) revealed the presence of two distinct spirochete species in *T. favus* sponges (Matcher *et al*., 2017), which were most closely related to clones Sp02-3 and Sp02-15 previously isolated from *T. favus* sponges (Walmsley *et al*., 2012a).

Numerically dominant spirochete populations have also been reported in a number of marine invertebrates such as corals and sea stars (Loudon *et al.,* 2023), molluscs (Kuhn 1981, Margulis and Hinkle 2006; Paster 2010; Sitnikova *et al.,* 2012), arthropods (Graña-Miraglia *et al.,* 2020) and the hindguts of termites (Lilburn *et al.,* 2001). Aside from *Tsitsikamma* and *Cyclacanthia* sponges, *Clathrina clathrus* is the only other sponge in which dominant spirochete populations have been documented (Neulinger *et al*., 2010; van der Water *et al.,* 2016). The functional potential of these spirochete populations remains unclear; however, it is speculated that they may be involved in the fixation of carbon and nitrogen in the red coral *Corallium rubrum* (van der Water *et al*., 2016), sea star *Pisaster orchraceus* (Lilburn *et al.,* 2001) as well as termites

Reticulitermes speratus (Rhinotermitidae), *Hodotermopsis sjoestedti* (Termopsidae), and *Mastotermes darwiniensis* (Iida *et al.,* 2000). While the role within sponges is currently unknown, the dominant spirochete OTUs in *T. favus* and *Cyclacanthia* spirochetes are hypothesized to fulfil specialized functions within their hosts. This hypothesis was borne of the fact that these spirochetes are not phylogenetically closely related to other sponge-associated spirochetes, and the exclusive presence of these spirochetes in latrunculid sponges was suggestive of possible the spirochetes being the biogenic source of pyrroloiminoquinones within these hosts (Matcher *et al*., 2017).

The aim of the research presented in this chapter was to determine the distribution and phylogenetic relatedness of spirochete populations in latrunculid sponges, relative to sympatric sponge species and further determine their potential role in the sponge holobiont.

3.2 Assessment of microbial communities present in latrunculid and non-latrunculid sponges endemic to the South African coastline

To investigate the presence of spirochete symbionts in sponge specimens collected to the South African coastline (including both latrunculid and non-latrunculid sponges), a number of collection sites were selected along the South-Eastern coastline of South Africa. Specimens were primarily collected from reefs within Algoa Bay (Gqeberha, formerly known as Port Elizabeth), but other collection sites included Tsitsikamma Marine Protected Area, the Amathole Marine Protected Area (East London), Sodwana Bay (KwaZulu-Natal) and the remote Bouvet Island in the Antarctic Ocean. The microbial communities of 155 marine sponges and 8 seawater samples were assessed using 16S rRNA gene fragment amplicons clustered into operational taxonomic units (OTUs) at a distance of 0.03 (Schloss *et al.,* 1999). A total of 13,012,208 reads were obtained from all 163 samples. A total of 9711 OTUs were subsequently recovered from the resultant amplicon libraries. Classification of these OTUs through alignment of the OTU sequences against the SILVA and nr databases revealed 142 OTUs were classified as members of phylum Spirochaetota, of which only 10 OTUs had an average abundance greater than 0.01% across all 163 sponge specimens (Appendix A - Table A2). OTU3 and OTU59 were the most abundant in the *Tsitsikamma* and *Cyclacanthia* sponges, with very low abundances in *Latrunculia apicalis* and *L. algoaensis* sponges (collected in Algoa Bay as well and the Antarctic Ocean), as well as in some *Mycale* specimens and a single sympatric *Phorbas* sp. sponge (Figure 3.1A). Alignment of the consensus sequences of these

OTUs revealed that these were most closely related to spirochete 16S rRNA clones Sp02-3 and Sp02-15, previously identified in *T. favus* sponges (Walmsley *et al.,* 2012a; Matcher *et al.,* 2017). As the *Mycale* specimens were found as encrusting species on the *Tsitsikamma* sponges, it is likely that these OTUs were the result of possible contamination. As there was only a single *Phorbas* specimen, additional specimens would be required to confirm the presence of these OTUs in this species of sponge and discredit their presence as the result of possible contamination. These two OTUs were otherwise absent from any other non-latrunculid sponge specimens collected from the same regions (Figure 3.1B). Despite the considerably low abundances of OTU3 and OTU59 in *L. apicalis* sponges collected off the coast of Bouvet Island in the South Atlantic Island (approximately 3000 kilometres from Algoa Bay), the presence of these spirochete OTUs and phylogenetically distinct spirochete OTUs in sympatric non-latrunculid sponges suggest that these Sp02-3 and Sp02-15 OTUs are specific to latrunculid sponges. Inspection of the phylogeny of these ten spirochete OTUs (Figure 3.1C) revealed six of the ten OTUs formed a clade with spirochete clones Sp02-3 and Sp02-15 previously identified in *T. favus* sponges (Walmsley *et al.,* 2012a; Matcher *et al.,* 2017). The remaining four appear to form two separate clades distantly related to other sponge-associated spirochetes, with OTU105 and OTU128, the most abundant OTUs in the non-latrunculid sponges, clustering between a clade containing spirochetes from *Clathrina clathrus* sponges and a clade containing free-living spirochetes isolated from soil, alkaline lakes and hypersaline environments.

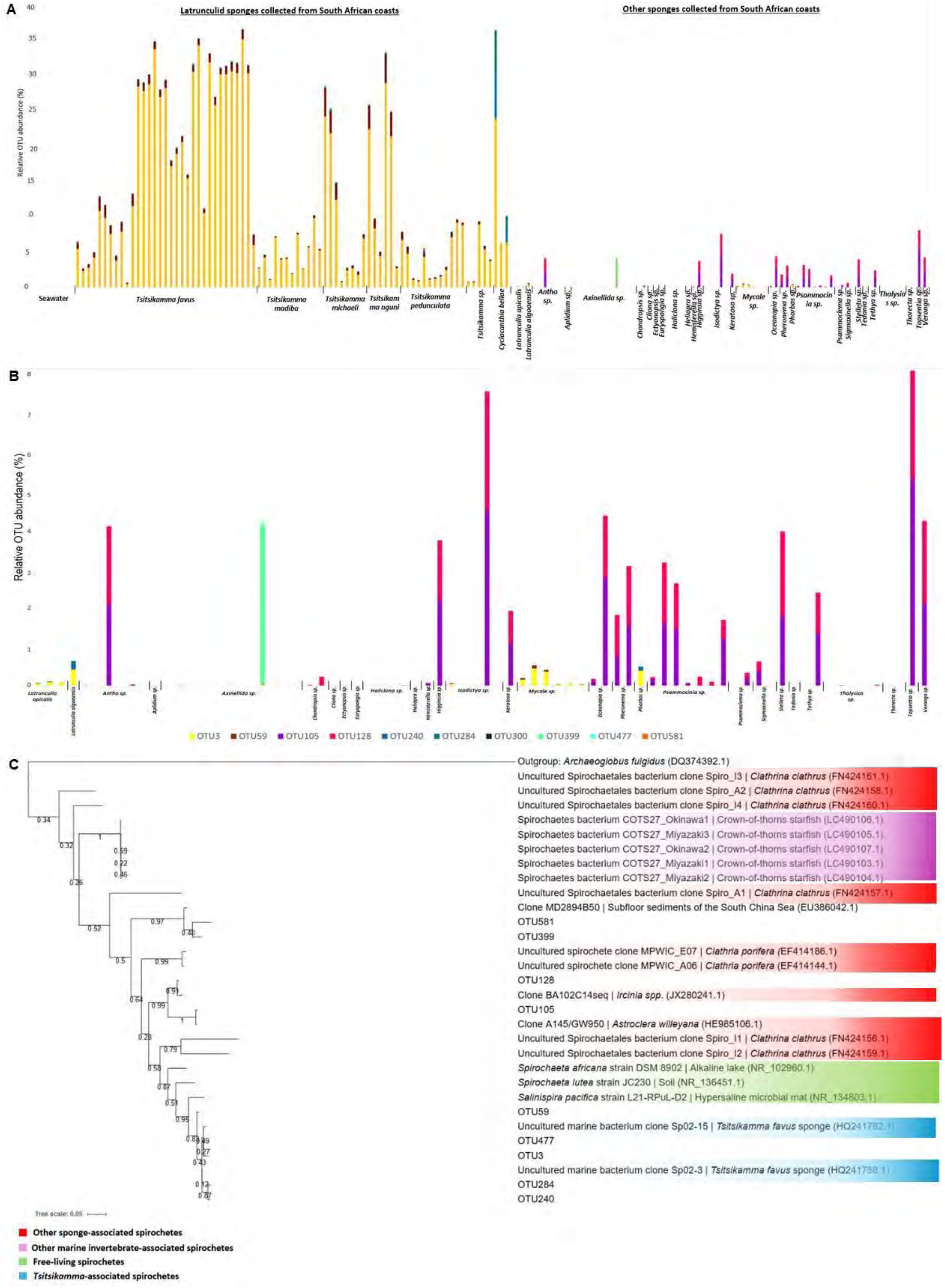

Figure 3.1: Distribution of spirochete populations across sponges collected from South African coasts and the Antarctic Southern Ocean. A) Relative abundance of OTUs classified as spirochetes clustered at a distance of 0.03. B) A magnified view of the spirochete

OTUs present in non-latrunculid sponges collected from the south-eastern coast of South Africa, as well as three *L. apicalis* sponges and one sympatric *L. algoaensis* sponge. C) Maximum-likelihood phylogeny of the top ten most abundant spirochete OTUs identified in the sponges included in this study conducted with 1000 bootstraps. Bootstrap values are indicated as decimals on the branch nodes and a scale bar is shown indicating the average number of nucleotide substitutions per site. The 16S rRNA sequence of *Archaeoglobus fulgidus* (GenBank accession number: DQ374392.1) was used as an outgroup. A colour key is used to represent the *Tsitsikamma*-associated spirochetes, free-living spirochetes and other sponge-associated spirochetes.

3.3 Characterization of the associated microbial communities in marine sponges of the Latrunculiidae family

The microbial communities of 79 sponges within the Latrunculiidae family were assessed using their 16S rRNA gene amplicon sequences. Clustering of these sequences at a distance of 0.01 (strain level) resulted in 17052 Operational Taxonomic Units (OTUs). This analysis revealed distinct microbial community profiles in the sponge species (Figure 3.2). OTUs representative of the dominant and highly conserved Tethybacterales population previously reported by Waterworth *et al.* (2021) were present in all sponges. OTU1, OTU2, OTU4, OTU9 and OTU20 were present across all 73 *Tsitsikamma* sponges, however, the distributions thereof varied according to the different sponge species. Alignment of the consensus sequences of these OTUs revealed that these were likely strains of the Tethybacterales symbionts (Sp02-1) previously identified in latrunculid sponges (Matcher *et al.*, 2017; Walmsley *et al.*, 2012a; Waterworth *et al.,* 2017). The distribution of these Tethybacterales symbionts was as follows: OTU1$_{0.01}$ was dominant in *T. favus, T. nguni* and *T. pedunculata* sponges, and OTU2$_{0.01}$ was dominant in *T. madiba, T. michaeli* and *C. bellae* sponges. The other three conserved OTUs (OTU4, OTU9 and OTU20) were secondary dominant populations of the Tethybacteria, where OTU4$_{0.01}$ was most abundant in *T. pedunculata* sponges, OTU9$_{0.01}$ was abundant in *T. madiba* and *T. michaeli* sponges and OTU20$_{0.01}$ was abundant in *T. pedunculata* sponges only.

Similarly, all latrunculid sponges collected from the South African coast included populations representative of the spirochete clones Sp02-3 and Sp02-15 previously isolated from *T. favus* specimen TIC2009-002 (Walmsley *et al.,* 2012a; Matcher *et al.,* 2017) (Figure 3.2A). Three spirochete representatives, OTU5$_{0.01}$, OTU6$_{0.01}$ and OTU7$_{0.01}$, were conserved across six

Tsitsikamma species, however, were present at very low abundances (less than 1%) in *C. bellae* and *L. apicalis* sponges. Pairwise analysis of similarity (ANOSIM) testing showed a statistically significant difference between the microbial communities associated with the sponge species, indicative of species-specific microbial communities ($p<0.05$) (ANOSIM data available in Appendix A – Table A4) (Figure 3.2B).

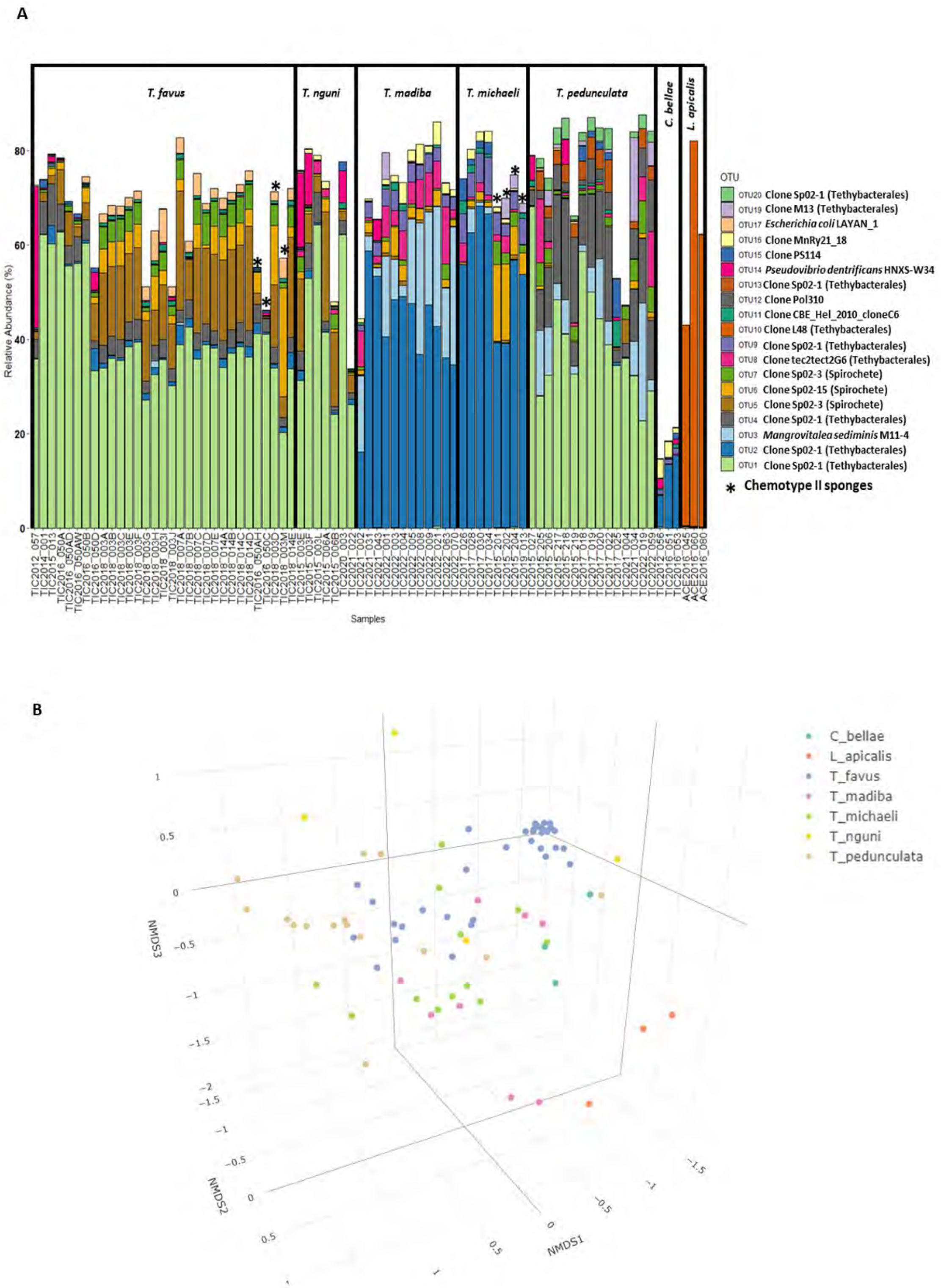

Figure 3.2: Distribution of microbial OTUs across latrunculid sponges. A) Bacterial community profiles associated with latrunculid sponges represented by OTUs clustered at a distance of 0.01. The top 20 OTUs with the highest relative abundances are shown. The closest relative (indicated with the colored key) was determined through alignment of the OTU nucleotide sequences against the NCBI-BLAST nucleotide (NR) database. B) Non-metric

multidimensional scaling plot of bacterial OTUs clustered at a distance of 0.01 showing the species-specific distribution of bacterial OTUs in latrunculid sponges. The species of each sponge are represented by a coloured key.

As shown in previous studies (Kalinski *et al*., 2019; Kalinski *et al.,* 2021; Kalinski *et al.,* 2022), much like the microbial communities, the chemical profiles present in latrunculid sponges are species-specific. Additionally, there are distinct chemotypes present in the *T. favus* and *T. nguni* sponges. Assessment of the chemical profiles associated with the sponges used in this study (Appendix A – Figure S1) showed that the distribution of pyrroloiminoquinones was sponge-specific. However, the *T. favus* and *T. michaeli* Chemotype II sponges were distinct ($p<0.05$) from *T. favus* and *T. michaeli* Chemotype I sponges (Appendix A – Figure S1, denoted by the asterisks) (ANOSIM data available in Appendix A – Table A5).

Inspection of the spirochete OTUs conserved across all the *Tsitsikamma* sponges (Figure 3.3) revealed that the distrubution of OTU5 and OTU6 varied between species. OTU5 was more dominant in *T. favus, T. nguni* and *T. pedunculata* sponges, accounting for 2 – 25% of all reads, and OTU6 was more dominant in *T. madiba, T. favus and T. michaeli* sponges, accounting for 1 – 22% of all reads (Figure 3.3A). Furthermore, it was noted that the abundance of OTU6 appeared greater than OTU5 in *T. favus* and *T. michaeli* sponges classified as Chemotype II (denoted with asterisks in Figure 3.2A). OTU7, an additional representative of spirochete clone Sp02-3, was dominant in *T. pedunculata sponges*. The phylogenetic relationship of these spirochete OTUs was assessed by aligning their 16S rRNA sequences against their closest matches obtained from the NCBI-BLAST nucleotide (NR) database, and sequences from spirochetes associated with other marine invertebrates (Wada *et al.,* 2020; van de Water *et al*., 2016), including the dominant spirochete associated with the distantly related *Clathrina clathrus* sponge (Neulinger *et al.,* 2010). The phylogeny was inferred using the Maximum-likelihood method, and showed that the latrunculid-associated spirochetes were distinct from other marine-associated spirochetes, yet clustered in a distinct clade with the 'uncultured marine clone Sp02-3' and 'uncultured marine clone Sp02-15' previously identified in *T. favus* isolate TIC2009-002 (Walmsley *et al.,* 2012a) as well as their closest culturable relative, *Salinispira pacifica* (GenBank Accession number: NR_134803.1), isolated from hypersaline microbial mats (Hania *et al*., 2015).

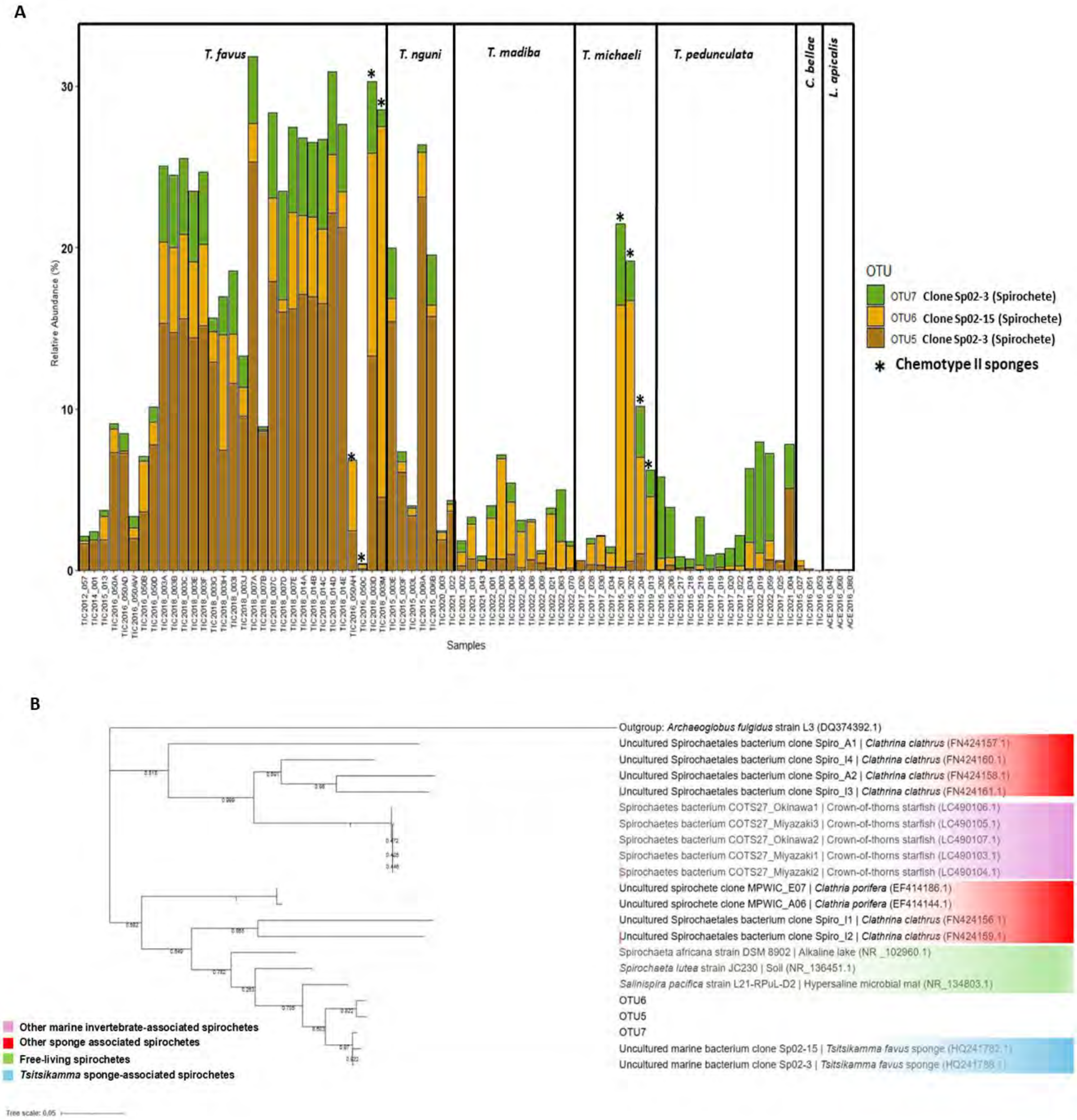

Figure 3.3: Assessment of the spirochete OTUs conserved across latrunculid sponges. A) The distribution of spirochete OTUs in latrunculid sponges B) Maximum-likelihood phylogeny of spirochete 16S rRNA sequences inferred using 1000 bootstrap replicates. Bootstrap values are indicated on the branch nodes and a scale bar is shown indicating the average number of nucleotide substitutions per site. The 16S rRNA sequence of *Archaeoglobus fulgidus* (GenBank accession number: DQ374392.1) was used as an outgroup. A colour key is used to represent the latrunculid-associated spirochetes, free-living spirochetes and other sponge-associated spirochetes.

To determine whether there was a correlation between the abundance of spirochete OTUs and the chemical profiles observed in the sponges (Appendix A – Figure S1) a correlation analysis was conducted using the top 50 most abundant OTUs with relative pyrroloiminoquinone abundances per sponge sample (Figure 3.4, Appendix A – Table A6). This was conducted by determining the correlation score per OTU-compound combination. The Sp02-3 spirochetes (OTU5, OTU7, OTU26) were positively correlated with Chemotype I pyrroloiminoquinones such as tsitsikammamines, as well as discorhadbin I and makaluvamines in *C. bellae* sponges. The Sp02-15 spirochete OTU6 was positively correlated with Chemotype II pyrroloiminoquinones that included makaluvamines, bromo-makaluvamines, bromo-orthoquinones, hydroxy-discorhabdins, and dibromo-discorhabdins. No other bacterial OTU was correlated with the change in chemotype observed in *T. favus* and *T. michaeli* sponges.

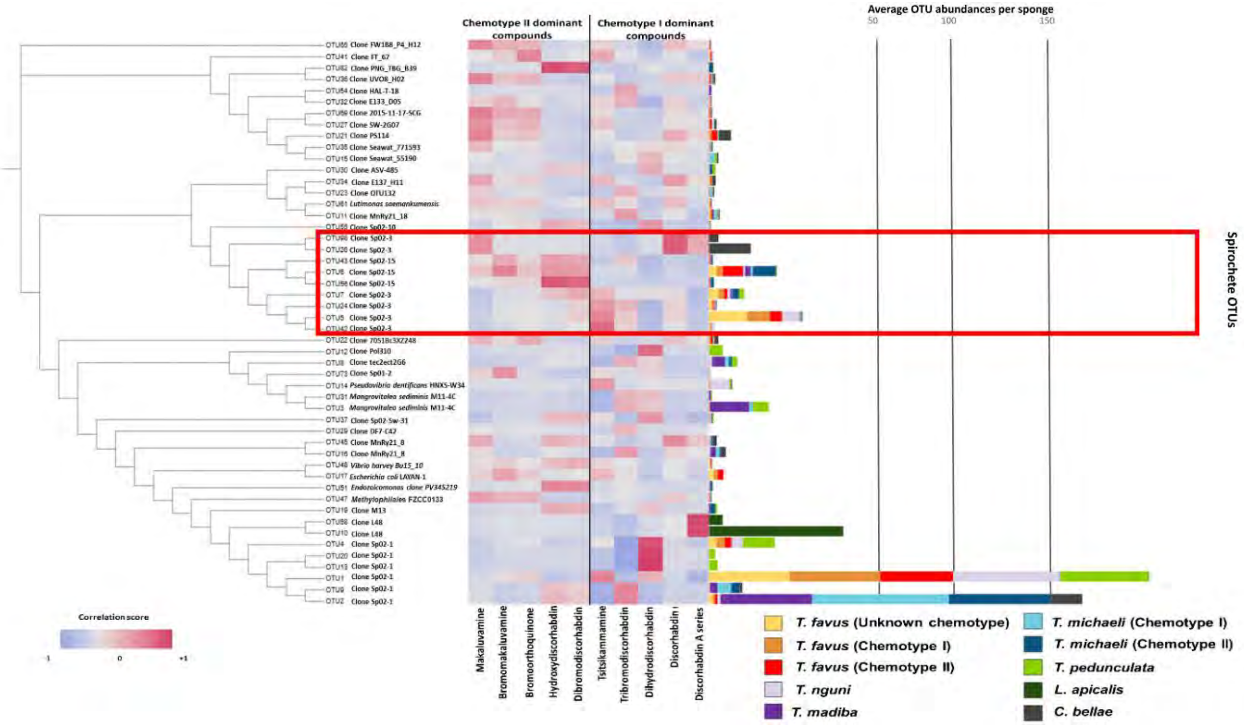

Figure 3.4: Co-correlation analysis of OTU abundance and pyrroloiminoquinone presence in latrunculid sponges shows that OTU6 is exclusively correlated with all compounds that are characteristic of Chemotype II sponges. The top 50 most abundant OTUs and their closest NCBI matches (left-hand dendrogram), along with the average OTU abundances per sponge (right-hand horizontal bar graphs) was used to assign a correlation score between -1 and 1 for each OTU-compound combination (colour-scale indicated). A negative score (represented by blue colour) represents a presence-absence relationship, and a positive score (represented by pink colour) represents a presence-presence relationship.

3.4 Characterization of spirochete genomic bins in latrunculid sponges:

To better understand the potential role of the spirochetes in the latrunculid sponges, we assembled and analysed shotgun metagenomic data from eight *Tsitsikamma* sponge specimens. The assembled contigs, from each metagenomic assembly, were binned into metagenome-assembled genomes (MAGs) using kmer frequency distribution profiles. Eight spirochete MAGs were extracted, one from each of the eight *Tsitsikamma* sponges: MAGs 050A_9, 050C_1, 003B_3, 003D_1, 003M_1, 013N_1, 009_1 and 059_1. The MAGs generated were named after the *Tsitsikamma* specimen from which they were extracted, and the numbers associated with each MAG were arbitrarily assigned during the binning process. The 16S and 23S rRNA gene sequences were extracted from individual MAGs using Barrnap 0.9. Alignment of the 16S rRNA gene sequences against the NCBI database revealed that five MAGs were representatives of the 'uncultured marine clone Sp02-3' previously identified in *T. favus* isolate TIC2009-002 (Walmsley *et al.,* 2012a), sharing above 99% sequence identity with the reference 16S rRNA gene sequence of Sp02-3 (HQ241788.1) (Table 3.1). The next closest matches from the NCBI database were from the 'uncultured marine clone Sp02-15' with a sequence identity match of 99.39%, followed by *Salinispira pacifica* strain L21-RPuL-D2 with a sequence identity match of 93.26%. A 16S rRNA gene sequence was not recovered from MAGs 003B_3, 013N_1 and 009_1. Assessment of the 23S rRNA gene sequences against the NCBI database revealed *Salinispira pacifica,* previously identified as the closest relative of Sp02-3 in South African latrunculid sponges (Walmsley *et al*., 2012b; Antunes *et al*., 2005; Matcher *et al.,* 2017), as the closest relative of the extracted *Tsitsikamma*-associated spirochete MAGs. Taxonomic classification of the MAGs against the GTDB-Tk database placed all eight genomes within the *Salinispira* genus (Appendix A – Table A7).

Table 3.1: Characteristics of putative representative genomes of latrunculid-associated spirochete symbiont species

Host	Sample	Bin	Size (Mbp)	Quality	16S rRNA (% ID)	23S rRNA (% ID)	Chemotype
T. favus	TIC2016-050A	050A_9	0.07	Low	Uncultured marine clone Sp02-3 (99.52%)	*Salinispira pacifica* L21-RPuL-D2 (89.54%)	Chemotype I
	TIC2016-050C	050C_1	1.79	Medium	Uncultured marine clone Sp02-3 (99.52%)	N/A	Chemotype II
	TIC2018-003B	003B_3	1.88	Medium	N/A	*Salinispira pacifica* L21-RPuL-D2 (89.54%)	Chemotype I
	TIC2018-003D	003D_1	2.49	High	Uncultured marine clone Sp02-3 (99.52%)	*Salinispira pacifica* L21-RPuL-D2 (89.58%)	Chemotype II
	TIC2018-003M	003M_1	2.55	High	Uncultured marine clone Sp02-3 (99.52%)	*Salinispira pacifica* L21-RPuL-D2 (89.58%)	Chemotype II
T. michaeli	TIC2019-013N	019N_1	2.33	High	N/A	*Salinispira pacifica* L21-RPuL-D2 (89.48%)	Chemotype II
	TIC2022-009	009_1	1.47	Medium	N/A	*Salinispira pacifica* L21-RPuL-D2 (91.25%)	Chemotype I
T. pedunculata	TIC2022-059	059_1	2.04	Medium	Uncultured marine clone Sp02-3 (99.11%)	*Salinispira pacifica* L21-RPuL-D2 (88.08%)	Chemotype I

3.5 Phylogeny of *Tsitsikamma*-associated spirochete genomes:

The phylogeny of all eight *Tsitsikamma* sponge-associated spirochete symbionts was inferred using the extracted 16S rRNA sequences as well as whole genome data. The recovered 16S rRNA gene sequences were aligned against the 16S rRNA sequences of the spirochete representative species OTU5 and OTU6, as well as *T. favus* isolate TIC2009-002 uncultured spirochete clones Sp02-3 or Sp02-15. Genomes 050A_9 and 059_1 clustered with OTU5 and the uncultured spirochete clone Sp02-3, while genomes 050C_1, 003D_1 and 003M_1 clustered with OTU6 and the uncultured spirochete clone Sp02-15 (Figure 3.5). This validated the distribution of strains Sp02-3 and Sp02-15 according to chemotypes, and further suggested that genomes 050A_9 and 059_1 may be representatives of clone Sp02-3 while genomes 050C_1, 003D_1 and 003M_1 may be representatives of Sp02-15 clone.

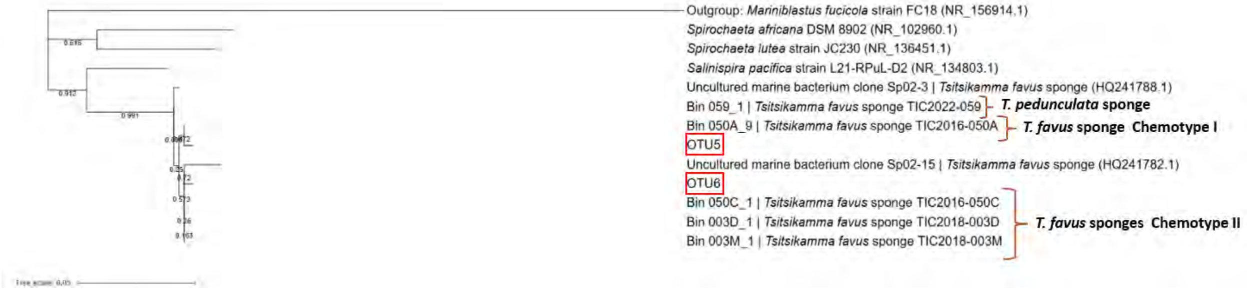

Figure 3.5: Phylogeny of spirochete 16S rRNA sequences obtained from *Tsitsikamma* sponge metagenomes and amplicon libraries. The phylogeny of 16S rRNA sequences from *Tsitsikamma* sponge-associated spirochetes follows that of the host sponge chemotypes. Maximum-likelihood method was used to infer phylogeny with 1000 bootstrap replicates. The 16S rRNA sequence of *Mariniblastus fucicola* (GenBank Accession number: NR_156914.1) was used as an outgroup to root the tree.

As an orthologous approach for investigating phylogeny at species and we calculated the pairwise average nucleotide identity (ANI) values for all eight spirochete MAGs. The ANI values were calculated using FastANI and visualized as a heatmap using ANIclustermap (v.1.2.0) (Figure 3.6). MAGs with ANI values greater than 95% are traditionally considered the same species (Rodriguez-R *et al.*, 2024). Although all *T. favus*-derived MAGs displayed ANI values greater than 95%, MAGs, there were two distinct clades of MAGs that shared greater than 97% identity: 050C_1, 003D_1, 003M_1 from *T. favus* Chemotype I sponges, and 050A_9 and 003B_3 from *T. favus* Chemotype II sponges. Interestingly, MAGs 013N_1 (*T. michaeli* host) and 059_1 (*T. pedunculata* host) displayed ANI values of 95.3% despite having been isolated from two distinct *Tsitsikamma* sponge species. A previous study by Kalinski *et al.* (2021) reported high levels of discorhabdin production in *T. pedunculata* sponges, similar to those present in *T. michaeli* sponges, thereby, providing a possible reason for the potential similarity of MAG 059_1 with MAG 013N, as both would have to adapt to similar chemical environments.

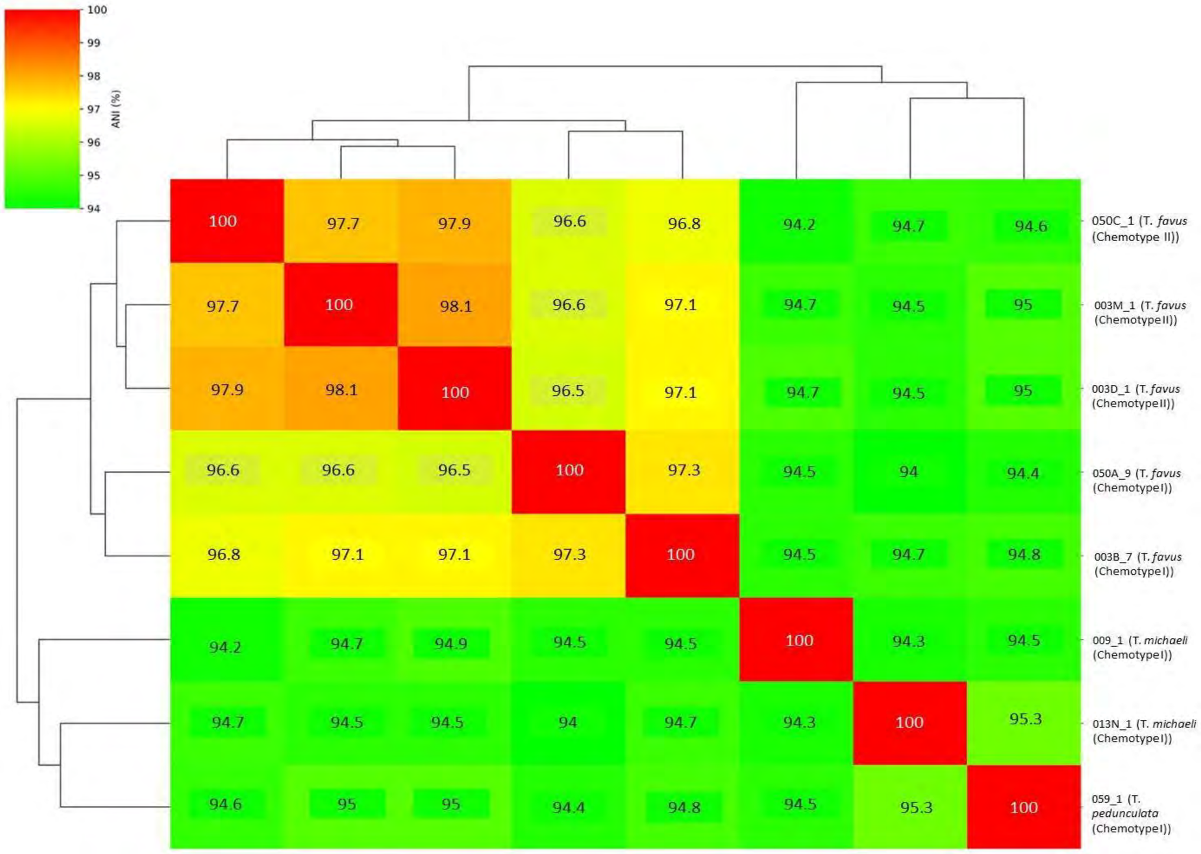

Figure 3.6: Comparative pairwise ANI values of *Tsitsikamma* sponge-associated spirochete genomes. The relatedness of *Tsitsikamma*-associated spirochetes appears to correlate with their reported chemotypes. A colour scale is used to represent the pairwise ANI values. The host species names and corresponding chemotypes of the different genomes are indicated.

Further analysis of the phylogeny of the *Tsitsikamma*-associated spirochete MAGs relative to other host-associated (N=14) and free-living (N=25) spirochetes within the Spirochaetacea*e* family was inferred using marker gene alignment via Phylophlan3 and RaxML (Figure 3.7). Notably, the phylogeny of the *Tsitsikamma*-associated spirochetes did not follow that of the phylogeny inferred from the 16S rRNA genes. The *Tsitsikamma*-associated spirochete genomes did not cluster with the other host-associated (sponges and termites) spirochetes but formed a distant clade with its closest relative *S. pacifica*, *Spirochaeta lutea, Spirochaeta* *africana* and other free-living (non-host-associated) spirochetes isolated from hypersaline microbial mats, seawater, and volcanic mud. Furthermore, the spirochetes associated with other sponges form a separate distinct clade from the termite-associated spirochetes that are more related to non-host associated spirochetes isolated from wetland sediments.

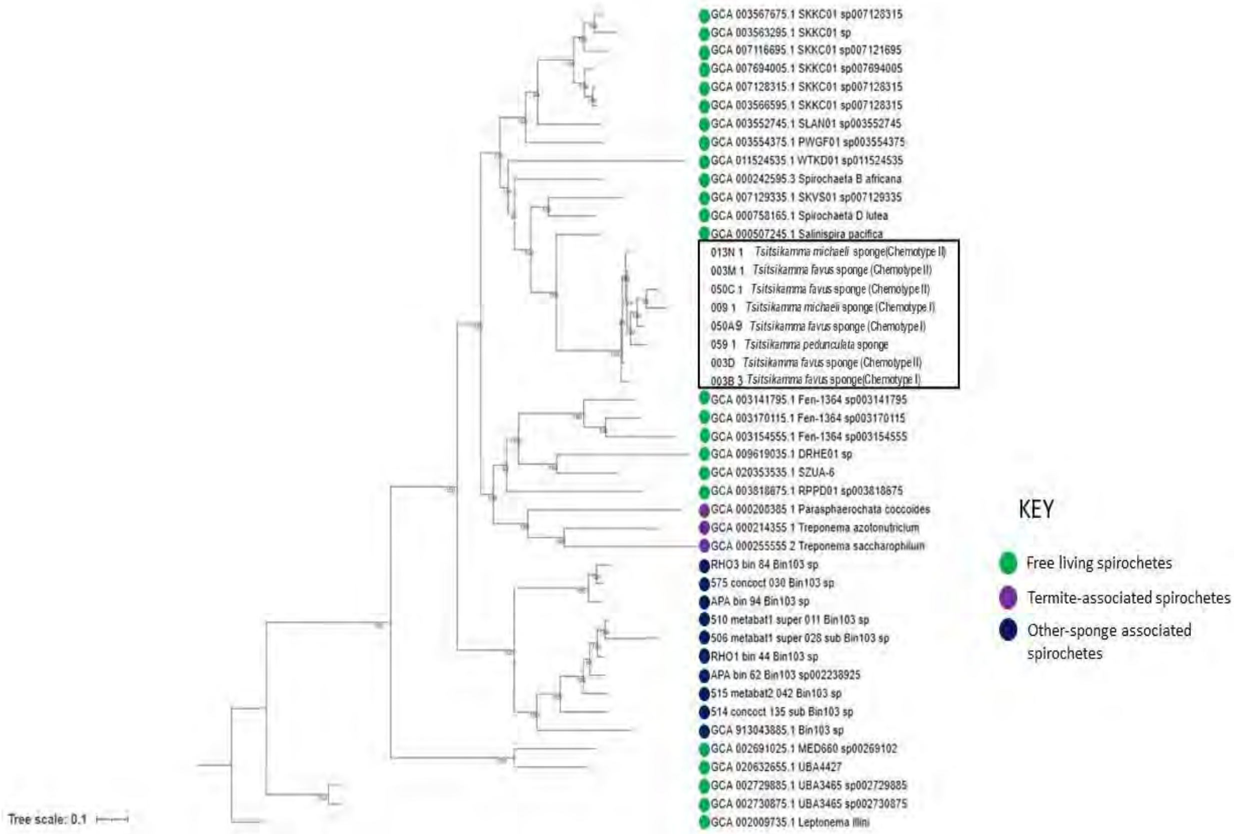

Figure 3.7: Phylogeny of sponge-associated spirochetes inferred using PhyloPhlan3 and RaxML. *Tsitsikamma*-associated spirochetes are highlighted using a black box with their respective hosts and chemotypes (where applicable). A colour key is used to represent free-living, other sponge-associated, and free-living reference spirochetes. Reference spirochete genomes are labelled as their accession number followed by their scientific name.

3.6 Estimated evolutionary divergence of sponge-associated spirochetes

The divergence pattern of sponge-associated spirochetes and their closest relatives was estimated using their rates of synonymous substitution (dS) calculated through alignment of amino acid and nucleotide coding sequences of 12 orthologous genes common to all 20 spirochete genomes (eight *Tsitsikamma* sponge-derived MAGs, ten sponge-associated spirochete reference genomes, and two seawater-derived spirochete reference genomes).

Visualization of the estimated divergence (Figure 3.8) revealed that the two groups began diverging from a common ancestor approximately 0.7 x rate[1] million years ago (mya), with the other sponge-associated spirochetes having diverged from their closest relative, a planktonic species (GCA_913043385.1), earlier than the *Tsitsikamma*-associated spirochetes diverged from *S. pacifica*. The *Tsitsikamma*-associated spirochetes are, thus, evolutionarily divergent from the other sponge-associated spirochetes as well as their closest relative *S. pacifica*. In addition, if we consider the point where the *Tsitsikamma*-associated spirochetes diverged from *S. pacifica* as the point in time when the closest ancestor of the *Tsitsikamma*-associated spirochetes began association with their latrunculid sponge hosts, then the *Tsitsikamma*-associated spirochetes only recently began diverging from one another as they began adapting to their host environments. Furthermore, the clades formed by the *Tsitsikamma*-associated species correlate with their respective hosts, however, the individual branches do not appear to follow the topology of the inferred phylogeny (Figure 3.7) nor chemotypes of the sponges (Figure 3.6).

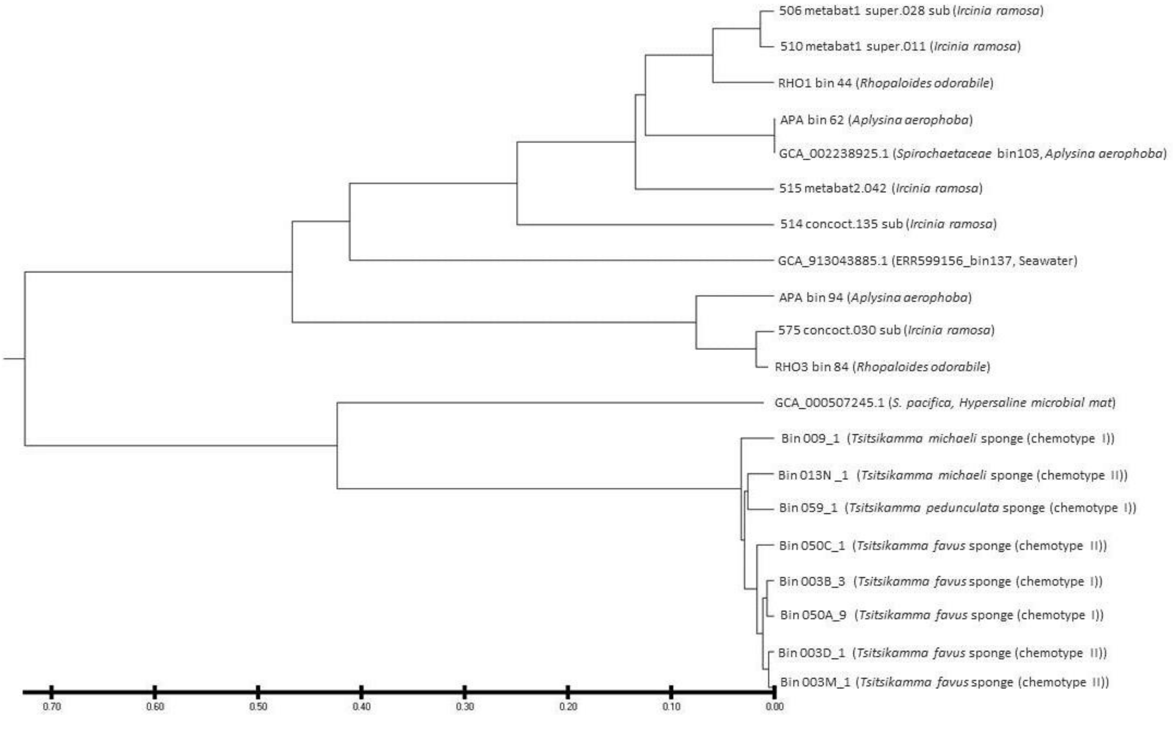

Figure 3.8: The estimated divergence pattern of sponge-associated spirochetes using pairwise synonymous substitution rates (dS) (values indicated on the x-axis) of sponge-

---

[1] Several rates of evolution have been provided on a limited number of bacteria. As all rates will be equally relative given the constant of the synonymous substitution rate, no single rate was applied as it is currently unknown which rate would be most appropriate for these particular bacteria.

associated spirochete genomes, based on the alignment of 12 orthologous genes. PAL2NAL and CodeML from the PAML package were used to calculate pairwise dS values and the resultant matrix was visualized in MEGA11 using the UPGMA method.

3.7 Comparative analysis of the functional capabilities of sponge-associated spirochetes

The functional capabilities of all 45 spirochete genomes (8 MAGs from this study, 37 reference genomes and MAGs) within the Spirochaetacea*e* family was predicted by assigning KEGG Orthologs (KO) annotations using KofamScan. KO counts were counted per genome and assigned to associated pathways within the KEGG database (Kanehisa and Goto, 2000) (Appendix A – Table A9). A non-metric multidimensional scaling (NMDS) analysis of the annotations per genome was performed to determine whether the functional gene repertoire differed between host-associated and free-living spirochetes (Figure 3.9). The *Tsitsikamma*-associated spirochetes appeared to be distinct from the other sponge-associated spirochetes and clustered towards the centre of the three groups representing free-living, other sponge-associated and other host associated, suggesting shared similarities in the functional potential between these groups. A pairwise ANOSIM of the data (Appendix A – Table A9) showed that the functional gene repertoires in the *Tsitsikamma*-associated spirochetes and other sponge associated spirochetes were significantly different ($p<0.05$) from each other as well as those from other hosts and those that were free-living. However, when comparing the R-values associated with each group, the *Tsitsikamma*-associated spirochetes may display an overlap in functional potential of free-living spirochetes ($R=0.2941$), unlike those associated from seawater ($R=0.6967$).

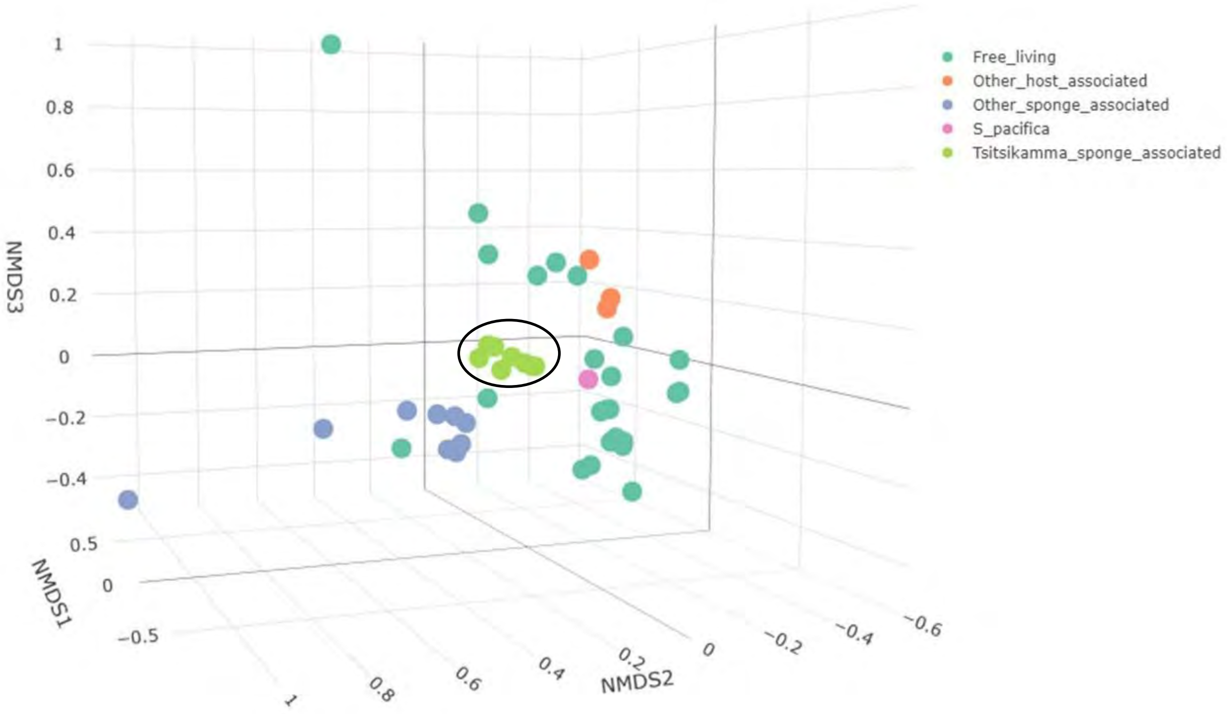

Figure 3.9: The functional gene potentials of sponge-associated and free-living spirochetes shows distinct gene profiles in *Tsitsikamma*-associated spirochetes. Non-metric multidimensional scaling plot of KEGG-annotated gene profiles in *Tsitsikamma* sponge-associated, other host-associated and free-living spirochetes. The *Tsitsikamma*-associated spirochetes (circled) appear to be at the interface of free-living and sponge-associated spirochetes.

A comparative analysis of the presence or absence of metabolic genes in both groups of sponge-associated-spirochetes and their closest relatives, was performed to determine whether there were genes common to all sponge-associated spirochetes (Figure 3.10). This was conducted by selecting genes that were present in *S. pacifica* (denoted by gene counts above one) and disregarding those that were absent from *S. pacifica* (gene counts of zero) (Appendix A – Table A8). The selected genes were then used to assess whether any genes present in *S. pacifica* were absent from both groups of sponge-associated spirochetes. No genes were shown to be absent in all sponge-associated spirochetes relative to *S. pacifica* (Appendix A – Table A8). However, to be exhaustive, genes correlating to bacterial chemotaxis, peptidoglycan biosynthesis, biofilm production, flagellar assembly, and quorum sensing pathways were examined as these are often associated with symbiotic adaptation (Chaban *et al.,* 2015; Dula *et al.,* 2012; Colin *et al.*, 2021; Preda and Săndulescu, 2019; Garde *et al.,* 2021). Inspection of the three high quality MAGs from *Tsitsikamma*-associated spirochetes (Bins 003M_1, 003D_1 and 013N_1) revealed that all the genes involved in flagellar assembly (i.e., the *flg, flh* and *fli* gene clusters) were conserved across the *Tsitsikamma*-associated spirochete genomes, while majority of these genes were absent in all spirochete genomes associated with other sponges. Conversely, the majority of the genes involved in quorum sensing were absent in the *Tsitsikamma*-associated spirochete MAGs but were present in the other sponge-associated spirochetes. Furthermore, inspection of the counts of quorum sensing genes across both groups showed that the *ddpD* and *ddpF* genes, oligopeptide permeases that form part of the ATP-binding transport cassette and transport peptides and sugars across the cell membrane (Tanaka *et al.,* 2018), were conserved across both groups of sponge-associated spirochetes, however, the abundance of these genes was considerably higher in the other sponge associated spirochete genomes (average of 102 gene counts for *ddpD* and 121 for *ddpF*) compared to the genomes of the *Tsitsikamma*-associated spirochetes (average of 11 gene counts for *ddpD* and 11 gene counts for *ddpF*) (Appendix A – A8). Interestingly, the representative MAGs of Sp02-3 and Sp02-15 encoded several genes for ABC membrane transporters, DNA metabolism, fatty acid biosynthesis and degradation, however, no genes were recovered that were associated with tryptophan biosynthesis, an amino acid from which pyrroloiminoquinone-containing natural products are derived (Figueroa *et al.,* 2024).

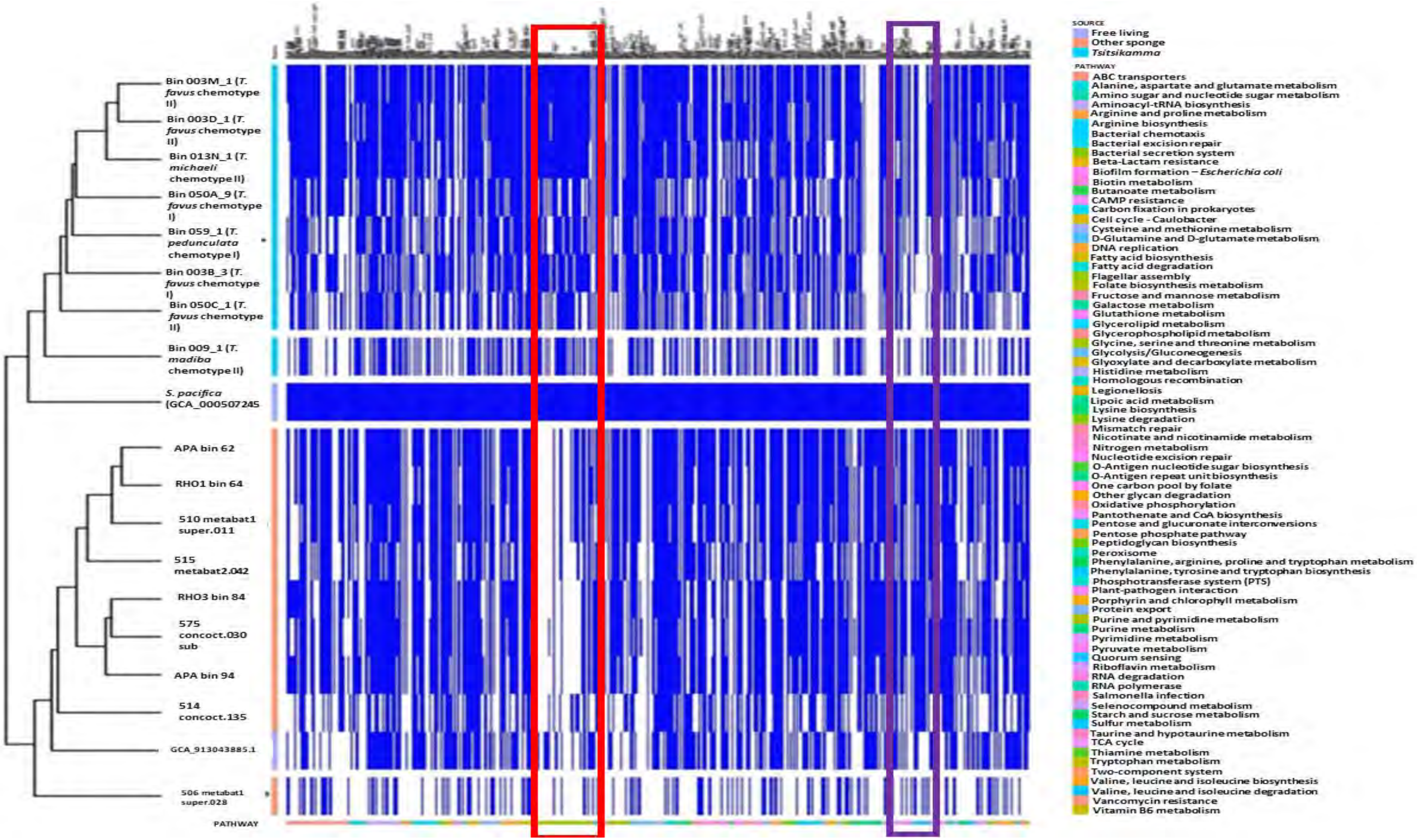

Figure 3.10: Presence/absence heatmap of genes lost from sponge-associated spirochetes and their closest free-living relatives. Only genes present in *S. pacifica* were used as a reference to determine which genes and corresponding pathways were missing from both *Tsitsikamma* sponge-

associated and other sponge associated spirochetes. Gene presence is indicated with blue boxes. A colour key illustrates the different pathways present as well as the sources from which the genes were obtained (free-living, *Tsitsikamma* sponge-associated or other sponge-associated). Genes corresponding to flagellar assembly (indicated with red box) and quorum sensing (indicated with purple box) pathways are shown.

3.8 Evidence of genome reduction in *Tsitsikamma*-associated spirochete MAGs

Further assessment of the sponge-associated spirochete MAGs was conducted to determine whether there was evidence of genome reduction, a hallmark characteristic of symbiosis (McCutcheon and Mora, 2012). The genome size, number of open reading frames (ORFs), GC content (%), coding density and the percentage of pseudogenes (number if pseudogenes per total number of ORFs) in all the sponge-associated spirochete genomes were assessed (Table 3.2) (A full list of the genome reduction statistics is available in Appendix A – Table A10). Symbionts undergoing genome reduction tend to have smaller genomes, with lower GC contents, and lower coding densities due to the lower number of functional genes and the presence of numerous pseudogenes (McCutcheon and Moran, 2012). A comparison of the two groups showed that the *Tsitsikamma*-associated spirochete genomes could potentially be undergoing genome reduction due to the smaller genome sizes, lower GC content, and smaller number of ORFs compared to the other sponge-associated spirochetes. Several medium quality genomes such as 050C_1, 009_1, 059_1, 514_concoct_135, 515_metabat2_042 and GCA913043885.1 appear to have a higher number of pseudogenes compared to the high quality genomes 003D_1, 003M_1, 013N_1, 510_metabat1_super_011, 575_concoct_135, APA bins 62 and 94, GCA002238925.1, RHO1_bin_44 and RHO3_bin_84. As pseudogenes are identified by truncation of or disruptions to open reading frames within a genome (Cheetham *et al*., 2020), medium quality genomes often exhibit a higher number of pseudogenes than higher quality genomes within the same size range due to fewer complete contigs (Cohen and Veksler-Lublinski, 2023). Nonetheless, all eight Tsitsikamma-associated spirochete genomes have likely begun undergoing genome reduction as evidenced by all eight possessing coding densities lower than the average 80% for bacterial genomes (McCutcheon and Moran, 2012), higher pseudogene counts, and lower GC content as observed in other reduced bacterial genomes.

Table 3.2: Quality assessment of *Tsitsikamma*-associated spirochete genomes for evidence of genome reduction

Genome	Size (Mbp)	Completeness (%)	GC Content (%)	Total ORFs	% Pseudogenes	Initial coding density	Coding density without pseudogenes
050A_9	0.07	3.47	6.00	2855	29.7	85.65	62.56
050C_1	1.79	86.92	50.1	2169	45.0	85.26	50.27
003B_3	1.88	88.92	50.2	2154	32.45	86.49	61.5
***003D_1**	2.49	98.93	49.63	2259	15.89	91.01	77.57
***003M_1**	2.55	98.04	49.37	2297	13.49	90.9	80.17
***013N_1**	2.33	93.6	45.05	2101	11.14	91.17	81
009_1	1.47	63.21	51.73	1898	40.3	84.39	53.87
059_1	2.04	85.25	50.29	2283	35.09	84.55	58.15
506 metabat1 super.028	2.75	55.84	66.28	2917	22.59	74.98	60.28
510 metabat1 super.011	5.23	94.53	67.53	4878	12.11	85.78	78.45
514 concoct.135	4.19	85.66	66.27	4208	23.36	71.39	57.63
515 metabat2.042	4.49	89.87	67.36	4355	16.09	83.96	73.23
***575 concoct.030**	4.48	95.33	68.58	5059	7.84	91.58	86.62
***APA bin 62**	4.84	98.13	59.5	4272	7.61	89.74	85.08
***APA bin 94**	6.27	96.93	66.35	5931	14.26	85.54	77.16
GCA_002238925.1	4.1	94.80	59.5	4327	11.05	89.06	81.86
RHO1 bin 44	5.73	97.93	67.40	5118	10.28	86.92	81.23
RHO3 bin 84	5.99	92.53	68.28	5481	9.62	90.34	84.7

*Represents high quality genomes according to MIMAG standards

3.9 Analysis of the biosynthetic potential of *Tsitsikamma*-associated spirochetes

A total of 504 biosynthetic gene clusters (BGCs) were identified across 18 sponge-associated spirochete MAGs using antiSMASH. Of these, 20 were predicted to encode a terpene BGC. Six of the eight *Tsitsikamma*-associated spirochetes were each predicted to include only a single BGC, all of which were predicted to encode a terpene. The remaining MAGs from samples TIC2016-050A and TIC2018-003B had no detectable BGCs, however, manual inspection of the scaffolds revealed two fragments of a terpene BGC in the unclustered regions of their respective metagenomes. Assessment of all 20 terpene BGCs from the sponge-associated spirochetes revealed that the BGCs identified in the other sponge spirochetes did not appear to share any homology with the BGCs identified in the *Tsitsikamma*-associated spirochetes (Figure 3.11). The terpene BGCs identified in the *Tsitsikamma*-associated spirochete genomes appear to be homologous with one another, bearing high similarity in organization of the genes as well as the amino acid sequences. A similar terpene producing BGC was identified in the genome of *S. pacifica*, the closest relative of the *Tsitsikamma*-associated spirochetes, however, the amino acid sequence similarity of shared biosynthetic genes only ranged between 37 – 50%. This indicates low similarity between the shared genes, thus suggesting that the terpene BGC in the *Tsitsikamma* spirochetes may be unique to these symbionts. Despite the low sequence and organizational similarities between the *Tsitsikamma*-associated and *S. pacifica* spirochete BGCs, it is likely that should the terpene secondary metabolite be produced by the latrunculid-associated spirochetes, it is likely to protect the symbiont or host from oxidative stress, as hypothesized for the *S. pacifica* bacterium (Hania *et al*., 2015).

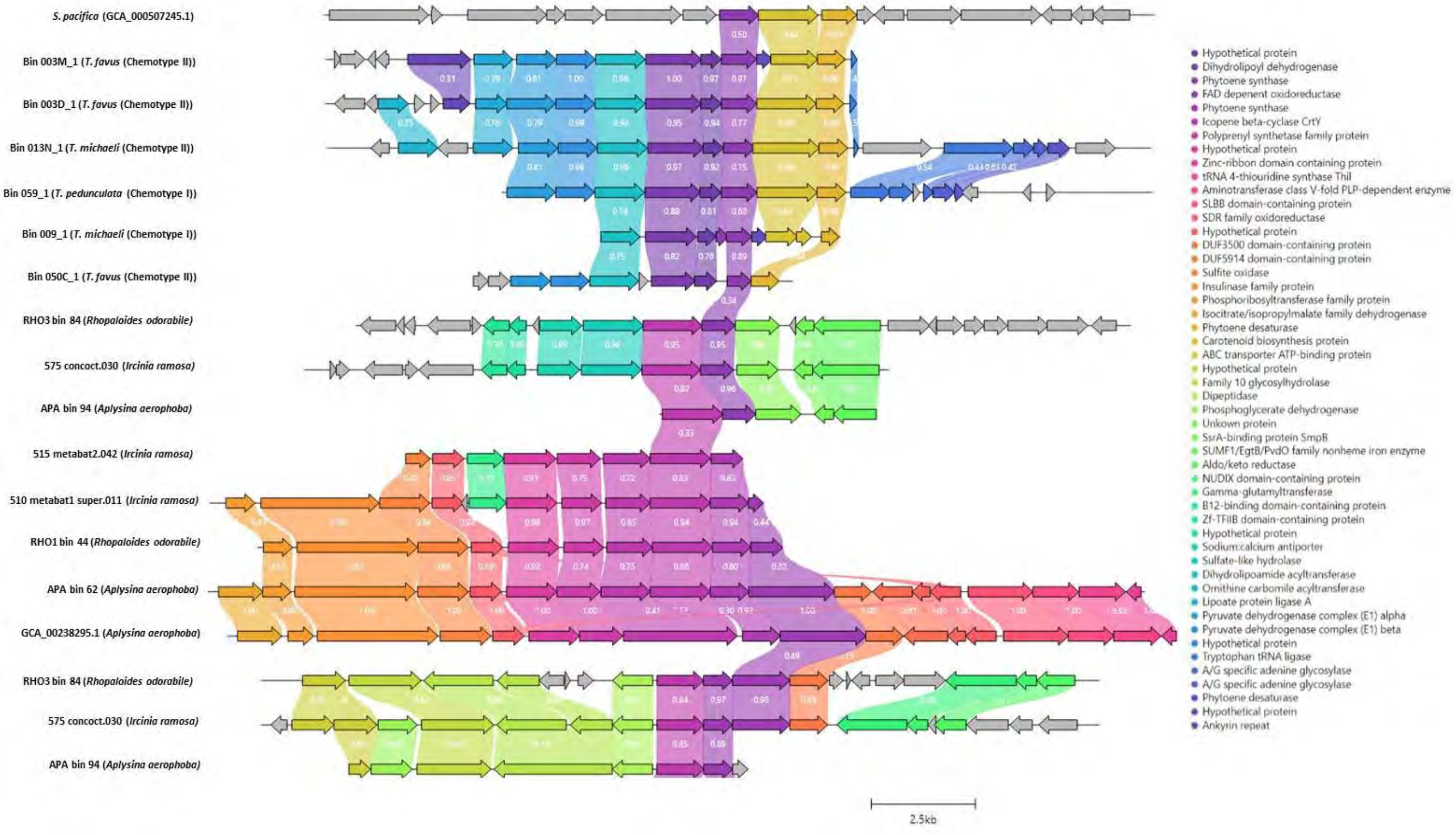

Figure 3.11: Pairwise comparison of amino acid sequence identity of terpene biosynthetic gene clusters from sponge-associated spirochetes. The pairwise similarity between genes is indicated between genes and the genes are colour coded according to their predicted functions. A scale bar is shown to indicate the gene sizes. Grey arrows indicate genes which do not show any pairwise similarity with other genes within the BGCs.

Chapter 4: Discussion and Conclusions

4.1 General discussion

Pyrroloiminoquinones are a major source of pharmaceutical potential due to the cytotoxicity, anti-microbial, anti-viral, anti-fungal and anti-inflammatory activities they possess (Antunes *et al*., 2005; Kalinski *et al*., 2019; Kalinski *et al.,* 2021; Kalinski *et al.,* 2022). The most prolific producers of these compounds are marine sponges of the Latrunculiidae family, a diverse group of sponges endemic to the Algoa Bay region of South Africa (Davies-Coleman *et al.,* 2023). Despite little progress being made to elucidate the biosynthetic origins of these compounds, sufficient evidence has been provided to support their biosynthetic origins as microbial (Jordan and Moore, 2016; Kalinski *et al.,* 2022). The microbial communities associated with sponges within this family are highly conserved (Walmsley *et al.,* 2012b; Matcher *et al*., 2017) and have been hypothesized to play a role in the biosynthesis of these compounds in *T. favus* and *Cyclacanthia* sponges (Matcher *et al.,* 2017).

The overall aim of the research presented in this study was thus to investigate the role of the dominant spirochete symbiont present in the highly conserved *Tsitsikamma favus* sponge microbial community and whether this was linked to the production of pyrroloiminoquinones in these sponges. First, the latrunculid sponge-associated bacterial communities were assessed to confirm the conservation of the dominant spirochete communities through the curation of 16S rRNA amplicon sequencing libraries. This was followed by the use of metagenomic analyses to elucidate the functional and biosynthetic potential of the conserved spirochete symbiont with respect to the possible biosynthesis of pyrroloiminoquinones.

Conservation and characterization of the dominant spirochete symbiont of latrunculid sponges

In this study 155 sponges were collected from Algoa Bay and sequenced to assess their associated bacterial communities. Of these, 79 sponges were representatives of the Latrunculiidae family and were confirmed to possess bacterial communities that were highly conserved, and numerically dominated by populations of Gammaproteobacteria and Spirochete symbionts that were representative species of the uncultured Tethybacterales symbiont clone Sp02-1 (Matcher *et al.,* 2017; Waterworth *et al.,* 2021) and the uncultured Spirochete clones Sp02-3 and Sp02-15 originally sequenced from a *T. favus* specimen TIC2009-002 (Walmsley *et al*., 2012a; Matcher *et al*., 2017; Waterworth, 2018). These representative species were otherwise absent from the communities associated with the remaining 76 sympatric non-latrunculid sponge specimens and 8 seawater samples. All three symbiont species were detected across the different *Tsitsikamma* sponge species representatives as well as in sponges representing the *Cyclacanthia bellae* and *Latrunculia apicalis* species, however, the microbial communities within the *Tsitsikamma* sponges versus those in *C. bellae* and *L. apicalis* sponges differed significantly ($p<0.05$). Both Sp02-3 and Sp02-15 strains were detected in all *Tsitsikamma* sponges and their respective OTUs abundances were exclusively correlated with the chemotype of their host sponge. Phylogenetic analysis of these spirochete OTUs revealed that both Sp02-3 and Sp02-15 were distinct from spirochetes associated with other marine invertebrates and the distantly related marine sponge *Clathrina clathrus* (Neulinger *et al*., 2010; van der Water *et al.,* 2016; Loudon *et al.,* 2023). Evans Peak, the reef from which the chemotype I and chemotype II *T. favus* and *T. michaeli* sponges were collected, is located inshore, in shallow water within the Algoa Bay and is under constant influence by the inflow of nutrient-rich freshwater from the river estuaries (Waterworth *et al*., 2017). It is tempting to speculate that the two chemotypes observed in *T. favus* and *T. michaeli* sponges could be the result of changes in the chemical composition of the freshwater inflows into Evans Peak, however, all *T. favus* sponges collected in 2016 representing both chemotypes were collected in close proximity to one another along the same reef, thus, rendering this an unlikely theory.

Phylogeny and estimated evolutionary divergence of spirochete symbiont

Metagenomic sequencing and binning of eight *Tsitsikamma* sponge specimens (spanning three species) resulted in a single spirochete MAG from each sponge sample. The topology of the phylogeny inferred from the 16S rRNA gene sequences from these MAGs correlated with the chemotypes of the respective hosts. However, this pattern was not repeated using conserved gene markers from the whole genome sequences. The phylogeny of the *Tsitsikamma* sponge spirochete genomes was inferred relative to spirochetes of the same family (Spirochaetaceae*),* where genomes were categorized as either free-living (non-host-associated) or associated with sponge or termite hosts. Spirochete species from wetland sediments, river estuaries and volcanic mats, which were selected as free-living spirochetes in the analyses, have been shown to thrive in diverse aquatic habitats as either aerobes or anaerobes (Isenbarger *et al*., 2008; Harris *et al*., 2013; Schneider *et al*., 2013). Termite-associated spirochetes such as those found in the hindguts of termite species *Parasphaerocata coccoides, Treponema azonutricium* and *Treponema saccharophilum,* are beneficial symbionts that contribute to the overall health and well-being of their hosts through the metabolism, fixation and general cycling of carbon and nitrogen (Breznak, 2002). Relative to these free-living and host-associated spirochetes, the MAGs for Sp02-3 and Sp02-15 were shown to be more closely related to free-living spirochetes isolated from hypersaline microbial mats, seawater, and volcanic mud, the closest relative being the anaerobic, aerotolerant bacterium *S. pacifica* (Hania *et al*., 2015). It is, therefore, highly likely that the same anaerobic and aerotolerant capabilities may be shared with the latrunculid spirochete MAGs.

A comparative analysis of the estimated divergences of the *Tsitsikamma*-associated spirochete genomes and the genomes of spirochetes associated with other sponges indicated that the two groups of sponge-associated spirochetes diverged from a common ancestor considerably early in their evolutionary history, approximately 0.7*rate million years ago, and that the other sponge-associated spirochetes diverged from their common ancestor earlier than the *Tsitsikamma*-associated spirochetes did. The *Tsitsikamma*-associated spirochetes only recently began diverging from their closest relative with *S. pacifica* at approximately 0.4*rate million years ago, at which point they were possibly acquired by their sponge hosts. Furthermore, if this is considered the approximate point from which the *Tsitsikamma*-associated spirochetes began diverging from the closest relative with *S. pacifica*, these spirochetes have only recently begun diverging from one another as they may be adapting to their sponge hosts. Adaptations leading to small-scale divergences between microbial populations of the same species are often the result of changes in the local environment (Rödin-Mörch *et al.,* 2020; Tuso *et al.,* 2020). This local environment not only describes the area from which these sponge specimens were collected but also includes the host environment, thus, highlighting the effect changes in either environment may have on the associated microbial communities. Comparison of the divergence estimates of the *Tsitsikamma*-associated spirochetes with that of the dominant Tethybacterales symbiont in *T. favus* sponges *Ca*. Ukwabelana africanus showed that the

Tethybacterales symbionts represent an older group of sponge symbionts, as they were likely acquired by their sponge hosts earlier at 0.6*rate million years ago (Waterworth *et al.*, 2021). Moreover, considering that these Tethybacterales symbionts are more generalist symbionts that are ubiquitous or broad host associated it is no surprise that they appear to have evolved and been acquired earlier than their spirochete counterparts, that are likely specialized symbionts as they appear to follow a narrow host range (Taylor *et al.*, 2003).

Functional potential of the spirochete symbiont and evidence of potential genome reduction

Characterisation of the dominant spirochete genome in *T. favus* sponges by Waterworth (2018) identified the presence of multiple genes involved in maintaining the symbiotic relationship between the spirochete and the sponge host such as ABC membrane transporters, lipopolysaccharides and mechanisms used to evade host predation systems such as efflux pumps (Wyrick, 2000; Garderes *et al.*, 2015; Webster and Thomas, 2016; Lorenzo *et al.*, 2017). The presence of these genes suggests the spirochete symbiont may utilise these genes to survive the toxicity present within the host sponge environment. Genome reduction is a phenomenon observed within bacterial symbionts who undergo changes to their genome as a method of adapting to the functional role they perform within their host organisms (Sloan and Morran, 2012). Similarly, genome streamlining describes the process by which bacterial symbionts undergo a reduction in genome size and the loss of non-essential genes, allowing for a reduction in energy expenditure (Sengupta *et al.*, 2024). The genome of *S. pacifica* was reported to have a GC content of 51%, 3500 open reading frames (ORFs) and 53 pseudogenes (Hania *et al.*, 2015). Consequently, the *Tsitsikamma*-associated spirochete MAGs 003D_1, 003M_1 and 013N_1 (high-quality MAGs) appear to be undergoing genome reduction, indicated by their low coding density values, high number of pseudogenes (relative to the number of predicted functional ORFs) and low GC content. This genome reduction is likely still in progress, as observed by the recent estimated divergence of these genomes from *S. pacifica* and further evolution of these *Tsitsikamma* species genomes from each other. The conservation of all genes encoding the biosynthesis of flagella in *Tsitsikamma* spirochete genomes compared to their absence from the spirochete genomes associated with other sponges confirms the likelihood of the *Tsitsikamma*-associated spirochetes still being in the process of undergoing genome reduction, as the production of flagella is most likely to ensure successful colonisation of the host environment (Otteman and Miller, 1997; Josenhans and Suerbaum, 2002; Sharndama and Mba, 2022). It is therefore possible that latrunculid-associated spirochetes may be extracellular symbionts that may possess a free-living stage, suggesting that they require an active mode of remaining associated with the sponge and its progeny. Although it is not entirely clear why quorum sensing may have been lost from the *Tsitsikamma*-associated spirochetes, it is tempting to speculate that these symbionts may no longer be reliant on quorum sensing when colonizing their host environments as they still maintained their flagella, and would not need to alter their behaviour when colonizing their host environment due to the introduction of changes in chemical stimuli or their environment (Mukherjee and Bonnie, 2019).

Biosynthetic potential of the spirochete symbiont

The genome of *S. pacifica*, the closest known relative of latrunculid-associated spirochetes, was shown to carry a terpene producing biosynthetic gene cluster (Hania *et al.,* 2015). Terpenes are carotenoid pigments that are often produced to combat oxidative stress (Britton *et al.*, 1998) and are likely responsible for the aerotolerant capabilities that *S. pacifica* possesses (Hania *et al.*, 2015). Both the *Tsitsikamma*-associated spirochetes and other sponge-associated spirochetes were found to carry terpene BGCs, however, very little homology was present between these groups. Similarly, the terpene BGCs in the *Tsitsikamma*-associated spirochetes and the BGC present in *S. pacifica* are distinct. This suggests that the terpene BGC in *Tsitsikamma*-associated spirochetes is likely novel. Despite the low levels of homology between the BGCs in the *Tsitsikamma*-associated spirochete genomes and *S. pacifica*, it is possible that the terpene BGC in the *Tsitsikamma*-associated spirochete genomes may act as an antioxidant and prevent oxidative stress within the host, similar to that observed in *S. pacifica* (Britton *et al.*, 1998; Hania *et al.,* 2015).

The data presented in this thesis indicates that the conserved spirochete in *Tsitsikamma* sponges is a relatively recent symbiont. The conservation of a single terpene BGC in the genomes would suggest that the encoded product is important to the fitness of the spirochete. While it is unclear what the function of the terpene product, it is possible that it helps protect the symbiont from oxidative stress. The maintenance of flagella would suggest that the bacterium is capable of movement but that this movement is not directed by quorum sensing but by a different method such as chemotaxis. The functional role of the spirochete remains unclear but may involve the production of the conserved terpene BGC product. What is clear, however, is that there is no evidence to suggest that the spirochetes associated with the latrunculid sponges are responsible for the production of the pyrroloiminoquinone compounds characteristic of these sponges. Instead, the spirochete appears to be co-speciating with its host, with additional responsive adaptation to the chemotype of its respective host. It has been shown that chemotype II sponges have deformed spicules (Kalinski *et al.,* 2019) and it is therefore reasonable to assume that this chemotype is detrimental to the health of the sponge. The increased abundance of Sp02-15 is correlated with this chemotype which would suggest that the distributions of the two spirochete strains between the different host chemotypes may influence sponge health and pyrroloiminoquinone production. Furthermore, it is likely that a change in the balance of these two strains could be detrimental to the health of the sponge host. This study did not show any functional differences between the two spirochete strains; however, it is likely that the assembled MAGs were a hybrid of both Sp02-3 and Sp02-15 strains, a limitation found during both the assembly and binning processes as a result of the short-reads used to generate contigs for assembly, resulting in difficulties reconstructing and identifying structural variations between the two MAGs (Sohn and Nam, 2018). Additional deep sequencing will be required to further investigate the differences between the two strains.

4.2 Future prospects

Amplicon and shotgun metagenomic sequencing of latrunculid sponges of the *Tsitsikamma* genus has revealed a correlation between the dominant spirochete strains, Sp02-3 and Sp02-15, and the distinct chemotypes in *T. favus* and *T. michaeli* sponges collected from Evans Peak, Algoa Bay. However, no genes predated to be directly involved in the production of the pyrroloiminoquinones associated with the different chemotypes, were identified within the *Tsitsikamma*-associated spirochete genomes. This indicates the need to conduct further analysis of other microbial genomes through additional metagenomic sequencing tools.

The results of this study have laid the foundation for further research into the biosynthetic origins of pyrroloiminoquinone synthesis in latrunculid sponges. Metatranscriptomics studies involve the sequencing of the RNA transcripts expressed by the entire sponge holobiont in order to provide an overview of all the genes expressed within a community (Aguiar-Pulido *et al.,* 2016; Mukherjee and Reddy, 2020). Therefore, metatranscriptomics analysis can provide insight into which functional genes are expressed under different environmental conditions. This data would provide an accurate representation of how stress and changes to environmental conditions may shape the microbial communities within a metagenome (Mukherjee and Reddy, 2020). Analysis of the metatranscriptomes from chemotype I and II *T. favus* and *T. michaeli* sponges would allow for the identification of which genes are expressed within the different sponge chemotypes, and whether the gene expression is by the host sponge or the bacterial symbionts thus providing further insight into how pyrroloiminoquinones are produced and elucidating the biosynthetic origins of these pharmaceutically important compounds.

The comparative assessment of the free-living and host-associated spirochete genomes in this study were only limited to a relatively small number of terrestrial and marine invertebrates. By expanding the study to include metagenomes from marine invertebrates besides sponges, these comparisons could be refined to determine whether dominant spirochete populations are associated with other marine invertebrates and provide further insight as to the phylogenetic and genetic relationships between latrunculid sponges and their symbiotic spirochetes. This could also reveal when and why these groups of marine invertebrate-associated spirochetes were acquired from the environment, how they subsequently evolved from one another and what the functional potential of these associated spirochetes are within their holobionts now.

4.3 Concluding remarks

Symbiotic relationships between marine sponges and their associated bacteria have long been linked to the production of secondary metabolites. As the largest contributor of pyrroloiminoquinones, the Latrunculiidae family of marine sponges, which play host to highly conserved associated microbial communities, have been a source of great effort for identifying the role symbionts may play in the production of these compounds. This study provides new knowledge of the nature of symbiotic relationships between marine sponge hosts and their symbiotic bacteria and emphasizes the acquisition of specialized bacterial symbionts to perform highly specialized roles within the sponge holobiont.

References

Abad, M.J., Bedoya, L.M., and Bermejo, P. (2012). Marine compounds and their antimicrobial activities. *Science against Microbial Pathogens: Communicating Current Research and Technological Advances*, 1293 – 1306.

Aburjania Z., Whitt J.D., Jang S., Nadkarni D.H., Chen H., Rose J.B., Velu S.E., Jaskula-Sztul R. (2020). Synthetic Makaluvamine Analogs Decrease c-Kit Expression and Are Cytotoxic to Neuroendocrine Tumor Cells. *Molecules.,* 25: 4940. doi: 10.3390/molecules25214940.

Aguiar-Pulido, V., Huang, W., Suarez-Ulloa, V., Cickovski, T., Mattee, K. and Narasimhan, G. (2016). Metagenomics, Metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol Bioinform Online*, 12(Suppl. 1): 5 – 16. Doi: 10.4137/EBO.S36436.

Alanjary, M., Steinke, K., Ziemert, N. (2019). AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* 47: W276–W282.

Alonso, E., Alvariño, R., Leirós, M., Tabudravu, J.N., Feussner, K., Dam, M.A., Rateb, M.E., Jaspars, M., Botana, L.M. (2016). Evaluation of the antioxidant activity of the marine pyrroloiminoquinone makaluvamines. *Mar. Drugs.* 14: 197. doi: 10.3390/md14110197.

Altenhoff, A.M., Levy, J., Zarowiecki, M., Tomiczek, B., Warwick Vesztrocy, A., Dalquen D.A., Müller, S., Telford, M.J., Glover, N.M., Dylus, D., Dessimoz, C. (2019). OMA standalone: orthology inference among public and custom genomes and trascriptomes. *Genome Res*., 29: 1152 – 1163.

Al-Khodor, S., Price, C.T., Kalia, A., Abu Kwaik, Y. (2010). Functional diversity of ankyrin repeats in microbial proteins. *Trends Microbiol,* 18: 132 – 139.

Amusengeri A., Bishop Ö.T. (2019). Discorhabdin N, a South African Natural Compound, for Hsp72 and Hsc70 Allosteric Modulation: Combined Study of Molecular Modeling and Dynamic Residue Network Analysis. *Molecules,* 24: 188. doi: 10.3390/molecules24010188.

Antunes, E.M., Beukes, D.R., Kelly, M., Samaai, T., Barrows, L.R., Marshall, K.M., Sincich, C., Davies-Coleman, M.T. (2004). Cytotoxic pyrroloiminoquinones from four new species of South African latrunculid sponges. *J. Nat. Prod.,* 67: 1268 – 1276. doi: 10.1021/np034084b.

Antunes, E.M., Brent, R.C., Davies-Coleman, M.T., Samaai, T. (2005). Pyrroloiminoquinones and related metabolites from marine sponges. *Natural Products Reports*, 22(1): 62 – 72.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., Ogota, H. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics,* 36: 2251 – 2252.

Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., Zhu, Q., Bolzan, M., Cumbo F., May, U., Sanders, J.G., Zolfo, M., Kopylova, E., Pasolli, E., Knight, R., Mirarab, S., Huttenhower, C., Segata, N. (2020). Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat Commun,* 11: 2500.

Astudillo-García, C., Slaby, B.M., Waite, D.W., Bayer, K., Hentschel, U. and Taylor, M.W. (2018). Phylogeny and genomics of SAUL, an enigmatic bacterial lineage frequently associated with marine sponges. *Environmental Microbiology,* 20, 561-576. https://doi.org/10.1111/1462-2920.13965

Aubart, K.M., Heathcock, C.H. (1999). A Biomimetic Approach to the Discorhabdin Alkaloids: Total Syntheses of Discorhabdins C and E and Dethiadiscorhabdin D. *J. Org. Chem.,* 64: 16 – 22. doi: 10.1021/jo9815397.

Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews in Microbiology*. 5, 782–91.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol,* 19: 455 – 477.

Bian, C., Wang, J., Zhou, W., Wu, W. and Guo, R. (2020). Recent advances in marine alkaloids from sponges. Chem. *Biodiversity*, 17: e2000186.

Beneteau, V., Besson, T. (2001). Synthesis of novel pentacyclic pyrrolothiazolobenzoquinolinones, analogs of natural marine alkaloids. *Tetrahedron Letters*, 42: 2673 – 2676.

Bergmann, W., Feeney, R.J. (1951). Contributions to the study of marine products. XXXII. The nucleosides of sponges. *J. Organ. Chem*., 16: 981-987.

Bergmann, W., Burke, D.C. (1956). Contributions to the Study of Marine Products. XL. The Nucleosides of Sponges. IV. Spongosine. *J. Organ. Chem*., 21(1343): 226 – 228.

Bergman, O., Haber, M., Mayzel, B., Anderson, M., Shpigel, M., Hill, R. & Ilan, M. (2011a). Marine-based cultivation of Diacarnus sponges and the bacterial community composition of wild and maricultured sponges and their larvae. *Marine Biotechnology*: 1-14.

Bergman, O., Mayzel, B., Anderson, M.A., Shpigel, M., Hill, R.T. & Ilan, M. (2011b). Examination of marine-based cultivation of three Demosponges for acquiring bioactive marine natural products. *Mar. Drugs*, 9: 2201 – 2219.

Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.E., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Medema, M.H. and Weber, T. (2023). AntiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures, and visualizations. *Nucleic Acids Research,* 51(W1): W46 – W50. doi: 10.1093/nar/gkad344.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., Tringe, S.G., Ivanova, N.N., Copeland, A., Clum, A., Becraft, E.D., Malmstrom, R.R., Birren, B., Podar, M., Bork, P., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson, W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J.G., Tighe, S., Konstantinidis, K.T., Liu, W.-T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Eren, A.M., Genome Standards Consortium, (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 35: 725 – 731.

Boyd, C.D., Smith, T.J., El-Kirat-Chatel, S., Newell, P.D., Dufrêne, Y.F., O'Toole, G.A. (2014). Structural features of the *Pseudomonas fluorescens* biofilm adhesin LapA required for LapG-dependent cleavage, biofilm formation, and cell surface localization. *J Bacteriol,* 196: 2775 – 2788.

Brain, C.B., Prave, A.R., Hoffman, K.-H., Fallick, A.E., Botha, A., Herd, D.A., Sturrock, C., Young, I., Condon, D.J. and Allison, S.D. (2012). The first animals: Ca. 760 my sponge-like fossils from Namibia. *South African Journal of Science*, 108: 658 – 666.

Bray, J.R. and Curtis, J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr*., 27(4): 325 – 349. doi: 10.2307/1942268.

Breznak, J.A. (2002). Phylogenetic diversity and physiology of termite gut spirochetes. *Integrative and Comparative Biology.* 42(2): 313 – 318.

Britton, G., Liaaen-Jensen, S., Pfander, H. (1998). *Carotenoids volume 4: natural functions.* Basel: Birkhäuser.

Carney, J.R., Scheuer, P.J., Kelly-Borges, M. (1993). Makaluvamine G, a cytotoxic pigment from an Indonesian sponge *Histodermella* sp. *Tetrahedron.,* 49: 8483 – 8486. doi: 10.1016/S0040-4020(01)96256-8.

Casapullo, A., Cutignano, A., Bruno, I., Bifulco, G., Debitus, C., Gomez-Paloma, L., Riccio, R. (2001). Makaluvamine P, a New Cytotoxic Pyrroloiminoquinone from *Zyzzya* cf. *fuliginosa*. *J. Nat. Prod.,* 64: 1354 – 1356. doi: 10.1021/np010053+.

Cavanaugh, C. (2015). Microbial symbiosis: patterns of diversity in the marine environment. American Zoologist, 34(1), 78–89.

Chaban, B., Hughes, H.V. & Beeby, M. (2015). The flagellum in bacterial pathogens: For motility and a whole lot more. *Semin. Cell Dev. Biol.* 46: 91–103.

Chang, L.C., Otero-Quintero, S., Hooper, J.N.A., Bewley, C.A. (2002). Batzelline D and Isobatzelline E from the Indopacific Sponge *Zyzzya fuliginosa*. *J. Nat. Prod.* 65: 776 – 778. doi: 10.1021/np010581l.

Chaumeil, P-A., Mussig, A.J., Hugenholtz, P., Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics,* 36: 1925 – 1927.

Cheetham, S.W., Faulkner, G.J. & Dinger, M.E. (2020). Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat. Rev. Genet.,* 21, 191 – 201. doi: 10.1038/s41576-019-0196-1.

Chen, T., Xu, Y., Guo, H., Liu, Y., Hu, P., Yang, X., Li, X., Ge, S., Velu, S.E., Nadkarni, D.H. (2011). Experimental therapy of ovarian cancer with synthetic makaluvamine analog: in vitro and in vivo anticancer activity and molecular mechanisms of action. *PLoS ONE.,* 6: e20729. doi: 10.1371/journal.pone.0020729.

Chen, M.L., Becraft, E.D., Pachiadaki, M., Brown, J.M., Jarett, J.K., Gasol, J.M., Ravin, N.V., Moser, D.P., Nunoura, T., Herndl, G.J., Woyke, T., Stepanauskas, R. (2020). Hiding in Plain Sight: The Globally Distributed Bacterial Candidate Phylum PAUC34f. *Front. Microbiol.*, 11: 376. doi: 10.3389/fmicb.2020.00376

Christensen, A., Martin, G.D.A. (2017). Identification and bioactive potential of marine microorganisms from selected Florida coastal areas. *Microbiology Open:* 6(4), 00448. doi: 10.1002/mbo3.448

Chu, M-J., Li, M., Ma, H., Li, P-L., and Li, G-Q. (2022). Secondary metabolites from marine sponges of the genus Agelas: A comprehensive update insight in structural diversity and bioactivity. *RSC Adv.*, 12, 7789 – 7820. doi: 10.1039/D1RA08765G

Clooney, A.G., Fouhy, F., Sleator, R.D., O'Driscoll, A., Stanton, C., Cotter, P.D., Claesson, M.J. (2016). Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. *PLoS ONE*: 11(2), e0148028. doi: 10.1371/journal.pone.0148028

Cohen, N., Veksler-Lublinsky, I. (2023). A large-scale phylogeny-guided analysis of pseudogenes in *Pseudomonas aeruginosa* bacterium. Microbiol Spectr 11: e01704 – e01723. doi: 10.1128/spectrum.01704-23

Colin, R., Ni, B., Laganenka, L., Sourjik, V. (2021). Multiple functions of flagellar motility and chemotaxis in bacterial physiology. *FEMS Microbiol Rev*. 45(6): fuab038. doi: 10.1093/femsre/fuab038.

Copp, B.R. and Ireland, C.M. (1991). Wakayin: A Novel Cytotoxic Pyrroloiminoquinone Alkaloid from the Ascidian Clavelina Species. *J. Org. Chem.,* 56: 4596 – 4597. doi: 10.1021/jo00015a005.

Cowan, J., Shadab, M., Nadkarni, D.H., Kailash, K.C., Velu, S.E., Yusuf, N. (2019). A Novel Marine Natural Product Derived Pyrroloiminoquinone with Potent Activity against Skin Cancer Cells. *Mar. Drugs.* 17: 443. doi: 10.3390/md17080443.

Cragg, G.M., Newman, D.J. (2005). Biodiversity: A continuing source of novel drug leads. *Pure Appl. Chem*., 77: 7 – 24.

Croué, J., West, N.J., Escande, M.L., Intertaglia, L., Lebaron, P. and Suzuki, M.T. (2013). A single betaproteobacterium dominates the microbial community of the crambescidine-containing sponge Crambe. *Scientific Reports*, 3: 2583.

Dat, T.T.H., Steinhert, G., Cuc, N.T.K., Smidt, H. and Sipkema, D. (2021). Bacteria cultivated from sponges and bacteria not yet cul;tivated from sponges – A review. Sec. Aquatic Microbiology,

Daniels P.N., Lee H., Splain R.A., Ting C.P., Zhu L., Zhao X., Moore B.S., van der Donk W.A. (2022). A biosynthetic pathway to aromatic amines that uses glycyl-tRNA as nitrogen donor. *Nat. Chem.,* 14: 71 – 77. doi: 10.1038/s41557-021-00802-2.

Davis, R.A., Buchanan, M.S., Duffy, S., Avery, V.M., Charman, S.A., Charman, W.N., White, K.L., Shackleford, D.M., Edstein, M.D., Andrews, K.T. (2012). Antimalarial activity of pyrroloiminoquinones from the Australian marine sponge *Zyzzya* sp. *J. Med. Chem.,* 55: 5851 – 5858.

Dias, A.D., Urban, S. and Roessner, U. (2012). A historical overview of natural products in drug discovery. *Metabolites*, 2(2): 303 – 336.

De Bary, A. (1879). Die erscheinung der symbiose: Vortrag (Strassburg: Verlag von Karl J. Trübner).

De Voogd, N. J., Cleary, D. F. R., Polónia, A. R. M., Gomes, N. C. M. (2015) Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from thousand islands reef complex, West Java, Indonesia. *FEMS Microbiology Ecology*. 91, 1-12.

Dijiman, G.G. (2000). Evolving together: The biology of symbiosis. *Proceedings (Baylor University Medical Centre)*. 13:4, 381 – 390.

Dijiman, G.G. (2002). Evolving together: The biology of symbiosis. *Proceedings (Baylor University Medical Centre)*. 13:3, 217 – 226.

Dijoux, M.-G., Schnabel, P.C., Hallock, Y.F., Boswell, J.L., Johnson, T.R., Wilson, J.A., Ireland, C.M., van Soest, R., Boyd, M.R., Barrows, L.R. (2005). Antitumor activity and distribution of pyrroloiminoquinones in the sponge genus *Zyzzya*. *Bioorg. Med. Chem.,* 13: 6035 – 6044. doi: 10.1016/j.bmc.2005.06.019.

Dulla, G.F.J., Go, R.A., Stahl, D.A. & Davidson, S.K. (2012). V*erminephrobacter eiseniae* type IV pili and flagella are required to colonize earthworm nephridia. *ISME J.* 6: 1166–1175.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res,* 32: 1792 – 1797.

Edgar, R.C. (2022). Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun.,* 13: 6968.

Ehrlich. P.R. and Raven P.H. (1964). Butterflies and plants: a study in coevolution. *Evolution.* 18, 586–60.

El-Demerdash, A., Tammam, M.A., Atanasov, A.G., Hooper, J.N.A., Al-Mourabit, A., Kijjoa, A. (2018). Chemistry and biological activities of the marine sponges of the genera Mycale (Arenochalina), Biemna and Clathria. *Mar Drugs*, 16(6): 214.

El-Naggar, M., and Capon, R.J. (2009). Discorhabdins revisited: cytotoxic alkaloids from Southern Australian marine sponges of the genera Higginsia and Spongosorites. *J. Nat. Prod.,* 72: 460 – 464. doi: 10.1021/np8007667.

Elissawy, A.M., Dehkordi, E.S., Mehdinzhad, N., Ashour, M.L., and Pour, P.M. (2021). Cytotoxic alkaloids derived from marine sponges: a comprehensive review. *Biomolecules*, 11(2): 258.

Ezell S.J., Li H., Xu H., Zhang X., Gurpinar E., Zhang X., Rayburn E.R., Sommers C.I., Yang X., Velu S.E. (2010). Preclinical Pharmacology of BA-TPQ, a Novel Synthetic Iminoquinone Anticancer Agent. *Mar. Drugs,* 8: 2129 – 2141. doi: 10.3390/md8072129.

Faulkner, D.J. (1998). Marine natural products. *Nat. Prod. Rep.*, 15: 113 – 158. doi: 10.1039/a815113y.

Feng, H., Edwards, N., Anderson, C.M.H., Althaus, M., Duncan, R.P., Hsu, Y.-C., Luetje, C.W., Price, D.R.G., Wilson, A.C.C., Thwaites, D.T. (2019). Trading amino acids at the aphid: *Buchnera* symbiotic interface. *Proc Natl Acad Sci USA* 116: 16003 – 16011.

Fieschbach, M.A., Walsh, C.T., Clardy, J. (2008). The evolution of gene collectives: How natural selection drives chemical innovation. *Proceedings of the National Academy of Sciences of the United States of America*. 105(12), 4601-4608

Figueroa, J.R., Zhu, L., van der Donk, W. (2024). Unexpected transformations during pyrroloiminoquinone biosynthesis. *J. Am. Chem. Soc*., 146(20): 14235 – 14245.

Fiore, C.L., Jarett, J.K., Lesser, M.P. (2013). Symbiotic prokaryotic communities from different populations of the giant barrel sponge, Xestospongia muta. *Microbiology Open*, 2(6): 938 – 952. doi: 10.1002/mbo3.135

Firn, R.D. and Jones, C.G. (2003). Natural products – a simple model to explain chemical diversity. *J. Nat. Prod. Rep*., 20: 382 – 391.

Ford J., and Capon R.J. (2000). Disocrhabdin R: A new antibacterial pyrroloiminoquinone from two latrunculid marine sponges, *Latrunculia* sp. and *Negombata* sp. *J. Nat. Prod.,* 63: 1527 – 1528. doi: 10.1021/np000220q.

Fu, X., Ng, P.-L., Schmitz, F.J., Hossain, M.B., van der Helm, D., Kelly-Borges, M. (1996). Makaluvic acids A and B: Novel Alkaloids from the Marine Sponge *Zyzzya fuliginosus*. *J. Nat. Prod.,* 59: 1104 –1106. doi: 10.1021/np9604947.

Garate, L., Sureda, J., Agell, G. and Uriz, M.J. (2017). Endosymbiotic calcifying bacteria across sponge species and oceans. *Sci. Rep*., 7(43674). doi: 10.1038/srep43674

Garde, S., Chodisetti, P.K., Reddy, M. (2021). Peptidoglycan: Structure, Synthesis, and Regulation. *EcoSal Plus*. 9(2). doi: 10.1128/ecosalplus.ESP-0010-2020

Gauthier, M.A., Watson, J.R., Degnan, S.M. (2016). Draft Genomes Shed Light on the Dual Bacterial Symbiosis that Dominates the Microbiome of the Coral Reef Sponge Amphimedon queenslandica, *Front. Mar. Sci.*. doi: 10.3389/fmars.2016.00196.

Gavrilidou, A., Avcı, B., Galani, A., Schorn, M.A., Ingham, C.J., Ettema, T.J.G., Smidt, H., Sipkema, D. (2023). Candidatus Nemesobacterales is a sponge-specific clade of the candidate phylum Desulfobacterota adapted to a symbiotic lifestyle. *ISME J*., 17:1808–1818.

Genta-Jouve, G., Francezon, N., Puissant, A., Auberger, P., Vacelet, J., Pérez, T., Fontana, A., Al Mourabit, A., Thomas, O.P. (2011). Structure elucidation of the new citharoxazole from the Mediterranean deep-sea sponge *Latrunculia* (*Biannulata*) *citharistae*. *Magn. Reson. Chem.,* 49: 533 – 536. doi: 10.1002/mrc.2772.

Gerçe, B., Schwartz, T., Syldatk, C. & Hausmann, R. (2011). Differences between bacterial communities associated with the surface or tissue of mediterranean sponge species. *Microbial Ecology*: 1-14.

Gerwick, W.H. and Moore, B.S. (2012). Lessons from the past and charting the future of marine natural products drug discovery and medical biology. *Chemistry and Biology*, 19(1): 85 – 98.

Gilchrist, C.L.M., Chooi, Y-H. (2021). Clinker and clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*. 37(16): 2473 – 2475.

Gloeckner, V., Wehrl, M., Moitinho-Silva, L., Gernert, C., Schupp, P., Pawlik, J. (2014). The HMA-LMA dichotomy revisited: an electron microscopical survey of 56 sponge species. *Biology Bulletin*. 227, 78 – 88.

Goey, A.K.L., Chau, C.H., Sissung, T.M., Cook, K.M., Venzon, D.J., Castro, A., Ransom, T.R., Henrich, C.J., McKee, T.C., McMahon, J.B. (2016). Screening and biological effects of marine pyrroloiminoquinone alkaloids: Potential inhibitors of the HIF-1α/p300 interaction. *J. Nat. Prod.,* 79: 1267 – 1275. doi: 10.1021/acs.jnatprod.5b00846.

Gosselin, S., Fullmer, M.S., Feng, Y., Gogarten, J.P. (2022). Improving Phylogenies Based on Average Nucleotide Identity, Incorporating Saturation Correction and Nonparametric Bootstrap Support. *Syst Biol.,* 71: 396 – 409.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research*, W695 – W695. doi: 10.1093/nar/gkq313

Graña-Miraglia, L., Sikutova, S., Vancová, M. (2020). Spirochetes isolated from arthropods constitute a novel genus *Entomospira* genus novum within the order Spirochaetales. *Sci Rep.,* 10: 17053. doi:10.1038/s41598-020-74033-9

Grkovic T., Kaur B., Webb V.L., Copp B.R. (2006). Semi-synthetic preparation of the rare, cytotoxic, deep-sea sourced sponge metabolites discorhabdins P and U. *Bioorg. Med. Chem. Lett.,* 16: 1944 – 1946. doi: 10.1016/j.bmcl.2005.12.081.

Grkovic, T., Pearce, A.N., Munro, M.H.G., Blunt, J.W., Davies-Coleman, M.T., Copp, B.R. (2010). Isolation and characterization of diastereomers of Discorhabdins H and K and assignment of absolute configuration to Discorhabdins D, N, Q, S, T, and U. *J. Nat. Prod.,* 73: 1686 – 1693. doi: 10.1021/np100443c.

Grkovic, T., and Copp B.R. (2009). New natural products in the discorhabdin A- and B-series from New Zealand-sourced *Latrunculia* spp. sponges. *Tetrahedron,* 65: 6335 – 6340. doi: 10.1016/j.tet.2009.06.012.

Grkovic, T., Ruchirawat, S., Kittakoop, P., Grothaus, P.G., Evans, J.R., Britt, J.R., Newman, D.J., Mahidol, C., O'Keefe, B.R. (2017). A New Bispyrroloiminoquinone alkaloid from a Thai follection of *Clavelina* sp. *Asian J. Org. Chem.,* 10: 1647 – 1649. doi: 10.1002/ajoc.202100155

Gunasekera S.P., McCarthy P.J., Longley R.E., Pomponi S.A., Wright A.E., Lobkovsky E., Clardy J.J. (1999). Discorhabdin P, a new enzyme inhibitor from deep-water Caribbean sponge of the genus *Batzella*. *J. Nat. Prod.,* 62: 173 – 175. doi: 10.1021/np980293y.

Gunasekera, S.P., McCarthy, P.J., Longley, R.E., Pomponi, S.A., Wright, A.E. (1996). Secobatzellines A and B, two new enzyme inhibitors from a deep-water Caribbean Sponge of the genus *Batzella*. *J. Nat. Prod.,* 62: 1208 – 1211. doi: 10.1021/np990177a.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 29(8): 1072 – 1075. doi: 10.1093/bioinformatics/btt086.

Haeffner, B. (2003). Drugs from the deep: marine natural products as potential drug candidates. *Drug Discov. Today*, 8: 536 – 544.

Hania, B.W., Joseph, M., Schumann, P., Bunk, B., Fiebig, A., Sproer, C., Klenk, H.P., Fardeau, M.L., Spring, S. (2015). Complete genome sequence and description of *Salinispira pacifica* gen. *nov., sp.* nov., a novel spirochaete isolated form a hypersaline microbial mat. *Standards in Genomic Science*, 10(7). doi: 10.1186/1944-3277-10-7.

Harris, J.K., Caporaso, J.G., Walker, J.J., Spear, J.R., Gold, N.J., Robertson, C.E., Hugenholtz, P., Goodrich, J., McDonald, D., Knights, D., Marshall, P., Tufo, H., Knight, R., Pace, N.R., Kirk Harriss, J., Gregory Caporaso, J. (2013). Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. *ISME J.* 7: 50 – 60.

Harris, E.M., Strope, J.D., Beedie, S.L., Huang, P.A., Goey, A.K.L., Cook, K.M., Schofield, C.J., Chau, C.H., Cadelis, M.M., Copp, B.R. (2018). Preclinical evaluation of discorhabdins in antiangiogenic and antitumor models. *Mar. Drugs.,* 16: 241. doi: 10.3390/md16070241.

Helber, S.B., Hoeijmakers, D.J.J., Muhando, C.A., Rohde, S. and Schupp, P.J. (2018). Sponge chemical defenses are a possible mechanism for increasing sponge abundance on reefs in Zanzibar. *PLoS One*, 13, e0197617. https://doi.org/10.1371/journal.pone.0197617

Hentschel, U., Fieseler, L., Wehrl, M., Gernert, C., Steinert, M., Hacker, J. and Horn, M. (2003). Chapter three – Microbial diversity of marine sponges. *Progress in Molecular and Subcellular Biology*, 37. Springer, Berlin, Heidelberg.

Hentschel, U., Usher, K.M. and Taylor, M.W. (2006). Marine sponges as microbial fermenters. *FEMS Microbiology Ecology*, 55: 167 – 177. doi: 10.1111/j.1574-6941.2005.00046.x

Hooper, G.J., Davies-Coleman, M.T., Kelly-Borges, M., Coetzee, P.S. (1996). New alkaloids from a South African latrunculid sponge. *Tetrahedron Lett.* 37: 7135 – 7138. doi: 10.1016/0040-4039(96)01560-2.

Hu J., Schetz J.A., Kelly M., Peng J., Ang K.K.H., Flotow H., Yan Leong C., Bee Ng S., Buss A.D., Wilkins S.P., (2002). New antiinfective and human 5-HT2 receptor binding natural and semisynthetic compounds from the Jamaican sponge *Smenospongia aurea*. *J. Nat. Prod.,* 65: 476 – 480. doi: 10.1021/np010471e.

Hu, G-P., Yuan, J., Sun, L., She, Z.G., Wu, L.H., Lan, X.J., Zhu, X., Lin, Y.C., Chen, S.P. (2011). Statistical research on marine natural products based on data obtained between 1985 and 2008. *Mar. Drugs*, 9(4): 514 – 525.

Hu, Y., Chen, J., Hu, G., Yu, J., Zhu, X., Lin, Y., Chen, S., and Yuan, J. (2015). Statistical research on the bioactivity of new marine natural products discovered during the 28 years from 1985 to 2012. *Mar. Drugs*, 13(1): 202 – 221.

Hughes, C.C., MacMillan, J.B., Gaudencio, S.P., Jensen, P.R. and Fenical, W. (2009). The ammosamides: Structures of cell cycle modulators from a marine-derived *Streptomyces* species. *Angew. Chem. Int. Ed.,* 48: 725 – 727. doi: 10.1002/anie.200804890.

Isaacs, L.T., Kan, J., Nguyen, L., Videau, P., Anderson, M.A., Wright, T.L., Hill, R.T. (2009). Comparison of the Bacterial Communities of Wild and Captive Sponge *Clathria prolifera* from the Chesapeake Bay. *Marine Biotechnology*, 11: 758 – 770. doi: 0.1007/s10126-009-9192-3

Isenbarger, T.A., Finney, M., Ríos-Velázquez, C., Handelsman, J., Ruvkun, G. (2008). Miniprimer PCR, a new lens for viewing the microbial world. *Appl. Environ. Microbiol*. 74: 840 – 849.

Ishibashi, M., Iwasaki, T., Imai, S., Sakamoto, S., Yamaguchi, K. and Ito, A. (2001). Laboratory culture of the myxomycetes: Formation of fruiting bodies of *Didymium bahiense* and its plasmodial production of makaluvamine A. *J. Nat. Prod.,* 64: 108 – 110. doi: 10.1021/np000382m.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*., 9: 5114.

Jensen S., Fortunato, S.A.V., Hoffmann, F., Home, S., Rapp, H.T., Øvreås, L. and Torsvik, V.L. (2017). The relative abundance and transcriptional activity of marine sponge-associated microorganisms emphasizing groups involved in sulfur cycle. *Microbial Ecology,* 73: 668 – 676. https://doi.org/10.1007/s00248-016-0836-3.

Jeon, J., Na, Z., Jung, M., Lee, H., Sim, C.J., Nahm, K., Oh, K.-B., Shin, J. (2010). Discorhabdins from the Korean marine sponge *Sceptrella* sp. *J. Nat. Prod.,* 73: 258 – 262. doi: 10.1021/np9005629.

Jiménez, C. (2018). Marine natural products in medicinal chemistry. *ACS Med. Chem. Lett.,* 9(10): 959 – 961.

Johnson, J.S., Spakowicz, D.J., Hong, BY., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M., Sodergren, E., Weinstock, G.E. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun.,* 10: 5029. https://doi.org/10.1038/s41467-019-13036-1

Jordan P.A., Moore B.S. (2016). Biosynthetic Pathway Connects Cryptic Ribosomally Synthesized Posttranslationally Modified Peptide Genes with Pyrroloquinoline Alkaloids. *Cell Chem. Biol.,* 23: 1504 – 1514. doi: 10.1016/j.chembiol.2016.10.009.

Josenhans, C., and Suerbaum, S. (2002). The Role of Motility as a Virulence Factor in Bacteria. *Int. J. Med. Microbiol.* 291: 605 – 614. doi: 10.1078/1438-4221-00173.

Kalinski, J-C.J., Waterworth, S.C., Noundou, X.S., Jiwaji, M., Parker-Nance, S., Krause, R.W.M, McPhail, K.L. and Dorrington, R.A. (2019). Molecular networking reveals two distinct chemotypes in pyrroloiminoquinone-producing *Tsitsikamma favus* sponges. *Mar. Drugs*, 17(1): 60.

Kalinski, J-C.J., Waterworth, S.C., Krause, R.W.M., Parker-Nance, S. and Dorrington, R.A. (2021). Unlocking the diversity of pyrroloiminoquinones produced by Latrunculid sponge species. *Mar. Drugs,* 19(2): 68.

Kalinski, J-C.J., Polyzois, A., Waterworth, S.C., Noundou, X.S. and Dorrington, R.A. (2022). Current perspectives on pyrroloiminoquinones: distribution, biosynthesis, and drug discovery potential. *Molecules*. 27(24): 8724.

Kanehisa, M., Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res,* 28: 27 – 30.

Karimi, E., Slaby, B.M., Soares, A.R., Blom, J., Hentschel, U., and Costa, R. (2018). Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. *FEMS Microbiology Ecology,* 94: fiy074.

Kelly, M. & Samaai, T. (2002). Family Latrunculiidae Topsent, 1922. IN Hooper, J.N.A. & Van Soest, R.W.M. (Eds.) System Porifera: *A guide to the Classification of Sponges*. New York: Kluwer/Plenum Academic.

Kelly, M., Sim-Smith, C., Stone, R., Samaai, T., Reiswig, H., Austin, W. (2016). New taxa and arrangements within the family Latrunculiidae (Demospongiae, Poecilosclerida). *Zootaxa*, 4121(1): 1 - 48. doi: 10.11646/zootaxa.4121.1.1

Kennedy, J., Marchesi, J.R. & Dobson, A.D. (2007). Metagenomic approaches to exploit the biotechnological potential of the microbial consortia of marine sponges. *Appl. Microbio. Biotechn*., 75: 11 - 20.

Keyzers, R.A., Arendse, C.E., Hendricks, D.T., Samaai, T., Davies-Coleman, M.T. (2005). Makaluvic acids from the South African Latrunculid sponge *Strongylodesma aliwaliensis*. *J. Nat. Prod.*, 68: 506 – 510. doi: 10.1021/np049589w.

Khidr, L. (2009). Deep sequencing. *Sci. Transl. Med.,* 1: 8ec29. doi: 10.1126/scitranslmed.3000638.

Konstantinidis, K.T., Rosselló-Móra, R., Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J* 11: 2399 – 2406.

König, H., Dröge, S. (2010). Intestinal Spirochetes of Termites. In: Dubinsky, Z., Seckbach, J. (eds) All Flesh Is Grass. Cellular Origin, Life in Extreme Habitats and Astrobiology, vol 16. Springer, Dordrecht. doi: 10.1007/978-90-481-9316-5_3.

Kudryavtsev, D.S., Spirova, E.N., Shelukhina, I.V., Son, L.V., Makarova, Y.V., Utkina, N.K., Kasheverov, I.E., Tsetlin, V.I. (2018). Makaluvamine G from the Marine Sponge *Zyzzia fuliginosa* inhibits muscle nAChR by binding at the orthosteric and allosteric sites. *Mar. Drugs.,* 16: 109. doi: 10.3390/md16040109.

Kuhn, D.A. (1981). The genus *Cristispira*. In: Starr M., Stop M.P., Trueper H.G., Balows A., Schlegel H.G. (eds) The prokaryotes: a handbook on habitats, isolation, and identification of bacteria, Springer-Verlag, Berlin, pp 473–507.

Lam, C.F.C., Grkovic, T., Pearce, N.A., Copp, B.R. (2012). Investigation of the electrophilic reactivity of the cytotoxic marine alkaloid discorhabdin B. *Org. Biomol. Chem.,* 10: 3092. doi: 10.1039/c2ob07090a.

Lam, C.F.C., Cadelis, M.M., and Copp, B.R. (2017). Exploration of the influence of spiro-dienone moiety on biological activity of cytotoxic marine alkaloid discorhabdin P. *Tetrahedron,* 73: 4779 – 4785.

Lam C.F.C., Cadelis M.M., Copp B.R. (2020). Exploration of the Electrophilic Reactiovity of the Cytotoxic Marine Alkaloid Discorhabdin C and Subsequent Discovery of a New Dimeric C-1/N-13-Linked Discorhabdin Natural Product. *Mar. Drugs.,* 18: 404. doi: 10.3390/md18080404

Lang, G., Pinkert, A., Blunt, J.W., Munro M.H.G. (2005). Discorhabdin W, the first dimeric discorhabdin. *J. Nat. Prod.,* 68: 1796 – 1798. doi: 10.1021/np050333f.

Langfelder, P., Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. *J Stat Softw* 46.

Laport, M., Santos, O., Muricy, G. (2009). Marine sponges: Potential sources of new antimicrobial drugs. *Curr. Pharma. Biotechnol.*, 10: 86 – 105.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G. (2007).

ClustalW and ClustalX version 2. *Bioinformatics*, 23(21): 2947 – 2948. doi: 10.1093/bioinformatics/btm404.

Leal, M.C., Puga, J., Serôdio, J., Gomes, N.C., Calado, R. (2012). Trends in the discovery of new marine natural products from invertebrates over the last two decades—Where and what are we bioprospecting? *PLoS One*., 7: e30580.

Letunic, I. and Bork, P. (2021). Interactive Tree OF Life (iTol) v5: an online tool for phylogenetic analysis tree display and annotation. Nucleic Acids Res, 49: W293 – W296.

Levy, T., Marchand, L., Stroobant, V., Pilotte, L., Van den Eynde, B., Rodriguez, F., Delfourne, E. (2021). IDO1 and TDO inhibitory evaluation of analogues of the marine pyrroloiminoquinone alkaloids: Wakayin and Tsitsikammamines. *Bioorg. Med. Chem. Lett.,* 40: 127910. doi: 10.1016/j.bmcl.2021.127910.

Li, R. (2015). Marinopyrroles: Unique Drug Discoveries Based on Marine Natural Products. *Medicinal Research Reviews*, 36(1): 169 – 189.

Li, F., Pfeifer, C., Pérez-Victoria, I., Tasdemir, D. (2018). Targeted isolation of tsitsikammamines from the Antarctic deep-sea sponge *Latrunculia biformis* by molecular networking and anticancer activity. *Mar. Drugs.,* 16: 268. doi: 10.3390/md16080268.

Li, F., Janussen, D., Tasdemir, D. (2020). New Discorhabdin B dimers with anticancer activity from the Antarctic deep-sea Sponge *Latrunculia biformis*. *Mar. Drugs.,* 18: 107. doi: 10.3390/md18020107.

Li. F., Kelly, M., Tasdemir, D. (2021). Chemistry, Chemotaxonomy and Biological Activity of the Latrunculid Sponges (Order Poecilosclerida, Family Latrunculiidae). *Mar Drugs*, 19(1): 27. doi: 10.3390/md19010027.

Li, P., Lu, H., Zhang, Y., Zhang, X., Liu, L., Wang, M., and Liu, W. (2023). The natural products discovered in marine sponge-associated microorganisms: structures, activities, and mining strategy. *Front. Mar. Sci*., 1 – 10. doi: 10.3389/fmars.2023.1191858

Lilburn, T.G., Kim, K.S., Ostrom, N.E., Byzek, K.R., Leadbetter, J.R., Breznak, J.A. (2001). Nitrogen fixation by symbiotic and free-living spirochetes. *Science*, 292: 2495 – 2498.

Lill, R.E., Major, D.A., Blunt, J.W., Munro, M.H.G., Battershill, C.N., McLean, M.G., Baxter, R.L. (1995). Studies on the biosynthesis of discorhabdin B in the New Zealand sponge Latrunculia sp. J. *Nat. Prod*., 58: 306 – 311. doi: 10.1021/np50116a028.

Lin, S., McCauley, E.P., Lorig-Roach, N., Tenney, K., Naphen, C.N., Yang, A., Johnson, T.A., Hernandez, T., Rattan, R., Valeriote, F.A. (2017). Another look at pyrroloiminoquinone alkaloids-perspectives on their therapeutic potential from known structures and semisynthetic analogues. *Mar. Drugs.,* 15: 98. doi: 10.3390/md15040098.

Lohmann, J.S., Wagner, S., von Nussbaum, M., Pulte, A., Steglich, W., Spiteller, P. (2018). Mycenaflavin A, B, C and D: Pyrroloquinoline alkaloids from the fruiting bodies of the mushroom *Mycena haematopus*. *Chem. Eur. J.,* 24: 8609 – 8614. doi: 10.1002/chem.201800235.

Lopanik, N.B. (2014). Chemical defensive symbioses in the marine environment. *Functional Ecology*, 28: 328 – 340. doi: 10.1111/1365-2435.12160

Loudon, A.H., Park, J., Wegener Parfrey, L. (2023). Identifying the core microbiome of the sea star *Pisaster ochraceus* in the context of sea star wasting disease. FEMS Microbiology Ecology, 99(3): fiad005. doi:10.1093/femsec/fiad005.

Lu, B, Kwan, K, Levine, YA, Olofsson, PS, Yang, H, Li, J, Joshi, S, Wang, H, Andersson, U, Chavan, SS, Tracey, KJ. (2014). α7 nicotinic acetylcholine receptor signaling inhibits inflammasome activation by preventing mitochondrial DNA release. *Mol Med*., 20(1): 350 – 358. doi: 10.2119/molmed.2013.00117.

Mahon, A.R., Amsler, C.D., McClintock, J.B., Amsler, M.O., Baker, B.J. (2003). Tissue-specific palatability and chemical defenses against macropredators and pathogens in the common articulate brachiopod Liothyrella uva from the Antarct. Penins. *J. Exp. Mar. Biol. Ecol*., 290:197 – 210. doi: 10.1016/S0022-0981(03)00075-3.

Makarchenko, A.E., Utkina, N.K. (2006). UV-Stability and UV-Protective Activity of Alkaloids from the Marine Sponge *Zyzzya fuliginosa*. *Chem. Nat. Comp.,* 42: 78 – 81. doi: 10.1007/s10600-006-0040-7.

Maloof, A.C., Rose, C.V., Beach, R., Samuels, B.M., Calmet, C.C., Erwin, D.H., Poirier, G.R., Yao, N., and Simons, F.J. (2010). Possible animal-body fossils in pre-Marinoan limestones from South Australia. *Nature Geoscience*, 3: 653–659.

Manchanda, N., Portwood, J.L., Woodhouse, M.R., Seetharam, A.S., Lawrence-Dill, C.J., Andorf, C.M. and Hufford, M.B. (2020). GenomeQC: a quality assessment tool for genome assemblies and gene structure annotations. *BMC Genomics,* 21: 193. doi: 10.1186/s12864-020-6568-2

Manzano-Marín, A., and Latorre, A. (2016). Snapshots of a shrinking partner: genome reduction in *Serratia symbiotica*. *Sci Rep* 6: 32590.

Margulis, L., Sagan, D. (1997). Microcosmos: Four Billion Years of Microbial Evolution. Berkeley: University of California Press McDevitt-Irwin JM, Baum JK, Garren M, Vega Thurber RL. 2017. Responses of coral-associated bacterial communities to local and global stressors. *Frontiers in Marine Sciences*. 4:262.

Margulis, L., Hinkle, G. (2006). Large symbiotic spirochetes: *Clevelandina, Cristispira, Diplocalyx, Hollandina* and *Pillotina*. *Procaryotes*. doi:10.1007/0-387-30747-8_43

Margulis, L. (2010) Symbiogenesis. A new principle of evolution rediscovery of Boris Mikhaylovich Kozo Polyansky (1890–1957). *Paleontological Journal*. 44:1525

Martins, A., Vieira, H., Gaspar, H. and Santos, S. (2014). Marketed marine natural products in the pharmaceutical and cosmeceutical industries: tips for success. *Mar. Drugs*, 12(2): 1066 – 1101.

Marulanda-Gomez, A., Ribes, M., Franzenburg, S., Hentschel, U. (2023). Transcriptomic responses of Mediterranean sponges upon encounter with seawater or symbiont microbial consortia. *bioRxiv* (preprint). doi: 10.1101/2023.11.02.563995

Matchado, M.S., Rühlemann, M., Reitmeier, S., Kacprowski, T., Frost, F., Haller, D., Baumbach, J. and List. (2024). On the limits of 16S rRNA gene-based metagenome prediction and functional profiling. *Microbial Genomics*, 10(1). doi: 10.1099/mgen.0.001203.

Matcher, G.F., Waterworth, S.C., Walmsley, T.A., Matsatsa, T., Parker-Nance, S., Davies-Coleman, M.T., Dorrington, R.A. (2017). Keeping it in the family: Coevolution of latrunculid sponges and their dominant bacterial symbionts. *Microbiology Open*, 6(2): e00417. doi: 10.1002/mbo3.417

Mayer, A.M.S., Glaser, K.B., Cuevas, C., Jacobs, R.S., Kem, W., Little, R.D., McIntosh, J.M., Newman, D.J., Potts, B.C., Shuster, D.E. (2010). The odyssey of marine pharmaceuticals: a current pipeline perspective. *Trends in Pharmocological Sciences*, 31(6): 255 - 265.

McCauley, E.P., Smith, G.C., Crews, P. (2020). Unraveling structures containing highly conjugated pyrrolo[4,3,2-de]quinoline cores that are deficient in diagnostic proton NMR Signals. *J. Nat. Prod.,* 83: 174 – 178. doi: 10.1021/acs.jnatprod.9b01111.

McCutcheon, J.P., and Moran, N.A. (2011). Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol.,* 10: 13 – 26.

McFall-Ngai, M., Heath-Heckman, E.A., Gillette, A.A., Peyer, S.M., Harvie, E.A. (2012). The secret languages of coevolved symbioses: insights from the Euprymna scolopes-Vibrio fischeri symbiosis. *Semin Immunol*. 24(1): 3 - 8. doi: 10.1016/j.smim.2011.11.006.

Miller, I.J., Rees, E.R., Ross, J., Miller, I., Baxa, J., Lopera, J., Kerby, R.L., Rey, F.E., Kwan, J.C. (2019). Autometa: automated extraction of microbial genomes from individual shotgun metagenomes. *Nucleic Acids Res* 47: e57.

Mishra, B.B. and Tiwari, V.K. (2012). Natural products: an evolving role in future drug discovery. *Euro. J. Med. Chem*., 46(10): 4769 – 4807.

Mohamed, N.M., Enticknap, J.J., Lohr, J.E., Mcintosh, S.M. & Hill, R.T. (2008). Changes in bacterial communities of the marine sponge Mycale laxissima on transfer into aquaculture. *Applied and Environmental Microbiology*, 74: 1209 – 1222.

Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C., Wu, Y-C., McCormack, G. (2017). Predicting the HMA-LMA status in marine sponges by machine learning. *Frontiers in Microbiology*. 8, 752.

Mukherjee, S., Bassler, B.L. (2019). Bacterial quorum sensing in complex and dynamically changing environments. *Nat. Rev. Microbiol.* 17: 371 – 382. doi: 10.1038/s41579-019-0186-5

Mukherjee, A., Reddy, M.S. (2020). Metatranscriptomics: an approach for retrieving novel eukaryotic genes from polluted and related environments. *3 Biotech.,* 10: 71. doi: 10.1007/s13205-020-2057-1

Munekata, P.E.S., Pateiro, M., Conte-Junior, C.A., Domínguez, R., Nawaz, A., Walayat, N., Movilla, E., Lorenzo, J.M. (2021) Marine Alkaloids: Compounds with In Vivo Activity and Chemical Synthesis. *Mar. Drugs*, 19(7): 374.

Na, M., Ding, Y., Wang, B., Tekwani, B.L., Schinazi, R.F., Franzblau, S., Kelly, M., Stone, R., Li, X.-C., Ferreira, D. (2010). Anti-infective Discorhabdins from a deep-water Alaskan sponge of the genus *Latrunculia*. *J. Nat. Prod.,* 73: 383 – 387. doi: 10.1021/np900281r.

Nagata, H., Ochiai, K., Aotani, Y., Ando, K., Yoshida, M., Takahashi, I., Tamaoki, T. (1997). Lymphostin (LK6-A), a novel immunosuppressant from *Streptomyces* sp. KY11783: Taxonomy of the producing organism, fermentation, isolation, and biological activities. *J. Antibiot.,* 50: 537 – 542. doi: 10.7164/antibiotics.50.537.

Nakatani, S., Kiyota, M., Matsumoto, J., Ishibashi, M. (2005). Pyrroloiminoquinone pigments from Didymium iridis. *Biochem. Syst. Ecol.,* 33: 323 – 325. doi: 10.1016/j.bse.2004.06.015.

Niehaus, T.D., Elbadawi-Sidhu, M., de Crécy-Lagard, V., Fiehn, O., Hanson, A.D. (2017). Discovery of a widespread prokaryotic 5-oxoprolinase that was hiding in plain sight. *J Biol Chem,* 292: 16360 – 16367.

Neulinger, S.C., Störh, R., Thiel, V., Schmaljohann, R. and Imhoff, J.F (2010). New phylogenetic lineages of the *Spirochaetes* phylum associated with *Clathrina* species (Porifera). *J. Microbiol*., 48(4): 411 – 418. doi: 10.1007/s12275-010-0017-x.

Nijampatnam B., Nadkarni D.H., Wu H., Velu S.E. (2014). Antibacterial and Antibiofilm Activities of Makaluvamine Analogs. *Microorganisms,* 2: 128 – 139. doi: 10.3390/microorganisms2030128.

O'Brien, P.A., Andreakis, N., Tan, S., Miller, D.J., Webster, N.S., Zhang, G., Bourne, D.G. (2021). Testing cophylogeny between coral reef invertebrates and their bacterial and archaeal symbionts. *Mol Ecol*., 30: 3768 – 3782.

O'Brien, P.A., Tan, S., Frade, P.R., Robbins, S.J., Engelberts, J.P., Bell, S.C., Vanwonterghem, I., Milelr, D.J., Webster, N.S., Zhang, G. and Bourne, D.G. (2023). Validation of key sponge symbiont pathways using genome-centric metatranscriptomics. *Environ Microbiol*., 25: 3207 – 3224.

Ottemann, K.M., Miller, J.F. (1997). Roles for Motility in Bacterial-Host Interactions. *Mol Microbiol*. 24:1109 –1117. doi: 10.1046/j.1365-2958.

Parker-Nance, S., Hilliar, S., Waterworth, S., Walmsley, T., and Dorrington, R. (2019). New species in the sponge genus *Tsitsikamma* (*Poecilosclerida*, Latrunculiidae) from South Africa. *Zookeys.,* 126: 101 – 126. doi: 10.3897/zookeys.874.32268.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res,* 25: 1043 – 1055.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P-A., Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.,* 36: 996 – 1004.

Paul, V.J., Puglisi, M.P. (2004). Chemical mediation of interactions among marine organisms. *Nat. Prod. Rep*., 21:189–209. doi: 10.1039/b302334f.

Paul, V.J., Puglisi, M.P., Ritson-Williams, R. (2006). Marine chemical ecology. *Nat. Prod. Rep*., 23:153–180. doi: 10.1039/b404735b.

Pérez-Cobas, A.E., Gomez-Valero, L., Buchrieser, C. (2020). Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microb Genom*., 6(8): mgen000409. doi: 10.1099/mgen.0.000409.

Perry, N.B., Blunt, J.W., Munro, M.H.G. (1988a). Cytotoxic pigments from New Zealand sponges of the genus *Latrunculia*: Discorhabdins A, B and C. *Tetrahedron.,* 44: 1727 – 1734. doi: 10.1016/S0040-4020(01)86737-5.

Perry, N.B., Blunt, J.W., Munro, M.H.G. (1988b). Discorhabdin D, an Antitumor Alkaloid from the Sponges *Latrunculia brevis* and *Prianos* sp. *J. Org. Chem.,* 53: 4127 – 4128. doi: 10.1021/jo00252a052.

Peters, S. and Spiteller, P. (2007). Sanguinones A and B, blue pyrroloquinoline alkaloids from the fruiting bodies of the mushroom *Mycena sanguinolenta*. *J. Nat. Prod.,* 70: 1274 – 1277. doi: 10.1021/np070179s.

Peters, S., Jaeger, R.J.R., and Spiteller, P. (2008). Red pyrroloquinoline alkaloids from the mushroom *Mycena haematopus*. *Eur. J. Org. Chem.,* 2008: 319 – 323. doi: 10.1002/ejoc.200700739.

Popov, A.M., Utkina, N.K. Pyrroloquinoline alkaloids from *Zyzzya* sp. (1998). sea sponges: Isolation and antitumor activity characterization. *Pharm. Chem. J.,* 32: 298 – 300. doi: 10.1007/BF02580512.

Pourmohammadi, R., Abouei, J. and Anpalagan, A. (2021). Error analysis of the PacBio sequencing CSS reads. *The International Journal of Biostatistics*, 19(2): 439 – 453.

Price, G., and Syme, A. (2024). Genome-assessment-post-assembly. *WorkflowHub*. doi: 10.48546/WORKFLOWHUB.WORKFLOW.403.4

Preda, V.G., Săndulescu, O. (2019). Communication is the key: biofilms, quorum sensing, formation and prevention. *Discoveries (Craiova).* 7(3): e100. doi: 10.15190/d.2019.13.

Proksch P. (1994). Defensive roles for secondary metabolites from marine sponges and sponge-feeding nudibranchs. *Toxicon*., 32:639 – 655. doi: 10.1016/0041-0101(94)90334-4.

Pulte, A., Wagner, S., Kogler, H., Spiteller, P. (2016). Pelianthinarubins A and B, red pyrroloquinoline alkaloids from the fruiting bodies of the mushroom *Mycena pelianthina*. *J. Nat. Prod.,* 79: 873 – 878. doi: 10.1021/acs.jnatprod.5b00942.

Ravi, R.K., Walton, K., Khosroheidari, M. (2018). MiSeq: A Next Generation Sequencing Platform for Genomic Analysis. In: DiStefano, J. (eds) Disease Gene Identification. Methods in Molecular Biology, vol 1706. Humana Press, New York, NY. doi.: 10.1007/978-1-4939-7471-9_12

Rice, P., Longden, I., Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet,* 16: 276 – 277

Rhoads, A., Au, K.F. (2015). PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinformatic*s., 13(5): 278 – 289. doi: 10.1016/j.gpb.2015.08.002.

Robbins, S.J., Song, W., Engelberts, J.P., Glasl, B., Slaby, B.M., Boyd, J., Marangon, E., Botté, E.S., Laffy, P., Thomas, T., Webster, N.S. (2021). A genomic view of the microbiome of coral reef demosponges. *ISME J.*, 15: 1641 – 1654.

Rödin-Mörch, P., Palejowski, H., Cortazar-Chinarro, M., Kärvemo, S., Richter-Boix, A., Höglund, J., and Laurila, A. (2021). Small-scale population divergence is driven by local larval environment in a temperate amphibian. *Heredity*, 126: 279 – 292. doi: 10.1038/s41437-020-00371-z.

Rodriguez-R, L.M., Conrad, R.E., Viver, T., Feistel, D.J., Lindner, B.G., Venter, S.N., Orellana, L.H., Amann, R., Rossello-Mora, R., Konstantinidis, K.T. (2024). An ANI gap within bacterial species that advances the definitions of intra-species units. *mBio* 15: e02696-23. doi: 10.1128/mbio.02696-23.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F. (2016). VSEARCH: a versatile open-source tool for metagenomics, PeerJ, 4, e2584. doi: 10.7717/peerj.2584.

Rosenberg, E., Falkovitz, L. (2004). The Vibrio shiloi/Oculina patagonica model system of coral bleaching. *Annual Reviews on Microbiology.* 58, 143–59.

Ruocco, N., Esposito, R., Zagami, G., Bertolino, M., De Matteo, S., Sonnessa, M., Andreani, F., Crispi, S., Zupo, M. and Constantini, M. (2021). Microbial diversity in Mediterranean sponges as revealed by metataxonomic analysis. *Sci Rep* 11, 21151. doi: 10.1038/s41598-021-00713-9

Sagar, S., Kaur, M., Minneman, K.P. (2010). Antiviral lead compounds from marine sponges. *Mar. Drugs*, 8(10): 2619 – 2638.

Sahlin, K., Medvedev, P. (2021). Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis. *Nat. Commun.,* 12(2). doi: 10.1038/s41467-020-20340-8.

Samaai, T., Kelly, M.J., Ngwakum, B., Payne, R., Teske, P.R., Janson, L., Kerwath, S., Parker, D., and Gibbons, M.J. (2020). New Latrunculiidae (Demospongiae, Poecilosclerida) from the Agulhas ecoregion of temperate southern Africa. *Zootaxa*. 4896(3), 409 – 442.

Seemann T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics,* 30: 2068 – 2069.

Seemann, T. and Booth, T. (2018). Basic Rapid Ribosomal RNA Predictor (Barrnap). Available from: https://github.com/tseemann/barrnap. Accessed: 19/02/2024.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.*, 75: 7537 – 7541. doi: 10.1128/AEM.01541-09.

Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T.S., Bushuiev, R., Murray, K.J, Hoffmann, N., Lu, M., Sarvepalli, A., Zhang, Z., Fleischauer, M., Dührkop, K., Wesner, M., Hoogstra, S.J., Rudt, E., Mokshyna, O., Brungs, C., Ponomarov, K., Mutabdžija, L., Damiani, T., Pudney, C.J., Earll, M., Helmer, P.O., Fallon, T.R., Schulze, T., Rivas-Ubach, A., Bilbao, A., Richter, H., Nothias, L-F., Wang, M., Orešič, M., Weng, J-K., Böcker, S., Jeibmann, A., Hayen, H., Karst, U., Dorrestein, P.C., Petras, D., Du, X., Pluskal, T. (2023). Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol.,* 41: 447 – 449

Schmidt, E.W., Harper, M.K., Faulkner, D.J. (1995). Makaluvamines H-M and damirone C from the Pohnpeian sponge *Zyzzya fuliginosa*. *J. Nat. Prod.,* 58: 1861 – 1867. doi: 10.1021/np50126a008.

Schneider, D., Arp, G., Reimer, A., Reitner, J., Daniel, R. (2013). Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the Kiritimati Atoll, Central Pacific. *PLoS One.* 8: e66662.

Sengupta, A., Bandyopadhyay, A., Sarkar, D., Hendry, J.I., Schubert, M.G., Liu, D., Church, G.M., Maranas, C.D., Pakrasi, H.B. (2024). Genome streamlining to improve performance of a fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *mBio*, 15: e03530 – 23.

Sharpton, T.J. (2014). An introduction to the analysis of shotgun metagenomic data. *Front. Plant. Sci*., 5: 209.

Sharndama, H.C., Mba, I.E. (2022). Helicobacter Pylori: An Up-to-Date Overview on the Virulence and Pathogenesis Mechanisms. *Braz. J. Microbiol.* 53: 33 – 50. doi: 10.1007/s42770-021-00675-0

Shimoya, Y. (2022). ANIclustermap: A tool for drawing ANI clustermap between all-vs-all microbial genomes. Available from: https://github.com/moshi4/ANIclustermap. Access date: 04/03/2024.

Sitnikova, T., Michel, E., Tulupova, Y., Khanaev, I., Parfenova, V. and Prozorova, L. (2012). Spirochetes in gastropods from Lake Baikal and North American freshwaters: new multi-family, multi-habitat host records. *Symbiosis,* 56, 103–110. doi:10.1007/s13199-012-0167-1

Skropeta D. (2008). Deep-sea natural products. *Nat. Prod. Rep*., 25: 1131 – 1166. doi: 10.1039/b808743a.

Sloan, D.B. and Morran, N.A. (2012). Genome reduction and co-evolution between the primary and secondary bacterial symbionts of Psyllids. *Mol. Biol. Evol*, 29(12): 3781 – 3792.

Sohn, J. and Nam, J-W. (2018). The present and future of de novo whole-genome assembly. *Briefings in Bioinformatics*, 19(1): 23 – 40. doi: 10.1093/bib/bbw096

Somerfield, P.J., Clarke, R.K. and Gorley, R.N. (2021). A generalised analysis of similarities (ANOSIM) statistic for designs with ordered factors. *Austral Ecology*, 46(6): 901 – 910.

Souza, C.R.M., Bezerra W.P., Souto J.T. (2020) Marine alkaloids with anti-inflammatory activity: Current knowledge and future perspectives. *Mar. Drugs,* 18: 147.

Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9): 1312 – 1313. doi: 10.1093/bioinformatics/btu033.

Stierle, D.B., and Faulkner, D.J. (1991). Two New Pyrroloquinoline alkaloids from the sponge *Damiria* sp. *J. Nat. Prod.,* 54: 1131 – 1133. doi: 10.1021/np50076a038.

Suyama, M., Torrents, D., Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res,* 34: W609 –W612.

Tamura, K., Stecher, G., Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol.,* 38(7): 3022 – 3027. doi: 10.1093/molbev/msab120.

Tanaka, K.J., Song, S., Mason, K., Pinkett, H.W. (2018). Selective substrate uptake: The role of ATP-binding cassette (ABC) importers in pathogenesis. *Biochim. Biophys. Acta. Biomembr.*,1860(4): 868 – 877. doi: 10.1016/j.bbamem.2017.08.011.

Tasdemir, D., Mangalindan, G.C., Concepción, G.P., Harper, M.K., Ireland, C.M. (2001). 3,7-Dimethylguanine, a New Purine from a Philippine Sponge *Zyzzya fuliginosa*. *Chem. Pharm. Bull.,* 49: 1628 – 1630. doi: 10.1248/cpb.49.1628.

Taufa, T., Gordon, R.M.A., Ali Hashmi, M., Hira, K., Miller, J.H., Lein, M., Fromont, J., Northcote, P.T., Keyzers, R.A. (2019). Pyrroloquinoline derivatives from a Tongan specimen of the marine sponge *Strongylodesma tongaensis*. *Tetrahedron Lett.,* 60: 1825 – 1829. doi: 10.1016/j.tetlet.2019.06.014.

Taylor, M.W., Schupp, P.J., Dahllöf, I., Kjelleberg, S., Steinberg, P.D. (2003). Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. *Environmental Microbiology*, 6(2): 121 – 130. doi:10.1046/j.1462-2920.2003.00545.x.

Tempone, A.G., Pieper, P., Borborema, S.E.T., Thevenard, F., Lago, J.H.G., Croft, S.L., Anderson, E.A. (2021). Marine alkaloids as bioactive agents against protozoal neglected tropical diseases and malaria. *Nat. Prod. Rep*., 2214 – 2235.

Thomas, T.R.A., Kavlekar, D.P., LokaBharathi, P.A. (2010). Marine drugs from sponge-microbe association—A review. *Mar. Drugs*., 8: 1417 – 1468. doi: 10.3390/md8041417.

Thomas, T. Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., Olsen, J.B., Erwin, P.M., López-legentil, S., Luter, H., Chaves-Fonnegra, A., Costa, R., Schupp, P.J., Steindler, L., Erpenbeck, D., Gilbert, J., Knight, R., Costa, R., Schupp, P.J., Stendler, L., Erpenbeck, D., Gilbert, J., Knight, R., Ackermann, G., Victor Lopez, J., Taylor, M.W., Thacker, R.W., Montoya, J.M., Hentschel, U. and Webster, N.S. (2016). Diversity, structure, and convergent evolution of the global sponge microbiome. *Nature Communications*, 7: 11870. doi: 10.1038/ncomms11870

Tlili, C., Djebbi, K., Elaguech, M.A, Bahri, M., Zhou, D., Shi, B. and Wang, D. (2022). Next-Generation DNA Sequencing: Ion Torrent Sequencers Versus Nanopore Technology. In: Sawan, M. (eds) Handbook of Biochips. Springer, New York, NY. doi:10.1007/978-1-4614-3447-4_56

Topsent, E. (1922). Les megascleres polytylotes des Monaxonides et la parente des Latrunculiines. *Bulletin de I'Institut oceanographique Monaco*. 415: 1 – 8.

Toshiya, I., Moriya, O., Kuniyo, O., Toshiaki, K. (2000). Symbiotic spirochetes in the termite hindgut: phylogenetic identification of ectosymbiotic spirochetes of oxymonad protists. *FEMS Microbiology Ecology*, 34(1): 17 – 26. doi: 10.1111/j.1574-6941.2000.tb00750.x

Tusso, S., Nieuwenhuis, B.P.S., Weissensteiner, B., Immler, S., Wolf, J.B.W. (2021). Experimental evolution of adaptive divergence under varying degrees of gene flow. *Nat Ecol Evol* 5: 338 – 349. https://doi.org/10.1038/s41559-020-01363-2

Urban, S., Hickford, S.J.H., Blunt, J.W., Munro, M.H.G. (2000). Bioactive marine alkaloids. *Curr. Org. Chem.* 4: 765 – 807. doi: 10.2174/1385272003376085.

Utkina, N.K., Gerasimenko, A.V., Popov, D.Y. (2003). Transformation of tricyclic makaluvamines from the marine sponge *Zyzzya fuliginosa* into damirones. *Russ. Chem. Bull.* 52: 258 – 260. doi: 10.1023/A:1022489508349.

Utkina, N.K., Makarchenko, A.E., Denisenko, V.A., Dmitrenok, P.S. (2004). Zyzzyanone A, a novel pyrrolo[3,2-f]indole alkaloid from the Australian marine sponge *Zyzzya fuliginosa*. *Tetrahedron Lett.,* 45: 7491 – 7494. doi: 10.1016/j.tetlet.2004.08.057.

Utkina, N.K., Makarchenko, A.E., Denisenko, V.A. (2005). Zyzzyanones B-D, Dipyrroloquinones from the Marine Sponge *Zyzzya fuliginosa*. *J. Nat. Prod.,* 68: 1424 – 1427. doi: 10.1021/np050154y.

Van der Water, J.A.J.M., Melkonian, R., Junca, H., Voolstra, C.R., Reynaud, S., Allemand, D., Ferrier- Pages, C. (2016). Spirochaetes dominate the microbial community associated with the red coral Coralliumrubrum on a broad geographic scale, *Scientific Reports*, 6, e27277. doi: 10.1038/srep27277.

Van de Water, J., Melkonian, R., Junca, H., Melkonian, R., Junca, H., Voolstra, C.R., Reynaud, S., Allemand, D., Ferrier-Pagés, C. (2016). Spirochaetes dominate the microbial community associated with the red coral *Corallium rubrum* on a broad geographic scale. *Sci Rep,* 6: 27277. doi: 10.1038/srep27277

Venables, D.A., Concepción, G.P., Matsumoto, S.S., Barrows, L.R., Ireland, C.M. (1997). Makaluvamine N: A new pyrroloiminoquinone from *Zyzzya fuliginosa*. *J. Nat. Prod.,* 60: 408 – 410. doi: 10.1021/np9607262.

Wada, Y., Fujioka, H., Kita, Y. (2010). Synthesis of the Marine Pyrroloiminoquinone Alkaloids, Discorhabdins. *Mar. Drugs.,* 8: 1394 – 1416. doi: 10.3390/md8041394.

Wada, N., Yuasa, H., Kajitani, R., Gotoh, Y., Ogura, Y., Yoshimura, D., Toyoda, A., Tang, S-L., Higashimura, Y., Sweatman, H., Forsman, Z., Bronstein, O., Eyal, G., Thongtham, N., Itoh, T., Hayashi, T., Yasuda, N. (2020). A ubiquitous subcuticular bacterial symbiont of a coral predator, the crown-of-thorns starfish, in the Indo-Pacific. *Microbiome* 8:123.

Walmsley, T.A. (2012a). An investigation into the bacterial diversity associated with South African Latrunculid sponges that produces bioactive secondary metabolites. *PhD Thesis*. Rhodes University. 1 – 157.

Walmsley, T.A., Matcher, G.F., Zhang, F., Hill, R.T., Davies-Coleman, M.T., Dorrington, R.A. (2012b). Diversity of bacterial communities associated with the Indian Ocean sponge *Tsitsikamma* *favus* that contains the bioactive pyrroloiminoquinones, tsitsikammamine A and B. *Mar. Biotechnol.*, 14: 681 – 691.

Wang, G. (2006). Diversity and biotechnological potential of the sponge-associated microbial consortia. *J. Indust. Microbiol. Biotechn.,* 33: 545 – 551.

Wang, W., Rayburn, E.R., Velu, S.E., Nadkarni, D.H., Murugesan, S., Zhang, R. (2009). In vitro and in vivo anticancer activity of novel synthetic makaluvamine analogues. *Clin. Cancer Res.* 15: 3511 – 3518. doi: 10.1158/1078-0432.CCR-08-2689.

Wang W., Nijampatnam B., Velu S.E., Zhang R. (2016) Discovery and development of synthetic tricyclic pyrroloquinone (TPQ) alkaloid analogs for human cancer therapy. *Front. Chem. Sci. Eng.,* 10: 1 – 15. doi: 10.1007/s11705-016-1562-6.

Wang, Y., Zhao, Y., Bollas, A., Wang, Y., and Au, K.F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol.,* 39, 1348 –1365. doi: 10.1038/s41587-021-01108-x

Waterworth, S.C., Jiwaji, M., Kalinski, J-C.J., Parker-Nance, S., Dorrington, R.A. (2017). A place to call home: an analysis of the bacterial communities in two *Tethya rubra* Samaai and Gibbons 2005 populations in Algoa Bay, South Africa. *Mar. Drugs*, 15: 95.

Waterworth, S.C. (2018). An investigation into the biosynthetic origins of pyrroloiminoquinone production by latrunculid sponges. *PhD thesis*, Rhodes University.

Waterworth, S.C., Parker-Nance, S., Kwan, J.C. and Dorrington, R.A. (2021). Comparative genomics provides insight into the function of broad-host range sponge symbionts. *American Society for Microbiology,* 12: 5.

Webster, N.S., Negri, A. P., Munro, M.M. and Battershill, C.N. (2004). Diverse microbial communities inhabit Antarctic sponges. *Environmental Microbiology*, 6: 288 – 300.

Wier, A., Dolan, M., Grimaldi, D., Guerrero, R., Wagensberg, J., Margulis, L. (2002). Spirochete and protist symbionts of a termite (*Mastotermes electrodominicus*) in Miocene amber. *Proc Natl Acad Sci USA*, 99(3): 1410 – 1413. doi: 10.1073/pnas.022643899.

WoRMS Editorial Board (2017). World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2024-01-27. doi:10.14284/170.

Xie, H., Yang, C., Sun, Y., Igarashi, Y., Jin, T., Luo, F. (2020). PacBio long reads improve metagenomic assemblies, gene catalogs and genome binning. *Front. Genet.*, 11. doi: 10.3389/fgene.2020.516269.

Yan, P., Ritt, D.A., Zlotkowski, K., Bokesch, H.R., Reinhold, W.C., Schneekloth, J.S., Jr., Morrison, D.K., Gustafson, K.R. (2018). Macrophilones from the marine hydroid *Macrorhynchia philippina* can inhibit ERK cascade signaling. *J. Nat. Prod.,* 81: 1666 – 1672. doi: 10.1021/acs.jnatprod.8b00343.

Yang, A., Baker, B.J., Grimwade, J., Leonard, A., McClintock, J.B. (1995). Discorhabdin alkaloids from the Antarctic sponge *Latrunculia apicalis*. *J. Nat. Prod.*, *58*, 1596 – 1599. doi: 10.1021/np50124a020

Yang, Q., Farnco, C.M.M., Sorokin, S.J., Zhang, W. (2017). Development of a multilocus-based approach for sponge (phylum Porifera) identification: refinement and limitations. *Scientific Reports*, 7, 41422. doi: 10.1038/srep41422

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*, 24: 1586 –1591.

Zhang X., Xu H., Zhang X., Voruganti S., Murugesan S., Nadkarni D.H., Velu S.E., Wang M.-H., Wang W., Zhang R. (2012). Preclinical Evaluation of Anticancer Efficacy and Pharmacological Properties of FBA-TPQ, a Novel Synthetic Makaluvamine Analog. *Mar. Drugs.,* 10: 1138 – 1155. doi: 10.3390/md10051138.

Zhang, F., Blasiak, L.C., Karolin, J.O., Powell, R.J., Geddes, C.D., and Hill, R.T. (2015). Phosphorus sequestration in the form of polyphosphate by microbial symbionts in marine sponges. *Proceedings of the National Academy of Science USA.* 112, 4381 – 4386. https://doi.org/10.1073/pnas.1423768112.

Zhang, F., Jonas, L., Lin, H., and Hill, R.T. (2019). Microbially mediated nutrient cycles in marine sponges. *FEMS Microbiology Ecology.* 95, fiz155. https://doi.org/10.1093/femsec/fiz155

Zlotkowski, K., Hewitt, W.M., Yan, P., Bokesch, H.R., Peach, M.L., Nicklaus, M.C., O'Keefe, B.R., McMahon, J.B., Gustafson, K.R., Schneekloth, J.S., Jr. (2017). Macrophilone A: Structure elucidation, total synthesis, and functional evaluation of a biologically active iminoquinone from the marine hydroid *Macrorhynchia philippina*. *Org. Lett.,* 19: 1726 – 1729. doi: 10.1021/acs.orglett.7b00496.

Zou Y., Wang X., Sims J., Wang B., Pandey P., Welsh C.L., Stone R.P., Avery M.A., Doerksen R.J., Ferreira D. (2019). Computationally Assisted Discovery and Assignment of a Highly Strained and PANC-1 Selective Alkaloid from Alaska's Deep Ocean. *J. Am. Chem. Soc.,* 141: 4338 – 4344. doi: 10.1021/jacs.8b11403.

Zou Y., Hamann M.T. (2013). Atkamine: A New Pyrroloiminoquinone Scaffold from the Cold Water Aleutian Islands *Latrunculia* Sponge. *Org. Lett.* 15: 1516 – 1519. doi: 10.1021/ol400294v.

Appendix A: Supplementary data

A link to all the supplementary data generated in this study can be found as Excel spreadsheets in the Google Drive folder titled “MSc Appendix A - Supplementary Data”.

MSC Appendix A - Supplementary Data - Google Drive

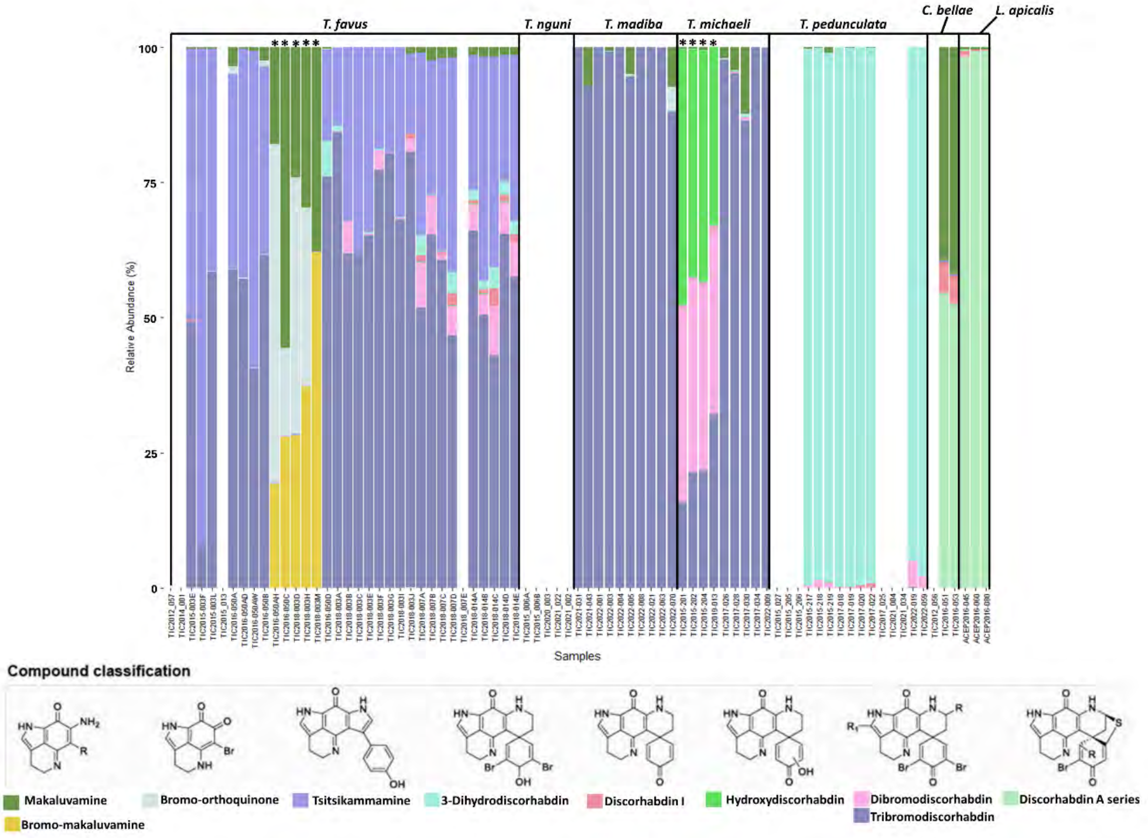

Figure S1: Relative abundance of different pyrroloiminoquinone classes per latrunculid sponge sample as indicated by the coloured key. Scaffold structures are provided for identified chemical groups. Blank regions indicated sponges that have not had their chemistry assessed. Sponges classified as Chemotype II are indicated with an asterisk.

Appendix B – Custom scripts

A link to all custom scripts and their respective input files can be found in the Github Repo under the profile Gabby9902 accessible using the following link:

Gabby9902/Custom-python-scripts-generated-for-MSc-thesis.: A list of all the custom programme and python scripts used to complete my MSc research.

Appendix C – Publications and conferences

Data from this thesis was presented at the following conference proceedings under the following titles:

Poster presentation at H3D Drug discovery symposium in 2022 titled: “Integrating genomics and metabolomics for the elucidation of secondary metabolite biosynthetic gene clusters by sponge microbiomes”.

Poster presentation at SASBI/SAGS Bio2024 conference in 2024 titled: “Genome-resolved metagenomic analysis of a conserved spirochete symbiont in pyrroloiminoquinone producing sponges”.

Below is a copy of the manuscript “The unique and enigmatic spirochete symbiont of latrunculid sponges” accepted at *mSphere* awaiting publication.

AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere

| Microbial Ecology | Research Article

The unique and enigmatic spirochete symbiont of latrunculid sponges

Samantha C. Waterworth,[1] Gabriella M. Solomons,[1] Jarmo-Charles J. Kalinski,[1] Luthando S. Madonsela,[1] Shirley Parker-Nance,[1,2,3] Rosemary A. Dorrington[1,3]

AUTHOR AFFILIATIONS See affiliation list on p. 15.

ABSTRACT Bacterial symbionts are critical members of many marine sponge holobionts. Some sponge-associated bacterial lineages, such as Poribacteria, sponge-associated unclassified lineage (SAUL), and Tethybacterales, appear to have broad-host ranges and associate with a diversity of sponge species, while others are more species-specific, having adapted to the niche environment of their host. Host-associated spirochete symbionts that are numerically dominant have been documented in several invertebrates including termites, starfish, and corals. However, dominant spirochete populations are rare in marine sponges, having thus far been observed only in *Clathrina clathrus* and various species within the Latrunculiidae family, where they are co-dominant alongside Tethybacterales symbionts. This study aimed to characterize these spirochetes and their potential role in the host sponge. Analysis of metagenome-assembled genomes from eight latrunculid sponges revealed that these unusual spirochetes are relatively recent symbionts and are phylogenetically distinct from other sponge-associated spirochetes. Functional comparative analysis suggests that the host sponge may have selected for these spirochetes due to their ability to produce terpenoids and/or possible structural contributions.

IMPORTANCE South African latrunculid sponges are host to co-dominant Tethybacterales and Spirochete symbionts. While the Tethybacterales are broad-host range symbionts, the spirochetes have not been reported as abundant in any other marine sponge except *Clathrina clathrus*. However, spirochetes are regularly the most dominant populations in marine corals and terrestrial invertebrates where they are predicted to serve as beneficial symbionts. Here, we interrogated eight metagenome-assembled genomes of the latrunculid-associated spirochetes and found that these symbionts are phylogenetically distinct from all invertebrate-associated spirochetes. The symbiosis between the spirochetes and their sponge host appears to have been established relatively recently.

KEYWORDS symbiosis, marine sponges, spirochetes, metagenomics, pyrroloiminoquinones, bacterial communities

Editor Paul D. Fey, University of Nebraska Medical Center College of Medicine, Omaha, Nebraska, USA

Address correspondence to Rosemary A. Dorrington, r.dorrington@ru.ac.za.

The authors declare no conflict of interest.

See the funding table on p. 15.

Received 9 October 2024

Accepted 23 October 2024

Published 21 November 2024

The development of symbiotic relationships with prokaryotes likely predates the emergence of marine sponges (phylum Porifera) during the Cambrian explosion ~540 million years ago (1, 2), and these associations have played a critical role in the evolution of modern sponge taxa (3, 4). Bacterial symbionts have co-evolved with their host to perform specific, specialized services that promote the health and fitness of the host (5). The symbionts are involved in nitrogen, sulfur, and phosphorus cycling (6–9), carbon cycling, detoxification (10, 11), and, in some cases, the production of bioactive secondary metabolites as chemical defenses against pathogens, predators, and competitors (12, 13). In return, the host provides its symbionts with a safe and nutrient-rich environment that promotes the fitness and survival of the symbiont (14). The taxonomic and functional diversity of sponge-associated microbiomes is generally host-specific, distinct from the surrounding water column, and is acquired by recruitment and enrichment from the environment (5, 15, 16). However, there are a small number of specialized symbionts acquired by vertical inheritance from the parent sponge that is broadly distributed across phylogenetically distant sponge hosts (17, 18), including the Poribacteria, the “sponge-associated unclassified lineage” (SAUL), and the recently discovered Tethybacterales symbionts (15, 19, 20).

The Tethybacterales represents a clade of cosmopolitan sponge-associated symbionts, comprising three families, namely the *Candidatus* Persebacteraceae, *Candidatus* Tethybacteraceae, and *Candidatus* Polydorabacteraceae (17, 20). As with the Poribacteria and Desulfobacteria, the Tethybacterales symbionts are present in phylogenetically diverse taxa that are primarily low-microbial abundance (LMA) sponge species, but these bacteria have also been detected in some high-microbial abundance species (17, 20). Characterization of metagenome-assembled genomes (MAGs) of different species of the three Tethybacterales families and their associated hosts also indicates that there were multiple acquisition events, and host adaptation and co-evolution began after each acquisition event (17).

Sponges of the family Latrunculiidae (Demospongiae and Poecilosclerida) are known to be prolific producers of cytotoxic pyrroloiminoquinone alkaloid compounds (21– 26) with pharmaceutical potential [reviewed in Kalinksi et al. (27)]. It has recently been discovered that there are two chemotypes present in the *Tsitsikamma favus* and *Tsitsikamma michaeli* latrunculid sponges (21, 28). Latrunculids are LMA sponges with highly conserved microbiomes that are dominated by Tethybacterales and Spirochete taxa (22, 29). The *T. favus* microbiome is dominated by two sponge-specific bacterial species defined by their 16S rRNA gene sequence, clones Sp02-1 and Sp02-3. The Sp02-1 symbiont has been recently characterized (17) and is classified as *Ca*. Ukwabelana africanus, a member of the *Ca*. Persebacteraceae family within the Tethybacterales (17). The *Ca*. U. africanus symbiont is phylogenetically related to symbionts in sponges across multiple orders within the Demospongiae and may be involved in the reduction of nitrogen and sulfur in the sponge holobiont (17).

Unlike *Ca*. U. africanus (Sp02-1), the co-dominant spirochete (Sp02-3) is not representative of a globally distributed, broad-host range sponge symbiont. Spirochetes have been reported as minor members of several sponge microbiomes (30–32), but numerically dominant populations of spirochetes in sponges have only been reported in Latrunculiidae species, which are endemic to the southeastern coast of South Africa, and in the distantly related *Clathrina clathrus* (Calcarea and Clathrinida), collected by Neulinger and colleagues from the Adriatic Sea off the coast of Croatia (33). In addition, spirochetes, presumed to be symbionts, have been detected in the embryonic and larval cells of the marine sponge *Mycale laevis*, but their role is currently unknown (34, 35). Numerically dominant spirochete species are, however, present in several other marine invertebrates, including sea anemones (36) and sea stars (37, 38), where a decreased abundance of certain spirochete populations correlates with an increased incidence of disease (38). Spirochaeta symbionts are also commonly present as dominant populations in corals (39–42) and in termite guts (43), where they may be involved in the fixation of carbon or nitrogen (41). A recent study investigating the association between coral hosts and their associated microbiota found that *Spirochaeta* was most abundant in the coral skeleton, hypothesizing that they may be key members in coral skeletal environment due to their ability to fix carbon and nitrogen (44).

The aim of the present study was to understand the relationship between latrunculid sponges and the Sp02-3 spirochete symbiont. Here, we report the characterization of eight spirochete MAGs from four *Tsitsikamma* sponge species and use comparative genomics to shed light on factors that may drive their conservation. Comparative analysis relative to publicly available genomes and MAGs of the Spirochaetaceae family suggests that the Sp02-3 spirochetes are distinct from all other sponge-associated spirochetes.

RESULTS AND DISCUSSION

Previous studies identified two closely related spirochete species, Sp02-3 and Sp02-15, in the *T. favus* microbiome (22). Subsequently, the Sp02-3 symbiont was shown to be present in the microbiomes of other *Tsitsikamma* species and *Cyclacanthia bellae* (29). Our aim in this study was to characterize the genome of the Sp02-3 symbiont to better understand its role in the sponge holobiont.

Survey of microbial communities in latrunculid sponges and other sponge species endemic to the South African coast

To survey the prevalence of spirochetes in sponge collected off the South African coastline, we clustered 16S rRNA gene fragment amplicons sourced from 155 marine sponges and eight seawater samples into operational taxonomic units (OTUs) at a distance of 0.03 in mothur (45). These sponges were collected primarily from reefs within Algoa Bay, South Africa but also included samples from the Tsitsikamma National Park, the Amathole Marine Protected Area in the Indian Ocean, and the remote Bouvet Island in the Southern (Antarctic) Ocean (Table S1).

A total of 9,711 OTUs were recovered from the 163 amplicon libraries. We identified spirochete OTUs with classifications from alignment of the OTUs against the SILVA and nr databases (Table S2). A total of 142 OTUs were classified within the Spirochaetota phylum, of which only 10 had an average abundance greater than 0.01% across all sponge specimens (Fig. 1A). OTU3 and OTU59 were most abundant in the *Tsitsikamma* and *Cyclacanthia* sponges. These OTUs were most closely related to spirochete 16S rRNA gene clones Sp02-3 and Sp02-15, previously identified in *T. favus* sponges (22). These two OTUs were present at low abundance in the *Latrunculia algoaensis* and *Latrunculia apicalis* sponge specimens (collected in Algoa Bay and the Southern Antarctic Ocean), as well as in some *Mycale* specimens and a single sympatric *Phorbus sp*. sponge (Fig. 1B). As the *Mycale* specimens were found as encrusting species on the *T. favus* sponges, we cannot discount the possibility of contamination between these two species. As we have only a single *Phorbus* sp. representative, additional specimens will be required to determine the significance of these spirochete OTUs in this genus or whether this was a result of contamination during collection. These two OTUs were otherwise absent in all other non-latrunculid sponges collected from sympatric regions. The presence, albeit low, of OTU3 and OTU59 in the *L. apicalis* sponges collected just off of Bouvet Island (~3,000 km/1,800 miles from Algoa Bay), and the presence of phylogenetically distinct spirochetes in sympatric non-latrunculid sponges of Algoa Bay would suggest that these Sp02-3 and Sp02-15 spirochetes are specifically associated with latrunculid sponges.

Spirochete OTUs, OTU105 and OTU128, were relatively abundant in other sponges collected from the South African coast. However, they were absent in latrunculid sponges and appeared more sporadic in their distribution among sponge specimens (Fig. 1B). These OTUs were most closely related to spirochetes detected in *Spongia officinalis* (OY759747.1) and *Astrosclera willeyana* (HE985144.1) sponges, respectively (Table S2). Inspection of phylogeny of these 10 OTUs (Fig. 1C) revealed that 6 of the 10 spirochete OTUs formed a clade with spirochete clones previously cloned from *T. favus* sponges (22). Of the remaining four, OTU105 and OTU128 (which were more abundant in non-Latrunculid sponge specimens) were part of distant clades of other sponge-associated spirochetes, while OTU581 and OTU399 belonged to a clade stemming from a variety of environments (Fig. 1C). Notably, a clone (Sp02sw36) isolated from the seawater extruded from *T. favus* sponges in 2012 (22) was closely related to spirochetes associated with crown-of-thorns starfish (37) and the dominant spirochete found in *Clathrina clathrus* sponges (33).

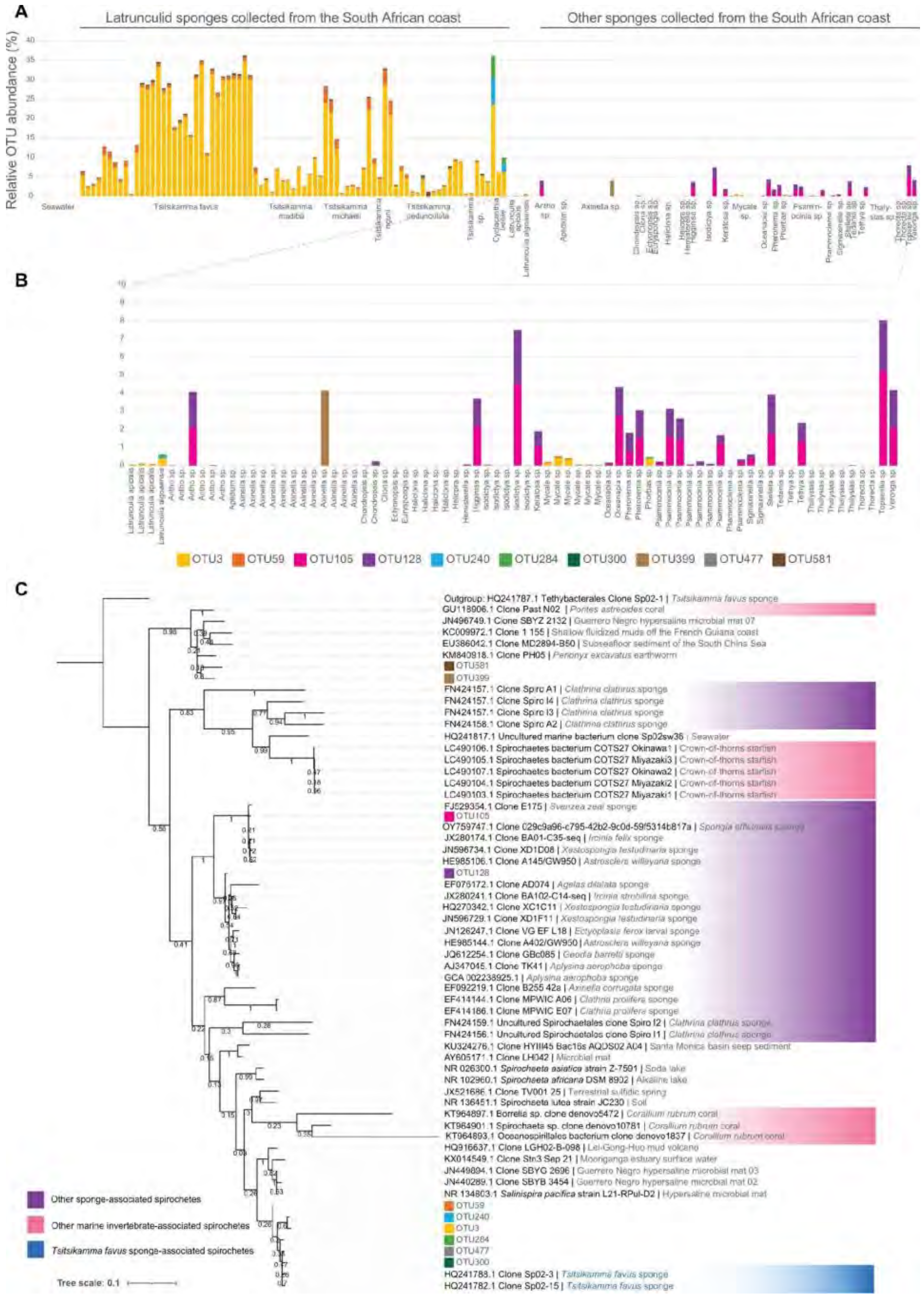

FIG 1 Spirochete population distribution in sponges collected from the South African coast and the Antarctic Southern Ocean. (A) The relative abundance of OTUs clustered at a distance of 0.03 and classified as spirochetes. (B) A magnified view of the spirochete OTUs present in non-latrunculid sponges collected from the southeastern coast of South Africa, three *L. apicalis* sponges collected from the Southern Ocean, and one sympatric *L. algoaensis* sponge. (C) Maximum-likelihood phylogeny (with 1,000 bootstraps) of the top 10 most abundant spirochete OTUs recovered from the sponges included in this study.

Characterization of *Tsitsikamma* sponge-associated spirochete MAGs

Eight sponges including five *T. favus* specimens (TIC2015-050A, TIC2015-050C, TIC2018-003B, TIC2018-003D, and TIC2018-003M) and one each of *T. michaeli* (TIC2019-013N), *Tsitsikamma madiba* (TIC2022-

009), and *Tsitsikamma pedunculata* (TIC2022-059) were selected for metagenomic analysis (Table S1). Following assembly, binning, and taxonomic classification, eight spirochete MAGs were identified, one from each of the eight *Tsitsikamma* sponge metagenomes: MAGs 050A_2, 050C_7, 003B_7, 003D_7, 003M_1, 059_1, 013N_1, and 009_1 (Table 1; Table S3). The 16S rRNA and 23S rRNA gene sequences from each MAG (if recovered) were aligned against the NR nucleotide database via online BLASTn (46).

The 16S rRNA gene sequences recovered from three MAGs all showed the greatest sequence identity with “uncultured marine clone Sp02-3,” representing the conserved spirochete symbiont previously identified in *T. favus* sponges (22, 29). All recovered 23S rRNA sequences shared the greatest sequence similarity with *Salinispira pacifica* L21-RPul-D2. This *S. pacifica* strain, isolated from a hypersaline microbial mat (47), was previously shown to be the closest known relative of the conserved spirochete Sp02-3 clone (22, 29). Finally, all eight *Tsitsikamma*-associated spirochete MAGs were taxonomically classified, via GTDB-Tk (48), within the *Salinispira* genus (Table S3). Therefore, we were confident that these MAGs represented the conserved spirochete symbiont (Sp02-3) previously reported in South African latrunculid sponges.

Phylogeny of *Tsitsikamma* sponge-associated spirochete MAGs

The 16S rRNA gene sequences recovered from three of the *Tsitsikamma*-associated spirochete MAGs were aligned against their closest matches in the NR database, as well as spirochetes from other marine invertebrates (37, 39), including the dominant spirochete present in the distantly related *Clathrina clathrus* sponges (33). Inferred maximum-likelihood phylogeny from the 16S rRNA gene alignment showed that the *Tsitsikamma*-associated spirochete MAGs were distinct from all other invertebrate-associated spirochetes (Fig. S1). The *Tsitsikamma*-associated spirochete MAGs formed a distinct clade but were most closely related to spirochetes detected in non-host-associated environments including hypersaline microbial mats, seawater, estuary water, and volcanic mud.

Since phylogeny inferred by a single marker gene can be limited, several orthogonal approaches were used to assess the phylogeny of the *Tsitsikamma* sponge-associated spirochete symbionts using whole-genome data. Initially, we employed autoMLST (49) in *de novo* mode, with both concatenated alignment (Fig. 2A) and coalescent tree (Fig. 2B) approaches, using 10 MAGs/genomes acquired from other sponge hosts, *Rhopaloeides odorabile*, *Ircinia ramosa*, and *Aplysina aerophoba* (50–52), as references. The resultant phylogenies from these two approaches had largely congruent topologies, with the *Tsitsikamma* sponge-associated Sp02-3 symbionts and other sponge-associated spirochetes forming two related but distinct clades (Fig. 2). The closest relative of the *Tsitsikamma*-associated spirochetes was *S. pacifica*, in agreement with the 23S rRNA gene phylogeny. The *Tsitsikamma*-associated spirochetes appeared phylogenetically clustered following their respective hosts, rather than geographically clustered. This contrasted with other sponge-associated spirochetes that did not seem to follow any discernible pattern of possible co-phylogeny or phylosymbiosis (Fig. 2).

As an orthogonal phylogenetic approach, we generated a phylogenetic tree using Phylophlan3 (53) and RaxML (54) (Fig. S2). Along with the eight *Tsitsikamma*-associated spirochete genomes and the 10 genomes of spirochetes associated with other sponges,

TABLE 1 Characteristics of putative representative genomes of *Tsitsikamma* sponge-associated spirochete symbiont MAGs

MAG	Size (Mbp)	Quality	16S rRNA (% ID)	23S rRNA (% ID)	Host	Sponge
003B_7	1.97	Medium	N/A[a]	*Salinispira pacifica* L21-RPul-D2 (89.54%)	*T. favus*	TIC2018-003B
050A_2	2.73	Low	Uncultured marine clone Sp02-3 (99.52%)	*S. pacifica* L21-RPul-D2 (89.54%)	*T. favus*	TIC2016_050A
003D_7	2.48	High	Uncultured marine clone Sp02-3 (99.52%)	*S. pacifica* L21-RPul-D2 (89.58%)	*T. favus*	TIC2018-003D
003M_1	2.74	High	N/A	*S. pacifica* L21-RPul-D2 (89.58%)	*T. favus*	TIC2018-003M
050C_7	1.72	Medium	Uncultured marine clone Sp02-3 (99.52%)	N/A	*T. favus*	TIC2016-050C
009_1	1.47	High	N/A	*S. pacifica* L21-RPul-D2 (91.25%)	*T. madiba*	TIC2022-009
013N_1	2.33	High	N/A	*S. pacifica* L21-RPul-D2 (89.48%)	*T. michaeli*	TIC2019-013N
059_1	2.04	Medium	N/A	N/A	*T. pedunculata*	TIC2022-059

[a] N/A, not applicable.

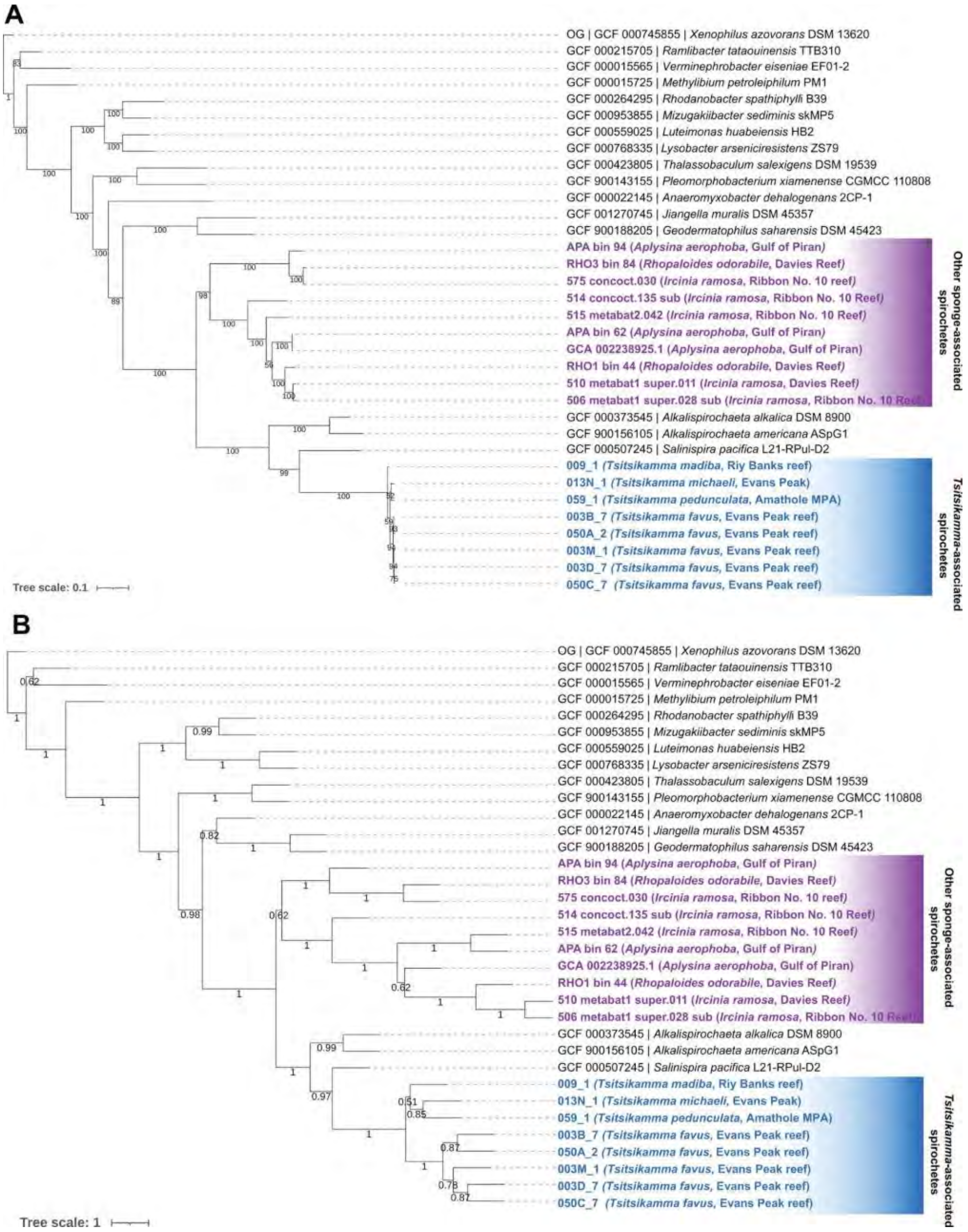

FIG 2 Phylogeny of sponge-associated spirochetes inferred with autoMLST in *de novo* mode using (A) concatenated alignment and (B) coalescent tree approaches. *Tsitsikamma*-associated spirochetes are highlighted in blue with their respective hosts. Other sponge-associated spirochetes are highlighted in purple with their associated hosts. All other reference spirochete genomes are listed in the format of "accession number | scientific name."

we included all Spirochaetaceae genomes from the National Center for Biotehnology Information (NCBI) database (*N* = 300) and all host-associated spirochete MAGs from the Joint Genome Institute (JGI) database (*N* = 44). Again, the *Tsitsikamma*-associated spirochetes formed a clade distinct from all other sponge-associated spirochete genomes. Additionally, in this analysis, we found that a MAG present in seawater (GCA 913043885.1) clustered with the other sponge-associated spirochetes. The origin of this particular genome, whether from a free-living spirochete or a sponge symbiont, remains uncertain due to potential annotation errors in the database. However, we have opted to follow the supplied annotation and presume that this MAG is likely representative of the closest free-living relative within the clade. Our phylogenetic analysis incorporated all publicly available genomes and MAGs of the Spirochaetaceae phylum, and therefore, this presumption is limited by the existing data set. We calculated pairwise average nucleotide identity (ANI) scores for all 363 spirochete genomes (Table S4). The *Tsitsikamma*-associated spirochetes shared between 93.9% and 98.2% ANI with each other (Table S5) and less than 75% ANI with any other spirochete, including their closest relative *S. pacifica*.

Estimated evolutionary divergence patterns of sponge-associated spirochetes

The divergence pattern of all sponge-associated spirochetes and their closest known free-living relatives was estimated using their rate of synonymous substitutions (dS) in orthologous genes present in all genomes. Visualization of the pairwise dS revealed that the *Tsitsikamma*-associated spirochetes are evolutionarily divergent from even their closest relative, *S. pacifica* (Fig. 3). It appears that the other sponge-associated spirochetes may have begun diverging before the *Tsitsikamma*-associated spirochetes diverged from their free-living relative. The divergence pattern of the *Tsitsikamma*-associated spirochetes is congruent with the phylogeny of their sponge host and incongruent with geographic location, suggestive of phylosymbiosis. Finally, it appears that these spirochetes have only recently begun diverging from one another as they adapt to their sponge host, and their association with latrunculid sponges is more recent than that of the co-dominant Tethybacterales symbionts (17).

Comparative analysis of functional potential in spirochete genomes

The functional potential for all 363 spirochete genomes was predicted by assigning Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthologs (KO) annotations using KofamScan (58). KO counts per genome were mapped back to associated pathways detailed in the KEGG database (59) (Table S6). Dimension reduction of these counts per genome revealed distinct clusters suggestive of adaptation to the various environments from which these spirochetes were acquired (Fig. 4). The functional potential of the *Tsitsikamma*-associated spirochetes was distinct from spirochetes associated with other sponges and, interestingly, clustered more closely with the functional potential of spirochetes associated with oligochaete worms and spirochetes from hypersaline lake environments (Fig. 4).

An analysis of similarity (ANOSIM) of the same data (Table S7) showed that the functional gene repertoire of the *Tsitsikamma*-associated spirochetes and other sponge-associated spirochetes was significantly different ($P < 0.05$) from one another and from all other environments. However, when considering the associated R-values, the *Tsitsikamma*-associated spirochetes may exhibit some overlap in functional potential of spirochetes in hypersaline lakes ($R = 0.26$), sediment ($R = 0.31$), freshwater lakes ($R = 0.38$), termites ($R = 0.47$), and seawater ($R = 0.49$). This suggests that the functional repertoire of *Tsitsikamma*-associated spirochetes may be more akin to free-living species than host-associated.

The biosynthetic potential of Sp02-3 spirochetes

A total of 581 biosynthetic gene clusters (BGCs) were detected in all spirochete genomes (*N* = 363; Table S8) and clustered into gene cluster families (GCFs) at a maximum distance of 0.3 with BiG-SCAPE (60) (Fig. 5A). Six of the eight *Tsitsikamma*-associated spirochetes had only a single-predicted BGC. The remaining two MAGs, 003B_7 and 050A_2, which were of medium and low quality, respectively, had no detected BGCs, likely due to incomplete coverage of the genomes. All six BGCs were predicted to encode a terpene product and were clustered into a single GCF (GCF1). Three other GCFs (GCFs 2, 3, and 4), consisting of terpene BGCs from other sponge-associated spirochetes, were identified but did not appear to have any homology with the terpene BGC in the *Tsitsikamma*-associated Sp02-3 spirochetes (Fig. 5B). Additional BiG-SCAPE analyses were performed with less stringent cutoffs of 0.5 and 0.8, and no BGCs from other spirochete genomes or the MiBIG database were incorporated into a GCF with the terpene BGCs detected in the *Tsitsikamma*-associated spirochetes (Table S8), indicating that this BGC is likely novel. Nonetheless, the closest characterized relative of the *Tsitsikamma*-associated spirochetes, *S. pacifica*, produces an orange carotenoid-like pigment (terpenoid), which we assume is produced via the only terpene BGC present in the *S. pacifica* genome. Despite the low sequence and organizational similarities, the terpene, if produced in the latrunculid-associated spirochetes, may protect them or their host against oxidative stress, as hypothesized for the *S. pacifica* bacterium (47).

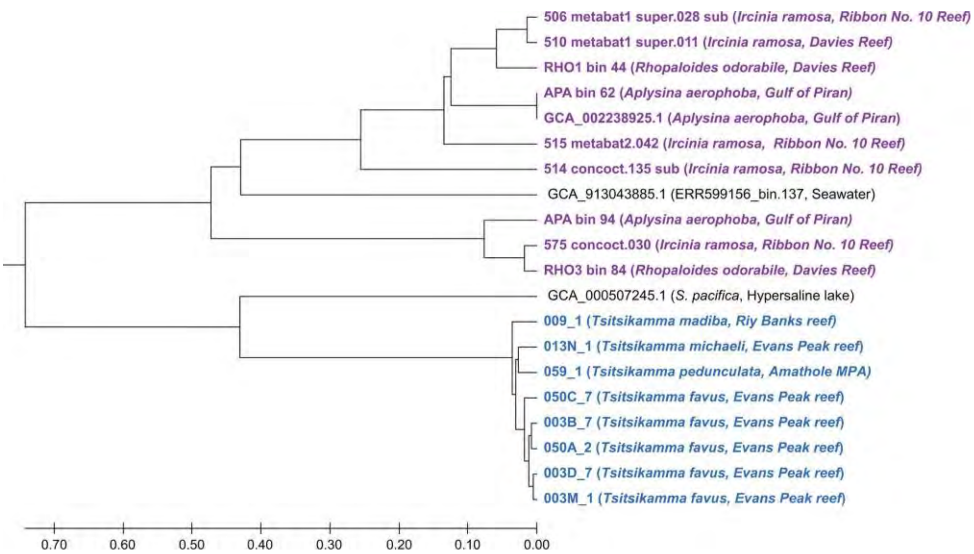

FIG 3 Unweighted pair group method with arithmetic mean (UPMGA) representation of pairwise dS of sponge-associated spirochete genomes, based on the alignment of 11 orthologous genes. PAL2NAL (55) and CodeML (56) from the PAML package were used to calculate pairwise dS values, and the resultant matrix was visualized in MEGA11 (57). The *Tsitsikammaa*ssociated spirochetes are colored in blue, and other sponge-associated spirochetes are colored in purple.

In our previous studies, we have reported the existence of two chemotypes that exist in the *T. favus* and *T. michaeli* sponge populations in Algoa Bay (21, 28). Chemotype I represents the majority of sponges, as the sponges appear visually healthy with turgid structure, and their spicules are in the canonical form. Furthermore, this Chemotype is defined by the presence of a variety of discorhabdins and tsitsikammamines (28). Conversely, the morphology of the Chemotype II sponges is considered abnormal where the tissues appear bruised and are soft to the touch (akin to rotten fruit), and many spicules are malformed (28, 61). This chemotype is further characterized by the presence of structurally simpler makaluvamines and brominated discorhabdins (21, 28).

Previous surveys of the microbial communities associated with 10 *T. favus* sponges and found no correlation between any bacterial population and the chemotypes (28). We have repeated the analysis with a larger cohort of *T. favus* and *T. michaeli* sponge specimens (*N* = 26). Using the same 16S rRNA gene amplicon data sets as presented in Fig. 1, but instead including only data from the latrunculid sponges with associated chemical data, the analysis was repeated, and OTUs were clustered at a maximum distance of 0.01 (Table S9) to disentangle the two spirochete strains previously identified in latrunculid sponges, Sp02-3 and Sp02-15 (22, 29). Using an indicator species analysis (Table S10), we found that a decrease in Sp02-3 representative OTU abundance (OTU3) and an increase in Sp02-15 representative OTU abundance (OTU6) correlated with Chemotype II sponges, relative to Chemotype I specimens (Fig. S3A and B; Table S10).

We conducted a correlation analysis of the top 50 most abundant OTUs with relative pyrroloiminoquinone abundance per sponge sample (Fig. S4; Table S11). The Sp02-3 spirochetes (OTU3) were positively correlated with the increased abundance of Chemotype I pyrroloiminoquinones and negatively correlated with the presence of

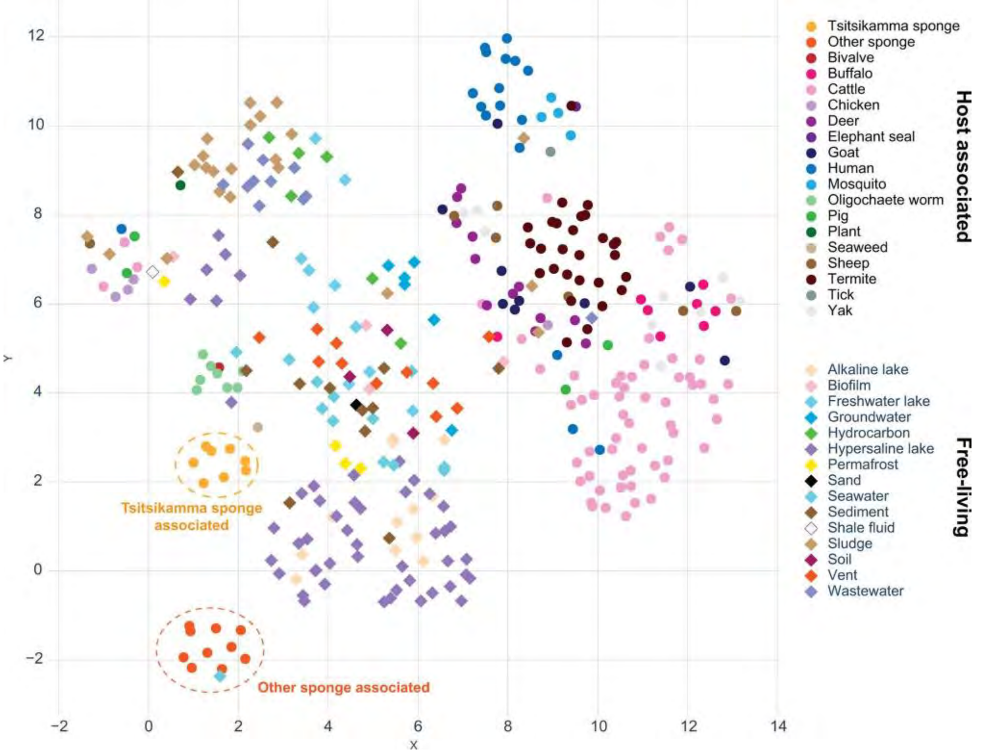

FIG 4 UMAP dimension reduction 2-dimensional representation of KEGG-annotated gene counts in all spirochete genomes. The isolation source of each genome is indicated by color and shape according to whether the isolation source is a living host (circles) or an abiotic environment (diamonds).

Chemotype II pyrroloiminoquinones. The converse was true of the Sp02-15 spirochetes (OTU6; Fig. S4; Table S11). As there was no evidence of BGCs for the production of pyrroloiminoquinones in the spirochete MAGs, this result suggests that the switch from Chemotype I to Chemotype II (the cause of which has yet to be identified) appears to negatively impact the Sp02-3 spirochete and allows the Sp02-15 spirochete to thrive in place.

Since the decrease in Sp02-3 similarly correlated with the incidence of deformed spicules, we considered whether it may play a role in spicule formation. The most closely related invertebrate-associated spirochete (Fig. 1 and 2) is a highly dominant and conserved spirochete in *Corallium rubrum* corals (39, 62). This spirochete is predicted to contribute to the coral's overall health of the coral (63) and to produce a pigmented carotenoid that influences the commercially prized color of this red coral, as the spirochete's presence correlates with the intensity of the observed red pigmentation (64). This spirochete was primarily found in the coenenchyme of the coral (64), which houses the sclerites (spicules) that are thought to act as initiation sites for the formation of the axial skeleton (65). Finally, the formation of spicules in a primary coral polyp is associated with a change in color from white to light pink (66). It is thus possible that the *C. rubrum*-associated spirochete may be involved in spicule formation as shown with the calcibacteria in *Hemimycale* sponges (pale orange to deep red in color) (67, 68) and hypothesized for the spirochetes in *Platygyra dadalea*, *Paragoniastrea australensis*, and *Porites lutea* sponges (44). While a speculative connection, as no MAG or genome is available for these spirochetes, this observation has prompted us to begin metatranscriptomic studies in conjunction with catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) experiments to determine the localization and potential structural role of spirochetes in latrunculid sponges from the South African coastline.

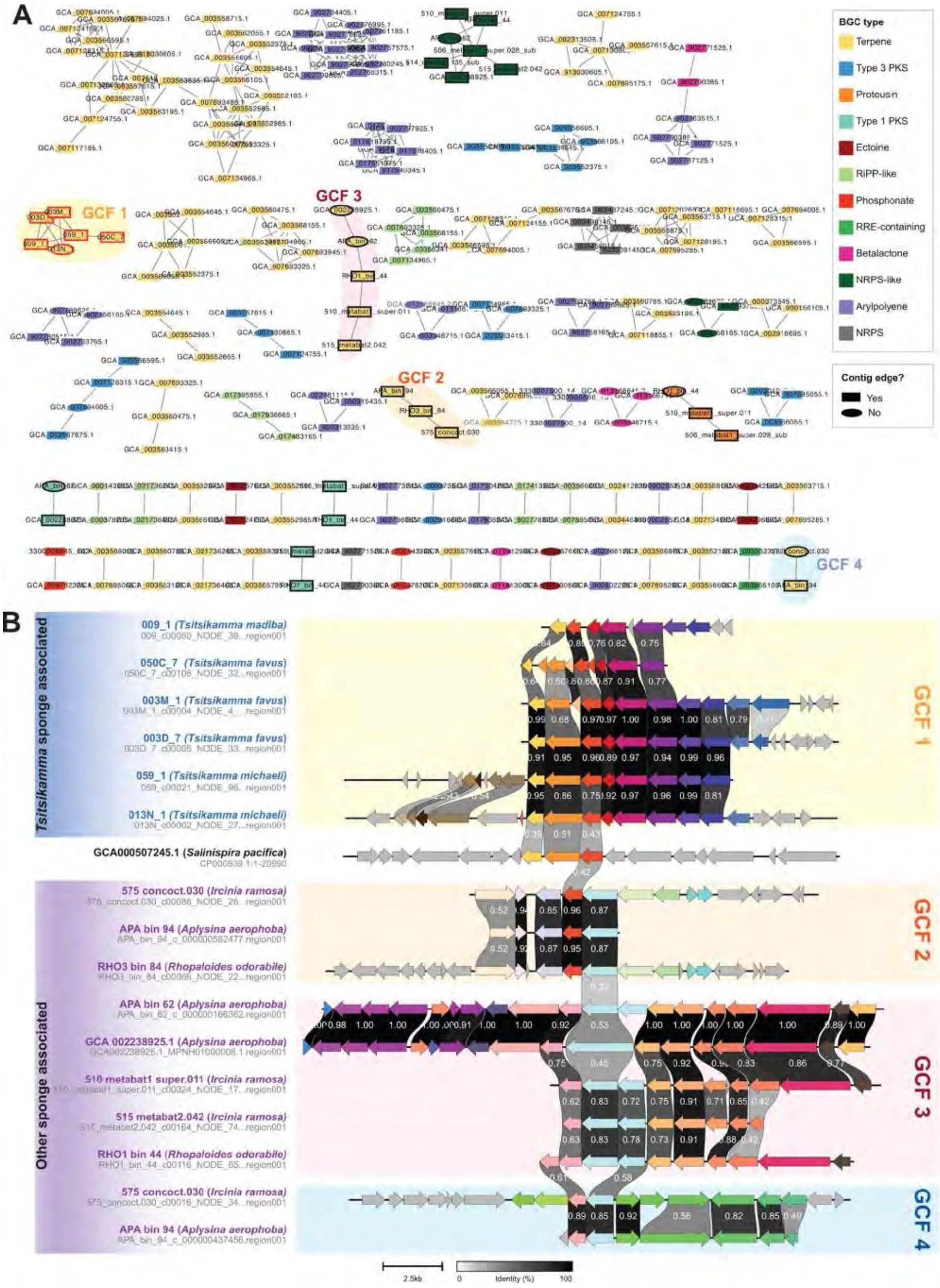

FIG 5 Assessment of biosynthetic potential in spirochetes. (A) Network visualization of BGCs from all spirochete genomes used in this study clustered into GCFs at a maximum distance of 0.3. BGCs from *Tsitsikamma*-associated spirochetes are highlighted with a red outline. BGCs from all other sponge-associated spirochetes are highlighted with a black outline. GCFs of interest are highlighted. (B) Pairwise comparison of amino-acid sequence identity of terpene BGCs from sponge-associated spirochetes. The pairwise similarity between genes is indicated between genes, and genes are colored according to their predicted function. The GCFs to which the BGCs belong have been indicated.

Conclusion

This study shows that the conserved Sp02-3 spirochete of latrunculid sponges is likely to be a relatively new symbiont that has begun co-evolving with its respective sponge hosts. The Sp02-3 symbiont is distinct from all other invertebrate-associated spirochetes, including non-dominant spirochetes associated with other marine sponges. Assessment of their functional potential suggests that the Sp02-3 spirochetes are functionally unique relative to other sponge-associated spirochetes. We found no evidence that they are directly involved in the production of the pyrroloiminoquinones characteristic of their host sponges. The close phylogenetic relatedness of the latrunculid-associated spirochetes to a dominant, conserved coral-associated spirochete hints at a possibly structural role within the sponges. However, additional experiments will be necessary to test this hypothesis.

MATERIALS AND METHODS

Sponge collection and taxonomic identification

Sponges were collected by SCUBA or remotely operated vehicle from multiple locations within the Tsitsikamma Marine Protected Area, Algoa Bay (Port Elizabeth), the Amathole Marine Protected Area (East London), and the Garden Route National Park. In addition, three *L. apicalis* specimens were collected by trawl net off Bouvet Island in the South Atlantic Ocean. Collection permits were acquired prior to collections from the Department of Environmental Affairs and the Department of Environment, Forestry, and Fisheries under permit numbers: 2015: RES2015/16 and RES2015/21; 2016: RES2016/11; 2017: RES2017/43; 2018: RES2018/44; 2019: RES2019/13; 2020: RES2020/31; 2021: RES2021/81; 2022: RES2022/70. Collection metadata are provided in Table S1. Sponge specimens were stored on ice during collection and moved to −20°C on return to the lab. Subsamples of each sponge, collected for DNA extraction, were preserved in RNALater (Invitrogen) and stored at −20°C. Sponge specimens were identified through inspection of gross morphology, spicule analysis, and molecular barcoding, as performed previously (21, 28, 29, 61).

Bacterial community profiles in latrunculid sponges

The V4–V5 of the 16S rRNA gene was PCR amplified from 79 latrunculid sponges collected between 1994 and 2022 (See Table S1 for collection data). Amplicons were sequenced using the Illumina MiSeq platform and curated using mothur (v.1.48.0) (45). All raw amplicon read data can be accessed under accession number PRJNA508092. Briefly, sequences that were shorter than 250 nt in length, longer than 350 nt in length, had homopolymeric runs of 7 nt or more, had ambiguous bases, or had a sliding window quality average lower than 20, were removed from the data sets. Chimeric sequences were detected using VSEARCH (69) and removed from the data set. Sequences were then classified via alignment against the SILVA database (v138.1), and any sequences classified as “chloroplast,” “mitochondria,” “unknown,” “archaea,” or “eukaryota” were removed. Sequences were clustered into OTUs at a distance of 0.03, and read counts thereof were converted to relative abundance (Table S2). Representative sequences of each OTU were aligned against the SILVA database (v138.1) in mothur and against the nt prokaryotic database using standalone blastn (70), using parameters -max_hsps 1 -max_target_seqs 1 to return only the first match. Descriptions and isolation sources for each returned accession were retrieved using the esearch, efetch, and xtract methods from the stand-alone entrez package (71). Spirochete OTUs were subset out and aligned with reference sequences from the NCBI nucleotide database using MUSCLE (v. 5.1) (72, 73), and phylogeny was inferred from the alignment using the maximum-likelihood method with 1,000 bootstrap replicates in MEGA11 (74). Finally, the same analysis was repeated but using only the raw amplicon read data from latrunculid sponges, and the OTUs were clustered at a distance of 0.01. In all other respects, the analyses were identical.

Chemical analysis and chemotype identification

Sponge extracts were prepared by extraction with methanol, drying *i. vac.*, and resuspension in methanol at 1–10 mg/mL. LC-MS/MS data were acquired on a Bruker ESI-Q-TOF Compact (Bruker, Bremen) in positive ionization mode coupled to a Dionex Ultimate 3000 Chromatograph (ThermoScientific, Sunnyvale, CA, USA) and using reversed-phase C18 columns and mobile phases consisting of water and acetonitrile with 0.1% formic acid each, using one of two methods (see Supplementary Methods for details). The data were converted to mzXML format and analyzed using MZmine3 (75) to assemble an aligned feature list (see Supplementary Methods for details). The feature list was filtered based on comparison of *m/z* values and MS/MS spectra to known or putative pyrroloiminoquinones. Peak area values were normalized to the overall pyrroloiminoquinone signal per sample and aggregated to the pyrroloiminoquinone class to summarize the latrunculid pyrroloiminoquinone profiles.

Correlation of spirochete populations and sponge chemotypes

An indicator species analysis was performed using the OTUs clustered at a distance of 0.01 for all *T. favus* and *T. micheali* sponges for which a chemotype had been assigned (16S_Chemotype_Indicator_Species_Analysis.R) to determine which OTUs, if any, were associated with the two chemotypes. The co-correlation analysis of the 50 most abundant OTUs (found as an average across all samples) was performed using the "cor" function (76) native to R using dataframes of OTU and compound abundances as input. A 16S rRNA gene sequence phylogeny was built from the representative sequences of the top 50 OTUs, aligned with MUSCLE (v 5.1) (72, 73), using the neighborjoining approach with 1,000 bootstraps in MEGA11 (74). The final tree was visualized in iTol (77) where the correlation matrix and the average OTU abundance per sponge species were visualized alongside the tree as data sets.

Metagenomic sequencing and analysis of individual *T. favus* specimens

The DNA extraction and metagenomic sequencing of four *T. favus* sponges that resulted in the recovery of four MAGs 050A_2, 050C_7, 003B_7, and 003D_7, classified as spirochetes, are described in reference (17). In addition to these samples, four additional metagenomes of *Tsitsikamma* sponges (TIC2018-003M, TIC2019-013N, TIC2022-009, and TIC2022-059) were sequenced. These sponges were selected for sequencing based on the apparent abundance of spirochete OTUs found via 16S rRNA gene amplicon sequence.

Total genomic DNA was extracted using the Zymo Research Quick DNA Fecal/ Soil Microbe Miniprep Kit (Catalog number: D6012) according to the manufacturer's specifications and stored at −4°C. Shotgun metagenomic IonTorrent libraries of 200 bp reads were prepared and sequenced using an Ion P1.1.17

chip. All metagenomes were assembled, binned, and processed as described in reference (17). Four additional spirochete genome MAGs (003M_1, 059_1, 013N_1, and 009_1) were extracted from the new data sets. MAGs were named after the *Tsitsikamma* sponge specimen from which they were extracted (e.g., 050A_2 is the MAG from sponge specimen TIC2016-050A). The numbers associated with each MAG are an arbitrary artifact of the binning process.

Acquisition of reference genomes and MAGs

Four spirochete MAGs associated with *Aplysina aerophoba* and *Rhopaloeides odorabile* sponges from a study by Robbins and colleagues (78) were downloaded from https://data.ace.uq.edu.au/public/sponge_mags/, and five sponge-associated spirochete MAGs were acquired from the China National GeneBank DataBase from studies by O'Brien and colleagues (50, 51). One spirochete genome from an *Aplysina aerophoba* sponge was additionally downloaded from the NCBI database (GCA_002238925.1). Additionally, all other genomes classified within the Spirochaetaceae family were downloaded from the NCBI database (*N* = 300), and all host-associated spirochete MAGs were downloaded from the JGI database (*N* = 44). This resulted in a total of 354 reference genomes (Table S3).

Characterization of MAGs and genomes

All scripts used for bioinformatic analyses, along with their associated inputs, used in the following methods can be found at https://github.com/samche42/Spirochete. All MAGs and genomes used in this study were assessed using CheckM (v1.1.3) (79) and taxonomically classified using GTDB-Tk (v2.3.2) (48) against the Release 214.1 reference database. Basic metrics such as size, number of contigs, and N50 were calculated using bin_summary.py. The number of genes, pseudogenes, and coding density per genome were calculated using all_included_genome_characteristics.py. All metadata per genome or MAG can be found in Table S3.

Phylogeny of spirochete genome MAGs extracted from individual *Tsitsikamma* sponges

Ribosomal sequences (23S rRNA, 16S rRNA, and 5S rRNA) were extracted from individual

MAGs using barrnap (v 0.9) (80). The closest matches of recovered 16S sequences from sponge-associated MAGs were identified using BLASTn (v 2.7.1) (70). Resultant sequences were aligned using MUSCLE (v. 5.1) (72, 73), and phylogeny was inferred using the maximum-likelihood method with 1,000 bootstraps in MEGA11 (74). Phylogeny of the *Tsitsikamma*-associated spirochete MAGs was similarly inferred using whole-genome data via autoMLST (49) and PhyloPhlan3 (53). Amino acid sequences and nucleotide sequences for all genes were found in all genomes using prokka (v 1.13) (81). The phylogeny of all 362 MAGs and genomes (8 *Tsitsikamma*-associated spirochete MAGs and 354 references) was inferred using Phylophlan3. Phylophlan3 was run with diversity set to medium, with default values in the supermatrix_aa configuration. The resultant gene protein alignment was used in RaxML (v 8.2.12) (82) to build a phylogenetic tree with 1,000 bootstrap replicates using the PROTGAMMAAUTO model. The resultant tree was visualized in iTol (77). Genomes from Myxococcota (GCA_002691025.1) and

Deltaproteobacteria (GCA_020632655.1) were chosen as outgroups. These genomes had been downloaded from the NCBI database as their metadata indicated that they were classified within the Spirochaetaceae family. However, the taxonomic classification of these genomes with GTDB-Tk revealed that these genomes had likely been misclassified. These genomes were considered serendipitous choices for outgroups for the Phylophlan3 analysis. AutoMLST was deployed in *de novo* mode using concatenated alignments and coalescent trees of marker genes in two separate analyses. ModelFinder and IQ-TREE Ultrafast Bootstrap analysis were enabled in both analyses. All latrunculidassociated and other sponge-associated spirochete MAGs were included in this analysis. MAGs and genomes from JGI and NCBI were not used in this analysis as the number of query genomes is limited to 20, so we opted to include only sponge-associated spirochetes in this analysis. Resultant trees were downloaded in Newick format and revisualized in iTol (77). Finally, the pairwise ANI was calculated for all genomes using fastANI (v1.33) (57). If a pairwise alignment fraction was lower than 70% (83), the associated ANI score was nullified as the accuracy of the ANI score could not be trusted.

Estimated evolutionary divergence patterns of sponge-associated spirochetes

Using the Phylophlan3 (53) and autoMLST (49, 53) trees as guidance, orthologous genes from the eight *Tsitsikamma*-associated spirochetes, the 10 other sponge-associated spirochetes, and their closest relatives were identified using OMA (v. 2.6.0) (84). A total of 11 orthologs common to all genomes were found using count_OGs.py and aligned using MUSCLE (v 5.1) (72, 73). The corresponding nucleotide sequence for each gene was retrieved using streamlined_seqretriever.py, all stop codons were removed using remove_stop_codons.py, and nucleotide sequences were aligned using MUSCLE (v 5.1) (72, 73). Ortholog gene sequences were grouped per genome using merge_fasta_for_dNdS.py. The nucleotide and amino acid sequences (per genome) were each concatenated union function from EMBOSS (85) and aligned using PAL2NAL

(86). The alignment was used to estimate pairwise dS and thereby infer the pattern of divergence between these genomes using codeml from the PAML package (87).

Comparative analysis of functional potential in spirochete genomes

Genes were identified in all genomes/MAGs using Prokka (v 1.13) (81) and then annotated against the KEGG database using KOfamSCAN (58) with detail-tsv as the output format. Reliable annotations were extracted from these results based on the criteria that the annotation score is greater than the estimated threshold, and then reliable annotations per MAG/genome were counted and summarized using the kegg_parser.py script. This produced a table of KO counts per genome that was used as input for both ANOSIM processing and dimension reduction, via UMAP (88), for three- and 2-dimensional visualizations (dimension_reduction.py). A Jupyter notebook is provided in the GitHub repository for easy reproduction and an interactive 3D figure. To find statistically significant KEGG-annotated drivers of the different samples, we performed a re-purposed indicator species analysis with the number of KEGG annotations per KO per genome in place of OTU abundance. This was performed using the multiplatt method from the “indicspecies” package in R (89) with 1,000 permutations and specifying the point biserial correlation coefficient (“r.g”) as the association index as this both accounts for abundance data (rather than presence/absence data) and corrects for the different number of samples per host type.

The biosynthetic potential of sponge-associated spirochetes

A total of 547 BGCs were predicted from all spirochete genomes (*N* = 363) using antiSMASH (v. 6.0.1) (55) with --cb-general --cb-knownclusters --cb-subclusters --asf --pfam2go --smcog-trees options enabled and genes found with prodigal. The resultant putative BGCs were clustered twice using BiG-SCAPE (v 1.1.5) (60) at maximum distances of 0.3, 0.5, and 0.8. Network files of non-singleton GCFs were visualized in Cytoscape (56). Highlighted gene clusters of interest were visualized with clinker (90). Metadata for BGCs were extracted from individual GenBank files using antismash_summary.py.

ACKNOWLEDGMENTS

We would like to acknowledge Gwynneth Matcher (South African Institute for Aquatic Biodiversity, Aquatic Genomics Research Platform), Carel van Heerden, and Alvera Vorster (Stellenbosch University Central Analytical Facility) for next-generation sequencing technical support. We thank Ryan Palmer and Koos Smith (African Coelacanth Ecosystem Program) for logistics and technical support during sponge collections. We thank the South African Environmental Observation Network, Elwandle Coastal Node, and the Shallow Marine and Coastal Research Infrastructure for the use of their research platforms and infrastructure for their assistance in SCUBA collections and logistical support.

This research was supported by South African National Research Foundation grants to R.A.D., including the South Africa Research Chair Initiative (SARChI) grant (UID: 87583) and the SARChI-led Communities of Practice Program (UID: 110612). S.C.W. was supported by an NRF Innovation and Rhodes University Henderson Ph.D. scholarships. G.M.S. and L.S.M. were supported by NRF Masters and PhD scholarships, respectively. S.P.-N. was supported by an NRF PDP scholarship (UID: 101038). J.-C.J.K. was supported by funding awarded to R.A.D. by the South African Medical Research Council as well as the UK Medical Research Council, with funds received from the UK Government’s Newton Fund (Grant No.: 96185).

We declare no competing interests, financial or otherwise, in relation to the work described here. The opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to any of the above-mentioned donors.

AUTHOR AFFILIATIONS

[1]Department of Biochemistry and Microbiology, Rhodes University, Makhanda, South Africa

[2]South African Environmental Observation Network, Elwandle Coastal Node, Port Elizabeth, South Africa

[3]South African Institute for Aquatic Biodiversity, Makhanda, South Africa

PRESENT ADDRESS

Samantha C. Waterworth, National Cancer Institute, Frederick, Maryland, USA

Jarmo-Charles J. Kalinski, Department of Biochemistry, University of California, Riverside, California, USA

AUTHOR ORCIDs

Samantha C. Waterworth http://orcid.org/0000-0001-6436-0142

Rosemary A. Dorrington http://orcid.org/0000-0002-8694-367X

FUNDING

Funder	**Grant(s)**	**Author(s)**
National Research Foundation (NRF)	87583	Rosemary A. Dorrington
National Research Foundation (NRF)	110612	Rosemary A. Dorrington
South African Medical Research Council (SAMRC)	96185	Rosemary A. Dorrington
National Research Foundation (NRF)	101038	Shirley Parker-Nance

AUTHOR CONTRIBUTIONS

Samantha C. Waterworth, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review and editing | Gabriella M. Solomons, Data curation, Formal analysis, Investigation, Writing – review and editing | Jarmo-Charles J. Kalinski, Data curation, Formal analysis, Methodology, Writing – review and editing | Luthando S. Madonsela, Data curation, Methodology, Writing – review and editing | Shirley ParkerNance, Investigation, Methodology, Resources | Rosemary A. Dorrington, Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

All sequence data can be accessed under accession number PRJNA508092 in the NCBI SRA database. All scripts used for analysis and visualization can be accessed at https:// github.com/samche42/Spirochete.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Fig. S1 (mSphere00845-24-s0001.eps). 16S phylo tree.

Fig. S2 (mSphere00845-24-s0002.eps). Phylo tree of MAGs with Phylophlan.

Fig. S3 (mSphere00845-24-s0003.eps). Spirochete OTU distributions.

Fig. S4 (mSphere00845-24-s0004.eps). OTU and compound abundance in chemotypes.

Supplemental tables, part 1 (mSphere00845-24-s0005.xlsx). Tables S1 to S3.

Supplemental tables, part 2 (mSphere00845-24-s0006.xlsx). Tables S4 to S6.

Supplemental tables, part 3 (mSphere00845-24-s0007.xlsx). Tables S7 to S11.

REFERENCES

Botting JP, Nettersheim BJ. 2018. Searching for sponge origins. Nat Ecol Evol 2:1685–1686. https://doi.org/10.1038/s41559-018-0702-4

Botting JP, Muir LA. 2018. Early sponge evolution: a review and phylogenetic framework. Palaeoworld 27:1–29. https://doi.org/10.1016/j.palwor.2017.07.001

Wilkinson CR, Smith DC. 1984. Immunological evidence for the Precambrian origin of bacterial symbioses in marine sponges. Proc R Soc Lond B 220:509–518. https://doi.org/10.1098/rspb.1984.0017

Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J, Moore BS. 2002. Molecular evidence for a uniform microbial community in sponges from different oceans. Appl Environ Microbiol 68:4431–4440. https://doi.org/10.1128/AEM.68.9.4431-4440.2002

Thomas T, Moitinho-Silva L, Lurgi M, Björk JR, Easson C, Astudillo-García C, Olson JB, Erwin PM, López-Legentil S, Luter H, Chaves-Fonnegra A, Costa R, Schupp PJ, Steindler L, Erpenbeck D, Gilbert J, Knight R, Ackermann G, Victor Lopez J, Taylor MW, Thacker RW, Montoya JM, Hentschel U, Webster NS. 2016. Diversity, structure and convergent evolution of the global sponge microbiome. Nat Commun 7:11870. https://doi.org/10.1038/ncomms11870

Jensen S, Fortunato SAV, Hoffmann F, Hoem S, Rapp HT, Øvreås L,

Torsvik VL. 2017. The relative abundance and transcriptional activity of marine sponge-associated microorganisms emphasizing groups involved in sulfur cycle. Microb Ecol 73:668–676. https://doi.org/10.

1007/s00248-016-0836-3

Karimi E, Slaby BM, Soares AR, Blom J, Hentschel U, Costa R. 2018. Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. FEMS Microbiol Ecol 94. https://doi.org/10.1093/femsec/fiy074

Zhang F, Blasiak LC, Karolin JO, Powell RJ, Geddes CD, Hill RT. 2015. Phosphorus sequestration in the form of polyphosphate by microbial symbionts in marine sponges. Proc Natl Acad Sci U S A 112:4381–4386. https://doi.org/10.1073/pnas.1423768112

Zhang F, Jonas L, Lin H, Hill RT. 2019. Microbially mediated nutrient cycles in marine sponges. FEMS Microbiol Ecol 95:fiz155. https://doi.org/10.1093/femsec/fiz155

de Voogd NJ, Cleary DFR, Polónia ARM, Gomes NCM. 2015. Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia. FEMS Microbiol Ecol 91:fiv019. https://doi.org/10.1093/femsec/fiv019

Hentschel U, Usher KM, Taylor MW. 2006. Marine sponges as microbial fermenters. FEMS Microbiol Ecol 55:167–177. https://doi.org/10.1111/j.

1574-6941.2005.00046.x

Helber SB, Hoeijmakers DJJ, Muhando CA, Rohde S, Schupp PJ. 2018. Sponge chemical defenses are a possible mechanism for increasing sponge abundance on reefs in Zanzibar. PLoS One 13:e0197617. https:// doi.org/10.1371/journal.pone.0197617

Lopanik NB. 2014. Chemical defensive symbioses in the marine environment. Funct Ecol 28:328–340. https://doi.org/10.1111/1365-

2435.12160

Engelberts JP, Robbins SJ, de Goeij JM, Aranda M, Bell SC, Webster NS. 2020. Characterization of a sponge microbiome using an integrative genome-centric approach. ISME J 14:1100–1110. https://doi.org/10.

1038/s41396-020-0591-9

Astudillo-García C, Slaby BM, Waite DW, Bayer K, Hentschel U, Taylor MW.

2018. Phylogeny and genomics of SAUL, an enigmatic bacterial lineage frequently associated with marine sponges. Environ Microbiol 20:561– 576. https://doi.org/10.1111/1462-2920.13965

Wang Y, Gong L, Gao Z, Wang Y, Zhao F, Fu L, Li X. 2023. Host-specific bacterial communities associated with six cold-seep sponge species in the South China Sea. Front Mar Sci 10. https://doi.org/10.3389/fmars.

2023.1243952

Waterworth SC, Parker-Nance S, Kwan JC, Dorrington RA. 2021. Comparative genomics provides insight into the function of broad-host range sponge symbionts. MBio 12:e0157721. https://doi.org/10.1128/ mBio.01577-21

Fieth RA, Gauthier M-EA, Bayes J, Green KM, Degnan SM. 2016. Ontogenetic changes in the bacterial symbiont community of the tropical demosponge *Amphimedon queenslandica*: metamorphosis is a

new beginning. Front Mar Sci 3. https://doi.org/10.3389/fmars.2016.

00228

Fieseler L, Horn M, Wagner M, Hentschel U. 2004. Discovery of the novel candidate phylum "*Poribacteria*" in marine sponges. Appl Environ Microbiol 70:3724–3732. https://doi.org/10.1128/AEM.70.6.3724-3732.

2004

Taylor JA, Palladino G, Wemheuer B, Steinert G, Sipkema D, Williams TJ, Thomas T. 2021. Phylogeny resolved, metabolism revealed: functional radiation within a widespread and divergent clade of sponge symbionts. ISME J 15:503–519. https://doi.org/10.1038/s41396-020-00791-z

Kalinski J-CJ, Krause RWM, Parker-Nance S, Waterworth SC, Dorrington

RA. 2021. Unlocking the diversity of pyrroloiminoquinones produced by

latrunculid sponge species. Mar Drugs 19:68. https://doi.org/10.3390/ md19020068

Walmsley TA, Matcher GF, Zhang F, Hill RT, Davies-Coleman MT, Dorrington RA. 2012. Diversity of bacterial communities associated with the Indian Ocean sponge Tsitsikamma favus that contains the bioactive pyrroloiminoquinones, tsitsikammamine A and B. Mar Biotechnol 14:681–691. https://doi.org/10.1007/s10126-012-9430-y

Hooper GJ, Davies-Coleman MT, Kelly-Borges M, Coetzee PS. 1996. New alkaloids from a South African latrunculid sponge. Tetrahedron Lett

37:7135–7138. https://doi.org/10.1016/0040-4039(96)01560-2

Antunes EM, Copp BR, Davies-Coleman MT, Samaai T. 2005. Pyrroloiminoquinone and related metabolites from marine sponges. Chem Inform 36. https://doi.org/10.1002/chin.200521269

Botić T, Defant A, Zanini P, Žužek MC, Frangež R, Janussen D, Kersken D, Knez Ž, Mancini I, Sepčić K. 2017. Discorhabdin alkaloids from Antarctic *Latrunculia* spp. sponges as a new class of cholinesterase inhibitors. Eur J

Med Chem 136:294–304. https://doi.org/10.1016/j.ejmech.2017.05.019

Ford J, Capon RJ. 2000. Discorhabdin R: a new antibacterial pyrroloiminoquinone from two latrunculiid marine sponges, *Latrunculia* sp. and *Negombata* sp. J Nat Prod 63:1527–1528. https://doi.org/10.1021/ np000220q

Kalinski J-CJ, Polyzois A, Waterworth SC, Siwe Noundou X, Dorrington RA. 2022. Current perspectives on pyrroloiminoquinones: distribution, biosynthesis and drug discovery potential. Molecules 27:27. https://doi. org/10.3390/molecules27248724

Kalinski J-CJ, Waterworth SC, Noundou XS, Jiwaji M, Parker-Nance S, Krause RWM, McPhail KL, Dorrington RA. 2019. Molecular networking reveals two distinct chemotypes in pyrroloiminoquinone-producing

tsitsikamma favus sponges. Mar Drugs 17:60. https://doi.org/10.3390/ md17010060

Matcher GF, Waterworth SC, Walmsley TA, Matsatsa T, Parker-Nance S, Davies-Coleman MT, Dorrington RA. 2017. Keeping it in the family: coevolution of latrunculid sponges and their dominant bacterial symbionts. Microbiologyopen 6:e00417. https://doi.org/10.1002/mbo3.

417

Villegas-Plazas M, Wos-Oxley ML, Sanchez JA, Pieper DH, Thomas OP, Junca H. 2019. Variations in microbial diversity and metabolite profiles of the tropical marine sponge *Xestospongia muta* with season and depth. Microb Ecol 78:243–256. https://doi.org/10.1007/s00248-018-1285-y 31. Isaacs LT, Kan J, Nguyen L, Videau P, Anderson MA, Wright TL, Hill RT. 2009. Comparison of the bacterial communities of wild and captive sponge *Clathria prolifera* from the Chesapeake Bay. Mar Biotechnol 11:758–770. https://doi.org/10.1007/s10126-009-9192-3

Taylor MW, Radax R, Steger D, Wagner M. 2007. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential.

Microbiol Mol Biol Rev 71:295–347. https://doi.org/10.1128/MMBR.

00040-06

Neulinger SC, Stöhr R, Thiel V, Schmaljohann R, Imhoff JF. 2010. New phylogenetic lineages of the *Spirochaetes* phylum associated with *Clathrina* species (Porifera). J Microbiol 48:411–418. https://doi.org/10.

1007/s12275-010-0017-x

Díez-Vives C, Koutsouveli V, Conejero M, Riesgo A. 2022. Global patterns in symbiont selection and transmission strategies in sponges. Front Ecol Evol 10. https://doi.org/10.3389/fevo.2022.1015592

Riesgo A, Taboada S, Sánchez-Vila L, Solà J, Bertran A, Avila C. 2015. Some like it fat: comparative ultrastructure of the embryo in two demosponges of the genus Mycale (order Poecilosclerida) from Antarctica and the Caribbean. PLoS One 10:e0118805. https://doi.org/

10.1371/journal.pone.0118805

Bonacolta AM, Connelly MT, Rosales SM, Del Campo J, Traylor-Knowles N. 2021. The starlet sea anemone, *Nematostella vectensis*, possesses body region-specific bacterial associations with spirochetes dominating the capitulum. FEMS Microbiol Lett 368:fnab002. https://doi.org/10.1093/femsle/fnab002

Wada N, Yuasa H, Kajitani R, Gotoh Y, Ogura Y, Yoshimura D, Toyoda A,

Tang S-L, Higashimura Y, Sweatman H, Forsman Z, Bronstein O, Eyal G, Thongtham N, Itoh T, Hayashi T, Yasuda N. 2020. A ubiquitous subcuticular bacterial symbiont of A coral predator, the crown-of-thorns starfish, in the Indo-Pacific. Microbiome 8:123. https://doi.org/10.1186/ s40168-020-00880-3

Høj L, Levy N, Baillie BK, Clode PL, Strohmaier RC, Siboni N, Webster NS, Uthicke S, Bourne DG. 2018. Crown-of-thorns sea star *Acanthaster* cf. *solaris* has tissue-characteristic microbiomes with potential roles in health and reproduction. Appl Environ Microbiol 84:e00181-18. https://doi.org/10.1128/AEM.00181-18

van de Water JAJM, Melkonian R, Junca H, Voolstra CR, Reynaud S, Allemand D, Ferrier-Pagès C. 2016. Spirochaetes dominate the microbial community associated with the red coral *Corallium rubrum* on a broad geographic scale. Sci Rep 6:27277. https://doi.org/10.1038/srep27277

Wessels W, Sprungala S, Watson S-A, Miller DJ, Bourne DG. 2017. The microbiome of the octocoral *Lobophytum pauciflorum*: minor differences between sexes and resilience to short-term stress. FEMS Microbiol Ecol 93. https://doi.org/10.1093/femsec/fix013

Lawler SN, Kellogg CA, France SC, Clostio RW, Brooke SD, Ross SW. 2016.

Coral-associated bacterial diversity is conserved across two deep-sea anthothela species. Front Microbiol 7:458. https://doi.org/10.3389/fmicb.

2016.00458

Park JS, Han J, Suh S-S, Kim H-J, Lee T-K, Jung SW. 2022. Characterization of bacterial community structure in two alcyonacean soft corals (*Litophyton* sp. and *Sinularia* sp.) from Chuuk, Micronesia. Coral Reefs 41:563–574. https://doi.org/10.1007/s00338-021-02176-w

Lilburn TG, Kim KS, Ostrom NE, Byzek KR, Leadbetter JR, Breznak JA. 2001. Nitrogen fixation by symbiotic and free-living spirochetes. Science 292:2495–2498. https://doi.org/10.1126/science.1060281

Ricci F, Tandon K, Black JR, Lê Cao K-A, Blackall LL, Verbruggen H. 2022. Host traits and phylogeny contribute to shaping coral-bacterial symbioses. mSystems 7:e0004422. https://doi.org/10.1128/msystems.

00044-22

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: opensource, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541. https://doi.org/10.1128/AEM.01541-09

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden

TL. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res 36:W5–

W9. https://doi.org/10.1093/nar/gkn201

Ben Hania W, Joseph M, Schumann P, Bunk B, Fiebig A, Spröer C, Klenk

H-P, Fardeau M-L, Spring S. 2015. Complete genome sequence and description of *Salinispira pacifica* gen. nov., sp. nov., a novel spirochaete isolated form a hypersaline microbial mat. Stand Genomic Sci 10:7. https://doi.org/10.1186/1944-3277-10-7

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/ btz848

Alanjary M, Steinke K, Ziemert N. 2019. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. Nucleic Acids Res 47:W276–W282. https://doi.org/10.

1093/nar/gkz282

O'Brien PA, Andreakis N, Tan S, Miller DJ, Webster NS, Zhang G, Bourne DG. 2021. Testing cophylogeny between coral reef invertebrates and their bacterial and archaeal symbionts. Mol Ecol 30:3768–3782. https:// doi.org/10.1111/mec.16006

O'Brien PA, Tan S, Frade PR, Robbins SJ, Engelberts JP, Bell SC, Vanwonterghem I, Miller DJ, Webster NS, Zhang G, Bourne DG. 2023. Validation of key sponge symbiont pathways using genome-centric metatranscriptomics. Environ Microbiol 25:3207–3224. https://doi.org/10.1111/1462-2920.16509

Slaby BM, Hackl T, Horn H, Bayer K, Hentschel U. 2017. Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization. ISME J 11:2465–2478. https://doi.org/10.1038/ ismej.2017.101

Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, Zhu Q, Bolzan M, Cumbo F, May U, Sanders JG, Zolfo M, Kopylova E, Pasolli E, Knight R, Mirarab S, Huttenhower C, Segata N. 2020. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun 11:2500. https://doi.org/10.1038/ s41467-020-16366-7

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455. https://doi.org/10.1093/ bioinformatics/btz305

Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP,

Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/ s41467-018-07641-9

Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36:2251–2252. https://doi.org/10.1093/bioinformatics/btz859

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.

27

Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon

JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P,

Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH. 2020. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol 16:60–68. https://doi.org/10.1038/s41589-019-0400-9

Parker-Nance S, Hilliar S, Waterworth S, Walmsley T, Dorrington R. 2019. New species in the sponge genus *Tsitsikamma* (Poecilosclerida,

Latrunculiidae) from South Africa. Zookeys 874:101–126. https://doi.org/

10.3897/zookeys.874.32268

van de Water J, Voolstra CR, Rottier C, Cocito S, Peirano A, Allemand D, Ferrier-Pagès C. 2018. Seasonal stability in the microbiomes of temperate gorgonians and the red coral *Corallium rubrum* across the Mediterranean sea. Microb Ecol 75:274–288. https://doi.org/10.1007/ s00248-017-1006-y

Tignat-Perrier R, van de Water JAJM, Allemand D, Ferrier-Pagès C. 2023. Holobiont responses of mesophotic precious red coral *Corallium rubrum* to thermal anomalies. Environ Microbiome 18:70. https://doi.org/10.

1186/s40793-023-00525-6

van de Water JAJM, Allemand D, Ferrier-Pagès C. 2024. Bacterial symbionts of the precious coral *Corallium rubrum* are differentially distributed across colony-specific compartments and differ among colormorphs. Environ Microbiol Rep 16:e13236. https://doi.org/10.1111/

1758-2229.13236

Perrin J, Vielzeuf D, Ricolleau A, Dallaporta H, Valton S, Floquet N. 2015. Block-by-block and layer-by-layer growth modes in coral skeletons. Am Min 100:681–695. https://doi.org/10.2138/am-2015-4990

Giordano B, Bramanti L, Perrin J, Kahramanoğulları O, Vielzeuf D. 2023. Early stages of development in Mediterranean red coral (*Corallium rubrum*): the key role of sclerites. Front Mar Sci 10. https://doi.org/10. 3389/fmars.2023.1052854

Uriz MJ, Agell G, Blanquer A, Turon X, Casamayor EO. 2012. Endosymbiotic calcifying bacteria: a new cue to the origin of calcification in metazoa? Evolution 66:2993–2999. https://doi.org/10.1111/j.1558-5646.

2012.01676.x

Garate L, Sureda J, Agell G, Uriz MJ. 2017. Endosymbiotic calcifying bacteria across sponge species and oceans. Sci Rep 7:43674. https://doi.org/10.1038/srep43674

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584. https://doi.org/10.7717/peerj.2584

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K,

Madden TL. 2009. BLAST+: architecture and applications. BMC

Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Maglott D, Ostell J, Pruitt KD, Tatusova T. 2005. Entrez Gene: genecentered information at NCBI. Nucleic Acids Res 33:D54–D58. https://doi.org/10.1093/nar/gki031

Edgar RC. 2022. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun 13:6968. https://doi.org/10.1038/s41467-022-34630-w

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120

Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrlund TS, Bushuiev R, Murray KJ, Hoffmann N, Lu M, et al. 2023. Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nat Biotechnol 41:447–449. https://doi.org/10.1038/s41587-023-01690-2

Langfelder P, Horvath S. 2012. Fast R functions for robust correlations and hierarchical cluste. J Stat Softw 46:i11.

Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. https://doi.org/10.1093/nar/gkab301

Robbins SJ, Song W, Engelberts JP, Glasl B, Slaby BM, Boyd J, Marangon E, Botté ES, Laffy P, Thomas T, Webster NS. 2021. A genomic view of the microbiome of coral reef demosponges. ISME J 15:1641–1654. https:// doi.org/10.1038/s41396-020-00876-9

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114

Seemann T. 2013. Barrnap 0.9: rapid ribosomal RNA prediction. https:// github.com/tseemann/barrnap.

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/ btu153

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Gosselin S, Fullmer MS, Feng Y, Gogarten JP. 2022. Improving phylogenies based on average nucleotide identity, incorporating saturation correction and nonparametric bootstrap support. Syst Biol 71:396–409. https://doi.org/10.1093/sysbio/syab060

Altenhoff AM, Levy J, Zarowiecki M, Tomiczek B, Warwick Vesztrocy A,

Dalquen DA, Müller S, Telford MJ, Glover NM, Dylus D, Dessimoz C. 2019. OMA standalone: orthology inference among public and custom genomes and transcriptomes. Genome Res 29:1152–1163. https://doi. org/10.1101/gr.243212.118

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. Trends Genet 16:276–277. https://doi.org/10.1016/s0168-9525(00)02024-2

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res 34:W609–W612. https://doi.org/10.1093/nar/gkl315 87. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591. https://doi.org/10.1093/molbev/msm088

McInnes L, Healy J, Saul N, Großberger L. 2018. UMAP: uniform manifold approximation and projection. JOSS 3:861. https://doi.org/10.21105/joss.00861

De Cáceres M, Legendre P. 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90:3566–3574. https://doi.org/10.1890/08-1823.1

Gilchrist CLM, Chooi Y-H. 2021. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37:2473– 2475. https://doi.org/10.1093/bioinformatics/btab007