Analysis of Actinobacterial Biodiversity in Marine Sediment from Gericke's Point (South Africa) and Screening of Isolates for Novel Antimycobacterial Compounds

by

Ashleigh Dreyer

Thesis submitted in fulfilment of the requirements for the degree of Master of Science in the Department of Molecular and Cell Biology, Faculty of Science, University of Cape Town, South Africa

July 2022

Plagiarism Declaration

I know the meaning of plagiarism and declare that all of the work in this thesis, save for that which is properly acknowledged, is my own.

Signature Signed by candidate

Date 20 July 2022

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The completion of this undertaking would not have been possible without the guidance, patience and support of my supervisor, Dr Paul Meyers. Thank you for sharing your extensive knowledge and expertise. Thank you for always going the extra mile for your students. You have made my postgraduate experience at UCT much easier and I am incredibly grateful.

Next, I would like to thank Dr Marilize Le Roes-Hill for providing me the opportunity to be a part of such an exciting research project and for always taking the time out to answer any questions I had. Your contribution as well as the contributions of the other members of the BTB research group to this thesis is greatly appreciated. I would especially like to thank Alaric Prins for his part in the metagenomic investigation.

Next, I would like to express my gratitude to Dr Daniel Watson for his time and expertise. Your contribution has made this research more impactful and is also greatly appreciated.

Special thanks go to my lab mate Sam Hewat who has been with me since the beginning of my postgraduate journey. Thank you for making everything "festive" and always reminding me to take care of myself. Many thanks also go to my past lab mates, Adeebah Rakiep and Caroline Robertson, who have also played a special part in this process. My lab experience would not have been the same without all of you.

Thank you to my parents and my siblings for their continued prayers and support over the years. I would especially like to thank my sister, Megan Dreyer, for sharing her own scientific knowledge and suggestions for PCR troubleshooting.

Lastly, I would like to thank the Almighty, without whom I would not be where I am today.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF

LIST OF ABBREVIATIONS

A	–	adenine (DNA base)
ANI	–	Average Nucleotide Identity
atpD	–	ATP synthase beta-subunit
BGC	–	biosynthetic gene cluster
BLAST	–	Basic Local Alignment Search Tool
bp	–	base pair (DNA)
BSA	–	bovine serum albumin
BTB	–	Biocatalysis and Technical Biology
C	–	cytosine (DNA base)
Ca	–	*candidatus*
CHPC	–	Centre for High Performance Computing
cm	–	centimetre
comb. nov.	–	*combination nova* (meaning “new combination”)
Da	–	Dalton
dDDH	–	digital DNA-DNA hybridisation
DDH	–	DNA-DNA hybridisation
DNA	–	deoxyribonucleic acid
dNTP	–	deoxyribonucleotide triphosphate
eDNA	–	environmental deoxyribonucleic acid
et al.	–	*et alia* (meaning “and others”)
EtAc	–	ethyl acetate
g	–	gram
G	–	guanine (DNA base)
GBDP	–	Genome Blast Distance Phylogeny
gDNA	–	genomic deoxyribonucleic acid
GGDC	–	Genome-to-Genome Distance Calculator
GNPS	–	Global Natural Products Social Molecular Networking database
gyrB	–	DNA gyrase subunit B gene
HGT	–	horizontal gene transfer
HPLC-MS	–	High Performance Liquid Chromatography-Mass Spectrometry
HSP	–	high-scoring segment pair

i.e.	–	*id est* (meaning “it is” or “that is”)
IDA	–	information dependent acquisition
ISP	–	International *Streptomyces* project
kb	–	kilobase
L	–	litres
LB	–	luria bertani (agar/broth)
MAG	–	metagenome-assembled genome
MB	–	Middlebrook 7H9 (agar/broth)
MDR-TB	–	multidrug-resistant TB
MEGA	–	Molecular Evolutionary Genetics Analysis
MeOH	–	methanol
mg	–	milligram
min	–	minute
ml	–	millilitre
MLSA	–	Multilocus sequence analysis
mM	–	millimolar
m	–	metre
mm^2	–	square millimetre
MRSA	–	methicillin-resistant *Staphylococcus aureus*
Mtb	–	*Mycobacterium tuberculosis*
ng	–	nanograms
NGS	–	next-generation sequencing
nm	–	nanometre
nr/nt	–	nucleotide collection database
NRPS	–	non-ribosomal peptide synthetase
nt	–	nucleotide (DNA)
OD	–	optical density
OGRI	–	overall genome relatedness index
OTU	–	Operational Taxonomic Unit
PCR	–	polymerase chain reaction
pH	–	potential Hydrogen
PKS	–	polyketide synthetase
R	–	reverse

recA	–	recombinase A gene
rRNA	–	ribosomal ribonucleic acid
rpm	–	revolutions per minute
rpoB	–	RNA polymerase beta-subunit
s	–	seconds
SDR-TB	–	single drug-resistant TB
SM	–	secondary metabolite
ssh	–	Secure shell protocol
subsp.	–	subspecies
T	–	thymine (DNA base)
TB	–	tuberculosis
TDR-TB	–	totally drug-resistant TB
TE	–	Tris-EDTA
trpB	–	tryptophan synthase beta-subunit
U	–	units
UV	–	ultra-violet
v/v	–	volume for volume
w/v	–	weight for volume
wgs	–	whole-genome shotgun contig database
WHO	–	World Health Organisation
XDR-TB	–	extensively drug-resistant TB
ZOI	–	zone of inhibition
°C	–	degrees Celcius
µl	–	microlitre
µg	–	microgram
µM	–	micromolar

ABSTRACT

Thirty-three (33) presumptive actinobacterial strains were isolated using traditional culture-based techniques from sediment taken from marine habitats (a subtidal zone, a rock pool and a beach area) at Gericke's Point (Garden Route National Park, Sedgefield, South Africa). Twenty-seven (27) of the 33 presumptive actinobacterial isolates were identified to the genus level: 26 *Streptomyces* strains and one *Nocardia* strain. The partial 16S-rRNA gene sequences obtained for each confirmed actinobacterial isolate were used to determine their phylogenetic positions within their respective genera. Further investigation of specific isolates was done utilising the *gyrB* gene to determine whether these isolates are clones. Metagenomic data generated from next-generation sequencing of 16S-rRNA amplicons were used to reveal the actinobacterial biodiversity of the Gericke's Point sediment that was not seen in the culture-dependent part of this study. A total of 1 541 544 actinobacterial partial 16S-rRNA gene sequences were identified using the SILVA 16S-rRNA gene database. Actinobacteria that could not be assigned to a class or order made up ~41% of the total actinobacterial strains found in the Gericke's Point sediment samples. The rest of the identified actinobacterial strains belonged to the orders *Candidatus* Microtrichales (~45%), *Candidatus* Actinomarinales (~9%), *Propionibacteriales* (~3%) and other actinobacterial orders that each made up less than one percent (<1%) of the actinobacterial strains found in Gericke's Point. The other actinobacterial orders include *Bifidobacteriales*, *Euzebyales*, *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales*, *Pseudonocardiales*, *Streptomycetales* and *Streptosporangiales*. This is one of the first detections of *Frankiales* strains in a marine environment. The majority (99%) of actinobacterial strains identified at Gericke's Point could not be assigned to a known genus. This represents an abundance of novel actinobacterial diversity that has yet to be revealed. Multidrug-resistance in *Mycobacterium tuberculosis* is a global threat to public health which has increased the need for new antibiotics to treat tuberculosis. In this study, all confirmed actinobacterial isolates and two presumptive actinobacterial isolates (29 strains in total) were screened for antimycobacterial activity against the non-pathogenic *Mycobacterium aurum* strain A+ using a standard agar overlay method. To investigate their spectrum of antibiotic activity, all isolates were also screened for activity against *Escherichia coli* ATCC 25922 and *Staphylococcus aureus* ATCC 25923. Twenty-one (21) isolates (20 *Streptomyces* strains and one unidentified strain) displayed strong to very strong antimycobacterial activity (defined as

a zone of growth inhibition of over 2000 mm^2). In addition, two *Streptomyces* strains displayed strong to very strong activity against *S. aureus* ATCC 25923. Compounds that displayed strong antimycobacterial activity were analysed using High Performance Liquid Chromatography-Mass Spectrometry and resulting mass spectra were compared to those of known compounds within the Global Natural Products Social Molecular Networking (GNPS) database. Eighteen (18) strains produced compounds with no matches in the GNPS database indicating these compounds could be novel. One strain produced a potential analogue of abyssomicin L (a rare antibiotic). Overall, the results obtained in this study emphasize the potential of marine environments as a source of novel actinobacteria and novel bioactive compounds.

CHAPTER 1:

INTRODUCTION

TABLE OF CONTENTS

1.1 ACTINOBACTERIA

The actinobacteria represent one of the largest bacterial phyla. They are free-living, Gram-positive organisms that have a typically high G+C content (≥60%) in their DNA (Barka *et al.*, 2016). Among Gram-positive bacteria, actinobacteria display the greatest morphological differentiation (Li et al., 2016). They are unicellular and have a variety of morphological features including coccus and rod shapes, as well as filamentous forms in which branched hyphae develop into a mycelium (Law *et al.*, 2020). Filamentous actinobacteria are commonly referred to as actinomycetes. Most actinobacteria are aerobic and are widely distributed in terrestrial and aquatic habitats (Servin *et al.*, 2008). In these habitats, actinobacteria carry out a variety of key functions. They contribute to plant biomass degradation and play essential roles in the carbon and nitrogen cycles of natural ecosystems (Lewin *et al.*, 2016; Liu *et al.*, 2017). Actinobacteria are agents of bioremediation as they are able to metabolize chemical compounds such as pesticides and accumulate heavy metals from soil (Alvarez *et al.*, 2017). They are able to act as probiotics providing health benefits to both animals and plants (Das *et al.*, 2008; Menendez *et al.*, 2019). Actinobacteria also play a significant role in agriculture as they are able to inhibit the growth of several plant pathogens (Anandan *et al.*, 2016; Bhatti *et al.*, 2017) and promote the growth of plants under stress through endophytic association (Salwan and Sharma, 2020). As producers of enzymes and clinically significant secondary metabolites, actinobacteria are considered of major importance for biotechnology and medicine. They account for ~70% of the naturally derived compounds that are currently in clinical use (Manivasagan *et al.*, 2014). In particular, marine actinobacteria have been identified as an important source of novel bioactive natural products.

1.1.1 Marine actinobacteria

The marine environment covers more than 70% of the Earth's surface and has been proven to be a rich source of microbial biodiversity (Mineta and Gojobori, 2016; Zhao, 2011). Marine microbial communities account for more than 80% of life on Earth and marine actinobacteria account for 10-40% of the bacterial diversity within these communities (Li and Qin, 2005; Anandan *et al.*, 2016). Actinobacteria have been shown to occur in various different marine habitats including marine sediment, the deep sea floor, hydrothermal vents, mangroves, coastal and intertidal regions (Manivasagan et al., 2013), as well as in marine algae and in

certain marine invertebrates such as sponges and corals (Mahmoud and Kalendar, 2016; Girão *et al.*, 2019; Liu *et al.*, 2019a). These habitats are exposed to unique stresses such as extreme pressures, temperatures, pH and salinity levels and even high levels of pollution (Barone *et al.*, 2014). Due to the unique environments they inhabit, it is assumed that marine actinobacteria have been exposed to extensive evolutionary forces and, as a result, have developed unique biochemical processes. Marine actinobacteria are thus considered to be a promising source of novel bioactive compounds. This fact, coupled with the deceleration of natural product research in terrestrial habitats, has led to an increasing interest in the exploration of marine actinobacteria (Li & Qin, 2005). As marine-derived natural products have become a more prominent area of research, more marine habitats are being explored all over the world (Blunt *et al.*, 2007). Despite this, it is still considered a largely untapped source of biodiversity and secondary metabolites. South Africa, in particular, has a coastline of approximately 2800 km yet very little is known about the biodiversity of actinobacterial strains in South African marine environments.

It was previously believed that marine actinobacterial isolates were from spores washed in from terrestrial environments and that these isolates were able to grow under, but did not require, the conditions associated with the marine environment (Goodfellow and Haynes, 1984; Ward and Bora, 2006). While this is the case for a significant number of actinobacterial strains isolated from the marine environment, particularly when sampling takes place in near-shore habitats under terrestrial influence, evidence has accumulated in recent years that demonstrate the existence of indigenous marine actinobacteria. Bull *et al*. (2005) described the criteria for defining indigenous marine bacteria as having the ability to grow optimally and be active at *in situ* salinity, pressure, temperature, and nutrient concentrations, while also displaying metabolic signatures unique to marine-adapted organisms. These bacteria are most likely to be found in deep-sea sediment. *Rhodococcus marinonascens* was one of the first actinobacterial species discovered to be obligately marine as it is unable to grow in freshwater media (Helmke and Weyland, 1984). Since then, Mincer *et al.* (2002) discovered the first major taxon of indigenous marine actinobacteria and showed it to be widely distributed across multiple different ocean habitats. This taxon was fully recognized with the establishment of the genus *Salinispora* which forms part of the *Micromonosporaceae* family (Maldonado *et al.*, 2005a). This genus is currently comprised of only 9 species with validly published names (https://lpsn.dsmz.de/genus/salinispora), suggesting that there is a low diversity of culturable marine actinobacteria. However, Maldonado *et al.* (2005b) suggested

that this is not due to actual low species diversity, but rather under sampling and inadequate isolation techniques. Therefore, the use of informed selective isolation procedures can aid in the isolation of members of novel taxa indigenous to the marine environment. Recent data suggest that other factors such as temperature and season should also be considered when sampling from marine habitats (Ladau *et al.*, 2013).

1.2 ACTINOBACTERIAL TAXONOMY

Bacterial taxonomy involves three key components; rank-based **classification** of organisms into taxa (groups) based on established criteria, **identification** based on previously classified organisms and the **naming** of organisms based on a set of formal rules laid out in the International Code of Nomenclature of Prokaryotes (Parker *et al.*, 2019). The full hierarchal taxonomic ranks from lowest to highest are as follows: species, genus, family, order, class, phylum and domain. Currently, the phylum *Actinobacteria* is made up of 6 classes, 46 orders, 79 families and a total of 425 genera with validly published names (Salam *et al.*, 2020). Initially, classification was solely based on cell morphology and other phenotypic markers (Lehmann & Neumann, 1896); however, this proved to be an unreliable basis for drawing taxonomic conclusions (Prescott *et al.*, 2008). Hence, modern bacterial taxonomy has adopted a polyphasic approach to classification which integrates data from phenotypic, genotypic and phylogenetic analyses (Colwell, 1970). Polyphasic taxonomy has helped establish stricter classification criteria which has led to marked improvements in the delineation of taxa at all levels (Zhi *et al.*, 2012). In addition to this, increased availability of genomic information has led to the development of genome-based methods that provide novel insights into the effects of horizontal gene transfer (HGT), gene duplication and gene decay. This has further increased the overall resolution of bacterial taxonomy (Coenye *et al.*, 2005). The interplay between data sets from traditional and whole-genome analyses provides a sound basis for describing the diversity of bacteria.

1.2.1 Phenotypic analyses

Although classification can no longer be based solely on phenotypic markers, it is still important to conduct phenotypic analyses in order to classify organisms fully. These include observing cell morphology (shape, flagella, Gram-staining etc.) and chemotaxonomic analysis. Chemotaxonomic analysis involves classifying organisms based on the chemical

constituents of their cells. This includes collecting information about the composition of sugars and amino acids in the cell wall and fatty acids, menaquinones and phospholipids in the cell membrane. As with other phenotypic analyses, chemotaxonomic methods can be laborious with some techniques also resulting in poor analytical resolution (Sutcliffe *et al.*, 2012). It was also stated in Tindall *et al.* (2010) that few scientists actually have the necessary skills to interpret chemotaxonomic data. For these reasons, it has recently been suggested that the practical value of chemotaxonomic data is non-existent for species delineation, when taking into consideration the higher taxonomic resolution of genome-based methods (Vandamme and Sutcliffe, 2021). These authors propose that chemotaxonomic data could still be useful in studies at higher taxonomic ranks only if current methodology is improved to yield more informative results. Generally, phenotypic analyses should also include environmental data to provide information about the ecology of a species. This is due to the fact that strains of the same species will exhibit fewer conserved genes than expected when inhabiting different ecological niches and similarly, due to the conditional nature of gene expression, strains of the same species might show different phenotypic characters in different environmental conditions (Konstantinidis and Tiedje, 2005; Prakash *et al.*, 2007).

1.2.2 Phylogenetic analyses

Phylogenetic analyses (comparative sequence analyses) were kick-started by the introduction of ribosomal RNA (rRNA) sequences into bacterial taxonomy (Woese and Fox, 1977). The 16S-rRNA gene in particular, is functionally stable and highly conserved in most bacteria; thus, it was established as a standard molecular marker to infer evolutionary relationships. However, it has a few shortcomings. The 16S-rRNA gene has a slow evolutionary rate which, while allowing comparison between very divergent bacteria, complicates comparison between closely related strains (Santos and Ochman, 2004). In addition to this, multiple copies of the 16S-rRNA gene are present in most bacterial phyla (Sun *et al.*, 2013) and differences between sequences within a single genome (intragenomic heterogeneity) could influence tree topology and overall phylogenetic resolution. It has been shown that using genes such as *rpoB*, *gyrB* and other single-copy genes for phylogenetic analyses can provide better resolution between closely related organisms (Case *et al.*, 2007).

It is also known that a tree constructed using a single genetic locus will not necessarily agree with a tree that represents the actual evolutionary pathway of an organism (Pamilo and Nei, 1988; Liu *et al.*, 2019b) Therefore, as there are now many bacterial whole-genome sequences

available, phylogenomic analysis presents itself as a more viable method to infer phylogenetic relationships between organisms. Unlike 16S-rRNA analysis, phylogenomic methods use a set of highly conserved bacterial genes rather than a single gene (Staley, 2006). These genes are selected for phylogenetic analysis by Multilocus Sequence Analysis (MLSA). MLSA is an approach used to infer phylogenetic relationships by comparing the concatenated sequences of multiple house-keeping genes between closely related strains (Hanage *et al*., 2006). The phylogenetic analyses can be supplemented with genomic information based on analyses of gene sequences, gene arrangement and gene expression. Phylogenomic analyses overcome the shortcomings of 16S-rRNA and other single-gene classifications to provide a more accurate representation of bacterial phylogeny (Takahashi *et al.*, 2009). Phylogenomic trees are currently being used to help recognize and rectify poorly classified taxa. An example of this was the phylogenomic revision of the actinobacterial family *Streptosporangiaceae*, which resulted in the inclusion of the genus *Sinosporangium* in the family and the reclassification of *Desertiactinospora gelatinilytica* as *Spongiactinospora gelatinilytica* comb. nov. (Ay *et al.*, 2020).

1.2.3 Genotypic analyses

1.2.3.1 DNA-DNA hybridization

In addition to aiding phylogenetic analyses, sequences obtained from 16S-rRNA gene amplification and sequencing are primarily used to identify bacteria to the genus level and identify potential novel species. However, due to its highly conserved nature, resolution based on 16S-rRNA gene analysis at the species level is often not possible (Fox *et al.*, 1992). For this reason, it has been suggested that when 16S-rRNA gene sequence similarities are high (>97%), it should be mandatory to include DNA/DNA hybridization (DDH) in the overall analysis in order to determine whether the two strains being compared belong to different genomic species (Meier-Kolthoff *et al.*, 2013). DDH is a measure of sequence similarity based on the melting and association of genomic DNA fragments of about 1kb in length (Stackebrandt and Goebel, 1994). G-C content (%) in bacterial DNA affects its melting temperature and, therefore, affects results obtained from DDH. The variation in G-C content is not more than 3% within a well-defined species and not more than 10% within a well-defined genus (Vandamme *et al.*, 1996). A standard of 70% or greater DNA-DNA similarity by DDH and less than a 5% difference in DNA melting temperature were set as the criteria for assigning two prokaryotic strains to the same species (Wayne *et al*., 1987). Even

though DDH was at one point considered a 'gold standard' for species delineation (Ramasamy *et al.*, 2014), many taxonomists believed a sequencing-based method should be developed that could replace DDH due to the limitations associated with this method. These limitations include high experimental error, problems with reproducibility and the fact that the technique does not lend itself to the generation of a cumulative database (Stackebrandt, 2003).

Digital DNA-DNA Hybridisation (dDDH) has since largely replaced traditional DDH. This method imitates DDH by using computer programs to fragment genome sequences *in silico* and identify high-scoring segment pairs (HSPs) between two genomes. Using these segment pairs, genome-to-genome distances are calculated using the Genome Blast Distance Phylogeny (GBDP) strategy described by Henz *et al.* (2005). The Genome-To-Genome Distance Calculator (GGDC) (https://ggdc.dsmz.de/) is a web server which implements all aspects of this approach (Auch *et al.*, 2010a). Regarding species boundaries, the same DNA-DNA similarity threshold as in traditional DDH (i.e. 70%) is used in dDDH analyses for assigning two strains to the same genomic species (Auch *et al.*, 2010b).

1.2.3.2 Average Nucleotide Identity

Average Nucleotide Identity (ANI) is another computer-based method used to compare the genomes of two bacterial strains to determine whether they belong to the same species. The genome of a query strain is divided into fragments *in silico* (similar to dDDH) and an individual nucleotide identity is calculated between each fragment sequence and the sequence of an intact reference genome (Lee *et al.*, 2016). Average nucleotide identity is the mean of these nucleotide identity values. Each genome in a pair serves as the query and reference sequence in separate analyses. Alternatively, both genomes can be fragmented simultaneously which enables a single comparative step (Palmer *et al.*, 2020). When ANI was first described, strains that exhibited >94% ANI were classified as belonging to the same species (Konstantinidis and Tiedje, 2005). This value corresponded to the 70% DDH standard. The threshold has since been increased to 95-96% (Richter and Rosselló-Móra, 2009). It has also been recommended that this threshold be tuned based on the strains being investigated (Ciufo *et al.*, 2018; Palmer *et al.*, 2020). For example, it can be lowered when investigating species observed to have relatively higher intraspecies diversity and conversely, increased when investigating species observed to have higher genome sequence similarity (Ciufo *et al.*, 2018).

Chun *et al.* (2018) proposed that a combination of dDDH and ANI values, previously referred to as an overall genome relatedness index (OGRI) by Chun and Rainey (2014), must be calculated against the type strains of all phylogenetically related species in order to robustly identify an organism as a known species or to propose that it represents a new one.

1.3 BIOACTIVE COMPOUNDS

Bioactive compounds are products of biosynthetic pathways and are produced by microbial organisms in order to give them a selective advantage within their respective ecological niches. These pathways are encoded by biosynthetic gene clusters (BGCs) and the bioactive compounds produced are referred to as secondary metabolites (SMs). This is because organisms do not require SMs to grow. The most dominant pathways for SM production are the non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) biosynthetic pathways (Albarano *et al.*, 2020). After the discovery of penicillin in 1928, the concept of synthesising and utilising bioactive compounds in clinical settings became a fully-fledged area of research. Since then, the large majority of compounds that have been characterised have been produced by actinobacteria (Jose *et al.*, 2021). Actinobacteria are best known as prolific producers of antibiotics, but have also been shown to produce other antimicrobials, anticholesterol compounds, antihelmintics, anti-plasmodial compounds, immunosuppressants, anticancer agents (Bhatti *et al.*, 2017; Mahapatra *et al.*, 2020) and other biotechnologically significant compounds (Figure 1.1). Most of these compounds (~80%) have been produced by members of the actinobacterial genus *Streptomyces* (Bérdy, 2012).

1.3.1 Dereplication and Genome mining

Traditionally, bioactive compound discovery involved screening for bioactivity, isolating active compounds and manual structure elucidation. This is not only time-consuming but also increases the chance of re-discovering known compounds (Mohamed *et al.*, 2016). Consequently, dereplication has become an important part of novel bioactive compound discovery. Dereplication is the rapid identification of known compounds in an extract to allow researchers to focus their efforts toward novel discoveries. This preliminary screening does not require the purification or structure elucidation of compounds of interest. The most commonly used technique for dereplication is High Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS). This technique allows one to obtain the individual mass

spectra of compounds in a crude extract, which can then be compared to spectra of known compounds in various databases to identify potential novel compounds.

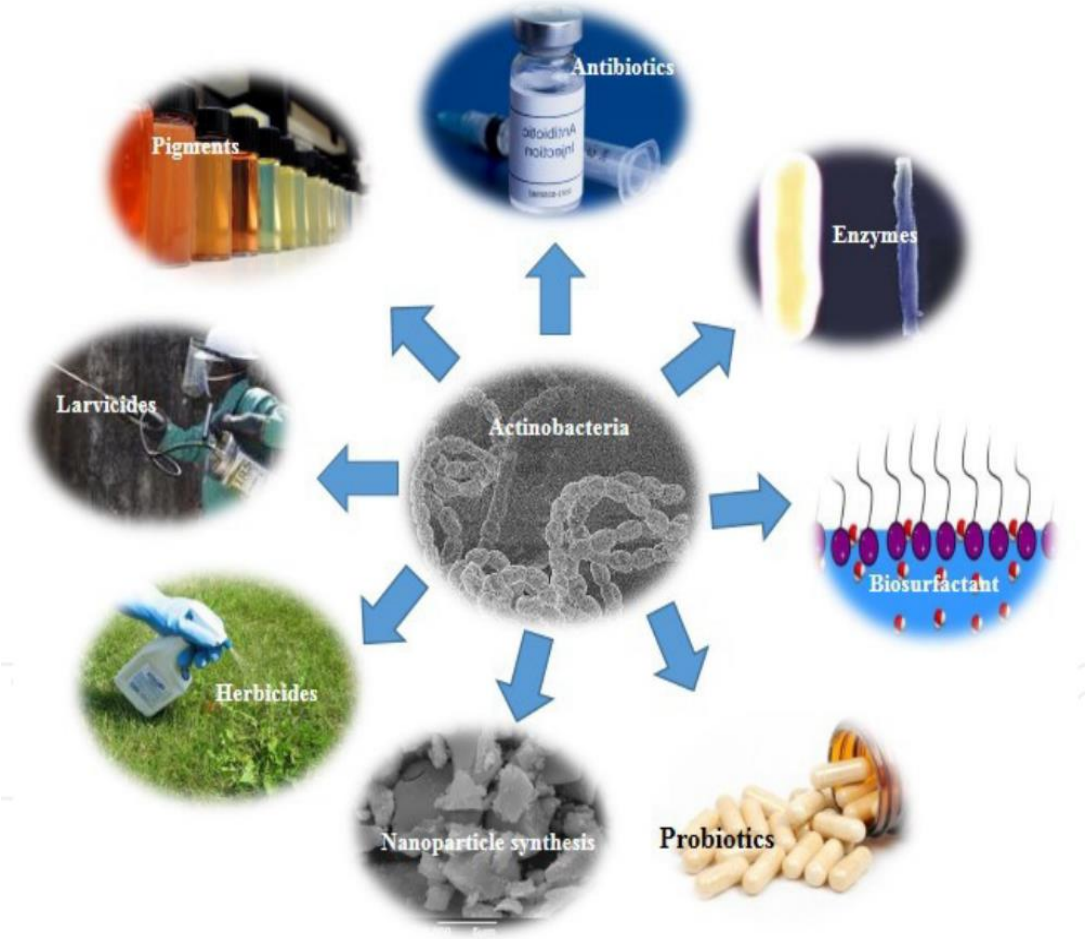

Figure 1.1. Biotechnological applications of actinobacteria. (Anandan *et al.*, 2016)

In a similar fashion, genome mining has also improved the efficiency with which novel bioactive compounds are discovered. Genome mining is an approach, made possible by the increased availability of whole-genome information, that uses various bioinformatics tools to predict BGCs and the SMs they encode (Kenshole *et al.*, 2021). This, again, avoids re-discovery of known compounds, while also allowing researchers to access chemical and structural information to rapidly characterize potentially novel compounds (Albarano *et al.*, 2020). Genome mining began in the genome of *Streptomyces coelicolor* A3(2) (the first available *Streptomyces* genome) where it was revealed that only ~18% of the secondary metabolites that *S. coelicolor* had the potential to produce have ever been detected during cultivation under standard laboratory conditions (Bentley *et al.*, 2002). This is due to 'silent' BGCs which encode SMs that cannot be expressed (or are expressed at very low

concentrations) using traditional cultivation techniques. Since genomic information can also provide a better understanding of genetic regulation and mechanisms of biosynthesis, this knowledge can be used to engineer expression of SMs from otherwise 'silent' BGCs to improve production of both novel and known bioactive compounds (Ren *et al.*, 2017). There are still many limitations associated with this concept and it remains an active area of research. These limitations include those associated with heterologous expression such as the lack of suitable heterologous hosts and the potential toxicity of SMs when expressed in these hosts (Singh *et al.*, 2021).

Genomes of marine actinobacteria have been shown to contain an abundance of novel BGCs (Dhakal *et al.*, 2017), which, in recent years, has resulted in the production of a variety of novel compounds (Fiedler *et al.*, 2005; Jensen *et al.*, 2005; Manivasagan *et al.*, 2014; Mahapatra *et al.*, 2020). One hundred and fifty (150) metabolites have been identified from members of the genus *Salinispora* alone, allowing it to serve as a prime example of the potential of indigenous marine actinobacteria to be a source of bioactive compounds (Kim *et al.*, 2020). Among these metabolites are novel antibacterial compounds such as arenimycin (Asolkar *et al.*, 2010), novel anti-inflammatory compounds such as arenamide A (Asolkar *et al.*, 2009) and novel anticancer compounds such as salinosporamide A (Feling *et al.*, 2003). Notably, salinosporamide A, a structurally novel metabolite produced by *Salinispora tropica*, is currently in phase III clinical trials. If approved, it will be the first natural product derived from a cultured marine actinobacterium to achieve such clinical relevance (Jensen, 2022). Members of the genus *Salinispora* have also been described as model organisms for SM discovery (Jensen *et al.*, 2015). For example, salinilactam A (a novel compound also isolated from *S. tropica*) has been used as a model to demonstrate the potential of genome mining to elucidate chemical structure (Udwary *et al.*, 2007).

1.3.2 Antibiotics

Actinomycin was the first antibiotic isolated from an actinomycete. It was isolated in 1940 by Selman Waksman from *Actinomyces antibioticus* (now called *Streptomyces antibioticus*). This event ushered in the so-called "golden era" of antibiotic discovery (1950-1960) as most of the known classes of antibiotics were discovered during this time (Figure 1.2). As mentioned previously, most of these compounds were isolated from *Streptomyces* species but non-*Streptomyces* genera (termed 'rare actinobacteria') have also proved to be good antibiotic producers (Jose and Jebakumar, 2013). The "golden era" ended when standard

screening techniques were no longer resulting in the discovery of novel antibacterial compounds and rediscovery of known compounds became more frequent (Bérdy, 2012). Frequent rediscovery was also due, in part, to the fact that standard culturing techniques were resulting in the re-isolation of known bacterial species. These issues, along with its associated costs, caused a decline in this field of research. However, the increasing issue of antibiotic resistance has forced a renaissance in antibiotic discovery.

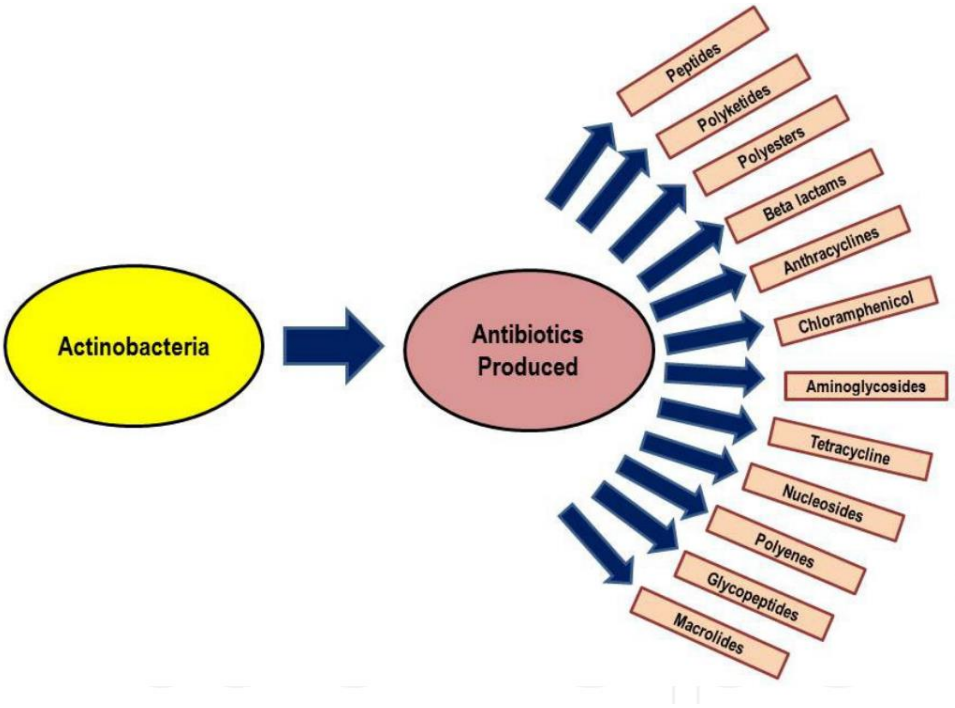

Figure 1.2. Classes of antibiotics produced by actinobacteria (Anandan et al., 2016)

1.3.2.1 Antibiotic resistance

Antibiotics target different bacterial structures and functions in order to selectively inhibit bacterial growth (Grasso *et al*., 2016). Antibiotic resistance occurs when bacteria acquire genes (HGT) or develop mutations that counteract this inhibition. Antibiotic resistance is usually caused by overuse and misuse of certain antibiotics (De Simeis and Serra, 2021). This phenomenon has caused the emergence of resistant pathogenic bacteria such as methicillin-resistant *Staphylococcus aureus* (MRSA) and fluoroquinolone-resistant *Pseudomonas aeruginosa* (Zhu *et al*. 2014). More significantly, it has caused the emergence of multi-drug resistant (MDR) pathogenic bacteria. The most well-known example is MDR tuberculosis

(MDR-TB) which, according to the World Health Organisation (WHO), remains a global public health crisis (https://www.who.int/tb/data/en/). TB is caused by the pathogenic actinobacterium *Mycobacterium tuberculosis (Mtb).* Due to antibiotic resistance, TB is currently treated using a combination of drugs (Singh and Chibale, 2021). This regimen has given rise to even more resistant forms of *Mtb*, namely, extensively drug-resistant *Mtb* (XDR-TB) and totally drug resistant *Mtb* (TDR-TB) (Jacobson *et al.*, 2010; Akbar *et al.*, 2013*)*. Increasing resistance creates an urgency to discover novel antibiotics as, without them, TB and other bacterial infections could become untreatable. It is important to note that, due to their cytotoxic nature, some antibacterial compounds are actually not useful in a clinical sense (though they are sometimes used in cancer treatment as antitumor compounds) (Missailidis, 2008). Hence, it is important to find novel antibiotics with varying mechanisms of action to increase the chance of discovering antibiotics that reach clinical relevance.

Genomic mining analyses have revealed that less than 10% of the genetic potential of antibiotic producers is currently being used (Mast and Stegmann, 2019). Thus, researchers are hopeful that increased selective isolation and screening of bacteria from under-explored habitats (such as the marine environment), when used in conjunction with selective cultivation strategies, insight into biosynthetic pathways and exploitation of triggers of antibiotic production, will lead to a new era of antibiotic drug discovery *(*Zhu *et al.*, 2014; Mast and Stegmann, 2019).

1.4 METAGENOMICS

The majority of microbes (>99%) are, as yet, unculturable using standard laboratory techniques (Schloss and Handelsman, 2003). It is particularly hard to culture indigenous marine bacteria (Li and Qin, 2005). This unculturable diversity serves as a very promising source of novel enzymes and bioactive compounds. Metagenomics is an approach that allows one to study the full range of microbial diversity regardless of whether the organisms being studied can be cultured. It involves the genetic analysis of a population of microorganisms by isolating DNA directly from an environmental sample (eDNA). Traditionally, eDNA is amplified, cloned into a culturable bacterium and sequenced based on a single taxonomic marker gene (Handelsman, 2005). However, this approach is labour intensive and most studies only analyse fewer than 100 clones per sample (Sanschagrin and Yergeau, 2014).

This results in an underestimation of the diversity present in a sample, hence why library preparation using next-generation sequencing (NGS) has become a more attractive option.

NGS generates large amounts of sequence data in a shorter period of time due to its ability to sequence in parallel i.e. multiple reads sequenced simultaneously (Behjati and Tarpey, 2013). It also evades the need to insert the DNA into a culturable host as isolated eDNA can be directly processed and collected in libraries for NGS-based applications (Sanschagrin and Yergeau, 2014). Metagenomics using NGS can be separated into 2 categories: shotgun and amplicon-based, though it has been proposed that the latter be referred to as 'metataxonomics' (Marchesi and Ravel, 2015). Amplicon-based metagenomics targets a single marker gene whereas shotgun metagenomics produces fragmented sequences (reads) that can be used to reconstruct metagenome-assembled genomes (MAGs). The assembly of MAGs involves generating consensus regions of DNA (contigs) and clustering them into individual groups which represent full or partial microbial genomes (binning) (Hugerth *et al.*, 2015; Yue *et al.*, 2020). Setubal (2021) stated that we are still in the early stages of tool and resource development to support the reconstruction and analysis of MAGs. However, once assembled, genome mining of these genomes will make it possible to assess potentially novel bioactive compounds from otherwise unculturable bacterial species while also providing insight into the functional and metabolic dynamics of a microbial community. It is important to note that most inferences based on metagenomic analyses are limited by their reliance on existing reference databases containing sequences of known origin and gene function (Simon and Daniel, 2009).

1.5 AIMS

This study forms part of a larger project involving two other research teams: Dr Marilize le Roes- Hill's Biocatalysis and Technical Biology (BTB) Research Group, Institute of Biomedical and Microbial Biotechnology, Cape Peninsula University of Technology (CPUT), Cape Town and the laboratory of Assoc. Prof. Denzil Beukes, School of Pharmacy, University of the Western Cape (UWC), Cape Town. The project sought to determine the actinobacterial diversity associated with selected marine sediments while also assessing the ability of culturable actinobacterial strains from these marine environments to produce novel bioactive compounds and oxidative enzymes. It also sought to determine the community composition of selected marine sediments (metagenomics) while also collecting information

on gene expression and the biochemical processes in which these communities are involved (metatranscriptomics). Marine sediment samples were collected from rock pools, beach/dunes and subtidal zones across six sampling sites by members of the BTB group (Figure 1.3). These sites are situated along the western, south western and southern coastline of South Africa. The BTB Research Group also isolated metagenomic eDNA from each sampling site. This metagenomic eDNA was processed and submitted for 16S-rRNA gene amplification using actinobacteria-specific primer sets in order to investigate the community composition of all sampling sites.

The first aim of my study in particular, was to investigate the ***culturable actinobacterial diversity*** and community composition of marine sediment samples from Gericke's Point (Garden Route National Park) at Sedgefield, South Africa (Figure 1.3), which was one of the six sampling sites chosen for this project. In order to investigate the actinobacterial diversity, actinobacterial strains were selectively isolated using standard plate cultivation techniques and subsequently identified to the genus level using 16S-rRNA gene sequence analysis. These sequences were used to construct phylogenetic trees to analyse the phylogeny of these strains. Metagenomic data generated by the BTB research group were analysed by the author and used to complement the culture-dependent investigation.

An additional aim of my part of the study was to assess the ability of isolated marine actinobacterial strains to produce novel antibacterial compounds. The antibiotic spectrum of the isolates was investigated by screening for antibiotic activity against two Gram positive bacteria (*Mycobacterium aurum* strain A+ and *Staphylococcus aureus* strain ATCC 25923) and a Gram negative bacterium (*E. coli* strain ATCC 25922) using agar overlays on plates. Agar samples from plates containing antibiotic compounds with strong antimycobacterial activity were collected by the author and then analysed by Dr Daniel Watson (Division of Clinical Pharmacology, Faculty of Health Sciences, University of Cape Town) using HPLC-MS to identify possible novel compounds for further study.

The data collected in my part of the larger study should provide some insights into the biodiversity present in selected marine habitats at Gericke’s Point and the ability of isolated strains to produce novel bioactive compounds (antibiotics). Combined with the other aspects of the overall project, these data help expand our knowledge of South African marine environments.

Figure 1.3. Sampling sites included in the larger project (indicated by red crosses). These images were compiled using the CapeFarmMapper tool provided by the Western Cape Department of Agriculture. (https://gis.elsenburg.com/apps/cfm/). A) Samples were taken from one site in Garden Route National Park (GRNP), one site in Agulhas National Park (ANP), two sites in Table Mountain National Park (TMNP), and two sites in West Coast National Park (WCNP). B-C) Closer look at the site sampled in this study at Gericke's Point, Garden Route National Park (GRNP), South Africa.

1.6 REFERENCES

Akbar Velayati, A., Farnia, P. and Reza Masjedi, M. (2013) 'The totally drug resistant tuberculosis (TDR-TB)', *International Journal of Clinical and Experimental Medicine*, *6*(4), p. 307.

Albarano, L., Esposito, R., Ruocco, N. and Costantini, M. (2020) 'Genome mining as new challenge in natural products discovery', *Marine Drugs*, *18*(4), p. 199.

Asolkar, R.N., Freel, K.C., Jensen, P.R., Fenical, W., Kondratyuk, T.P., Park, E.J. and Pezzuto, J.M. (2009) 'Arenamides A− C, cytotoxic NFκB inhibitors from the marine actinomycete *Salinispora arenicola*'. *Journal of natural products*, *72*(3), pp. 396-402.

Asolkar, R.N., Kirkland, T.N., Jensen, P.R. and Fenical, W. (2010) 'Arenimycin, an antibiotic effective against rifampin-and methicillin-resistant *Staphylococcus aureus* from the marine actinomycete Salinispora arenicola'. *The Journal of antibiotics*, *63*(1), pp. 37-39.

Alvarez, A., Saez, J.M., Costa, J.S.D., Colin, V.L., Fuentes, M.S., Cuozzo, S.A., Benimeli, C.S., Polti, M.A. and Amoroso, M.J. **(2017)** 'Actinobacteria: Current research and perspectives for bioremediation of pesticides and heavy metals', *Chemosphere*, *166*, pp. 41–62.

Anandan, R., Dharumadurai, D. and Manogaran, G. P. (2016) 'An Introduction to Actinobacteria', *Actinobacteria - Basics and Biotechnological Applications*. InTechOpen.

Auch, A. F., Klenk, H.-P. and Göker, M. (2010a) 'Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs', *Standards in Genomic Sciences*, *2*(1), pp. 142–148.

Auch, A.F., von Jan, M., Klenk, H.P. and Göker, M. **(2010b)** 'Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison', *Standards in genomic sciences*, *2*(1), pp. 117–134.

Ay, H., Saygin, H. and Sahin, N. (2020) 'Phylogenomic revision of the family *Streptosporangiaceae*, reclassification of *Desertactinospora gelatinilytica* as *Spongiactinospora gelatinilytica* comb. Nov. and a taxonomic home for the genus *Sinosporangium* in the family *Streptosporangiaceae*', *International Journal of Systematic and Evolutionary Microbiology*, *70*(4), pp. 2569–2579.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.P., Clément, C., Ouhdouch, Y. and van Wezel, G.P. **(2016)** 'Taxonomy, Physiology, and Natural Products of Actinobacteria', *Microbiology and Molecular Biology Reviews*, *80*(1), pp. 1–43.

Barone, R., De Santi, C., Palma Esposito, F., Tedesco, P., Galati, F., Visone, M., Di Scala, A. and De Pascale, D. (2014) 'Marine metagenomics, a valuable tool for enzymes and bioactive compounds discovery', *Frontiers in Marine Science*, *1*, p. 38.

Behjati, S. and Tarpey, P. S. (2013) 'What is next generation sequencing?', *Archives of Disease in Childhood: Education and Practice Edition*, *98*(6), pp. 236–238.

Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D. and Bateman, A. (2002) 'Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)'. *Nature*, *417*(6885), pp.141-147.

Bérdy, J. (2012) 'Thoughts and facts about antibiotics: Where we are now and where we are heading', *The Journal of Antibiotics*, *65*, pp. 385–395.

Bhatti, A. A., Haq, S. and Bhat, R. A. (2017) 'Actinomycetes benefaction role in soil and plant health'. *Microbial pathogenesis*, *111*, pp.458-467.

Blunt, J.W., Copp, B.R., Hu, W.P., Munro, M.H., Northcote, P.T. and Prinsep, M.R. (2007) 'Marine natural products', *Natural Product Reports*, *24*(1), pp. 31–86.

Bull, A.T., Stach, J.E., Ward, A.C. and Goodfellow, M. (2005) 'Marine actinobacteria: perspectives, challenges, future directions', *Antonie van Leeuwenhoek*, *87*(1), pp. 65–79.

Case, R.J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W.F. and Kjelleberg, S. (2007) 'Use of 16S-rRNA and *rpoB* genes as molecular markers for microbial ecology studies', *Applied and Environmental Microbiology*, *73*(1), pp. 278–288.

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., da Costa, M.S., Rooney, A.P., Yi, H., Xu, X.W., De Meyer, S. and Trujillo, M.E. (2018) 'Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes', *International Journal of Systematic and Evolutionary Microbiology*, *68*(1), pp. 461–466.

Chun, J. and Rainey, F. A. (2014) 'Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea', *International Journal of Systematic and Evolutionary Microbiology*, *64*, pp. 316–324.

Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., Brover, S., Schoch, C.L., Kimchi, A. and DiCuccio, M. (2018) 'Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI', *International journal of systematic and evolutionary microbiology*, *68*(7), pp. 2386-2392

Coenye, T., Gevers, D., De Peer, Y.V., Vandamme, P. and Swings, J. (2005) 'Towards a prokaryotic genomic taxonomy', *FEMS Microbiology Reviews*, 29(2), pp. 147–167.

Colwell, R. R. (1970) 'Polyphasic taxonomy of the genus *Vibrio*: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species.', *Journal of bacteriology*. American Society for Microbiology Journals, *104*(1), pp. 410–33.

Das, S., Ward, L. R. and Burke, C. (2008) 'Prospects of using marine actinobacteria as probiotics in aquaculture', *Applied Microbiology and Biotechnology*, *81*(3), pp. 419–429.

De Simeis, D. and Serra, S. (2021) 'Actinomycetes: A Never-Ending Source of Bioactive Compounds—An Overview on Antibiotics Production', *Antibiotics*, *10*(5). p. 483.

Dhakal, D., Pokhrel, A.R., Shrestha, B. and Sohng, J.K. (2017) 'Marine Rare Actinobacteria: Isolation, Characterization, and Strategies for Harnessing Bioactive Compounds', *Frontiers in Microbiology*, *8*, p. 1106.

Feling, R.H., Buchanan, G.O., Mincer, T.J., Kauffman, C.A., Jensen, P.R. and Fenical, W. (2003) 'Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinospora*'. *Angewandte Chemie International Edition*, *42*(3), pp. 355-357

Fiedler, H.P., Bruntner, C., Bull, A.T., Ward, A.C., Goodfellow, M., Potterat, O., Puder, C. and Mihm, G. (2005) 'Marine actinomycetes as a source of novel secondary metabolites'. *Antonie Van Leeuwenhoek*, *87*(1), pp. 37-42

Fox, G. E., Wisotzkey, J. D. and Jurtshuk, P. (1992) 'How Close Is Close: 16S-rRNA Sequence Identity May Not Be Sufficient To Guarantee Species Identity', *International Journal of Systematic Bacteriology*, *42*(1), pp. 166–170.

Girão, M., Ribeiro, I., Ribeiro, T., Azevedo, I.C., Pereira, F., Urbatzka, R., Leão, P.N. and Carvalho, M.F. (2019) 'Actinobacteria isolated from *laminaria ochroleuca*: A source of new bioactive compounds', *Frontiers in Microbiology*, *10*, p. 683.

Goodfellow, M. and Haynes, J.A. (1984). 'Actinomycetes in marine sediments'. *Biological, biochemical and biomedical aspects of actinomycetes*, pp.453-472.

Grasso, L. Lo, Martino, D. C. and Alduina, R. (2016) 'Production of Antibacterial Compounds from Actinomycetes', *Actinobacteria - Basics and Biotechnological Applications*. *214*(11), pp. 272-282.

Hanage, W. P., Fraser, C. and Spratt, B. G. (2006) 'Sequences, sequence clusters and bacterial species', *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1475), pp. 1917–1927.

Handelsman, J. (2005) 'Metagenomics: Application of Genomics to Uncultured Microorganisms', *Microbiology and Molecular Biology Reviews*, *68*(4), pp. 669-685.

Helmke, E. and Weyland, H. (1984) '*Rhodococcus marinonascens* sp. nov., an actinomycete from the sea', *International Journal of Systematic Bacteriology*, *34*(2), pp. 127–138.

Henz, S.R., Huson, D.H., Auch, A.F., Nieselt-Struwe, K. and Schuster, S.C. (2005) 'Whole-genome prokaryotic phylogeny', *Bioinformatics*, *21*(10), pp. 2329–2335.

Hugerth, L.W., Larsson, J., Alneberg, J., Lindh, M.V., Legrand, C., Pinhassi, J. and Andersson, A.F. (2015) "Metagenome-assembled genomes uncover a global brackish microbiome'. *Genome biology*, *16*(1), pp. 1-18.

Jacobson, K.R., Tierney, D.B., Jeon, C.Y., Mitnick, C.D. and Murray, M.B. (2010) 'Treatment outcomes among patients with extensively drug-resistant tuberculosis: systematic review and meta-analysis', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *51*(1), pp. 6–14.

Jensen, P.R., Mincer, T.J., Williams, P.G. and Fenical, W. (2005) 'Marine actinomycete diversity and natural product discovery'. *Antonie Van Leeuwenhoek*, *87*(1), pp. 43-48.

Jensen, P.R., Moore, B.S. and Fenical, W. (2015) 'The marine actinomycete genus *Salinispora*: a model organism for secondary metabolite discovery'. *Natural product reports*, *32*(5), pp. 738-751

Jensen, P. R. (2022) 'Microbe Profile: *Salinispora tropica*: natural products and the evolution of a unique marine bacterium', *Microbiology*, *168*(4), p. 001163.

Jose, P. A., and Jebakumar, S. R. D. (2013). 'Non-streptomycete actinomycetes nourish the current microbial antibiotic drug discovery'. *Frontiers in microbiology*, *4*, p. 240

Jose, P. A., Maharshi, A. and Jha, B. (2021) 'Actinobacteria in natural products research: Progress and prospects', *Microbiological Research*, *246*, p. 126708.

Kenshole, E., Herisse, M., Michael, M. and Pidot, S.J. (2021) 'Natural product discovery through microbial genome mining', *Current Opinion in Chemical Biology*, *60*, pp. 47–54.

Kim, H., Kim, S., Kim, M., Lee, C., Yang, I. and Nam, S.J. (2020) 'Bioactive natural products from the genus *Salinospora*: a review'. *Archives of Pharmacal Research*, *43*(12), pp. 1230-1258.

Konstantinidis, K. T. and Tiedje, J. M. (2005) 'Genomic insights that advance the species definition for prokaryotes.', *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), pp. 2567–2572.

Ladau, J., Sharpton, T.J., Finucane, M.M., Jospin, G., Kembel, S.W., O'dwyer, J., Koeppel, A.F., Green, J.L. and Pollard, K.S. (2013) 'Global marine bacterial diversity peaks at high latitudes in winter', *ISME Journal*, *7*(9), pp. 1669–1677.

Law, J.W.F., Letchumanan, V., Tan, L.T.H., Ser, H.L., Goh, B.H. and Lee, L.H. (2020) 'The Rising of "Modern Actinobacteria" Era', *Progress In Microbes & Molecular Biology*, *3*(1), pp. 1–6.

Lee, I., Kim, Y.O., Park, S.C. and Chun, J. (2016) 'OrthoANI: an improved algorithm and software for calculating average nucleotide identity'. *International Journal Of Systematic And Evolutionary Microbiology*, *66*(2), pp. 1100-1103.

Lehmann, K. B. and Neumann, R. O. (1896). 'Atlas und Grundriss der Bakteriologie und Lehrbuch der Speciellen Bakteriologischen Diagnostik', *Teil II, München: Lehmann*.

Lewin, G.R., Carlos, C., Chevrette, M.G., Horn, H.A., McDonald, B.R., Stankey, R.J., Fox, B.G. and Currie, C.R. **(2016)** 'Evolution and Ecology of Actinobacteria and Their Bioenergy Applications', *Annual Review of Microbiology*, *70*, pp. 235–254.

Li, Q., Chen, X., Jiang, Y. and Jiang, C (2016) 'Morphological Identification of Actinobacteria', *Actinobacteria - Basics and Biotechnological Applications*. pp. 59-86.

Li, W. (2009) 'Analysis and comparison of very large metagenomes with fast clustering and functional annotation', *BMC Bioinformatics*, *10*(1), pp. 1–9.

Li, X. and Qin, L. (2005) 'Metagenomics-based drug discovery and marine microbial diversity', *Trends in Biotechnology*, *23*(11), pp. 539–543.

Liu, T., Wu, S., Zhang, R., Wang, D., Chen, J. and Zhao, J. (2019a) 'Diversity and antimicrobial potential of Actinobacteria isolated from diverse marine sponges along the Beibu Gulf of the South China Sea', *FEMS Microbiology Ecology*, *95*(7), p. 89.

Liu, L., Anderson, C., Pearl, D. and Edwards, S.V. (2019b) 'Modern Phylogenomics: Building Phylogenetic Trees Using the Multispecies Coalescent Model', *Methods in Molecular Biology*, *1910*, pp. 211–239.

Liu, X., Cong, J., Lu, H., Xue, Y., Wang, X., Li, D. and Zhang, Y. (2017) 'Community structure and elevational distribution pattern of soil Actinobacteria in alpine grasslands', *Acta Ecologica Sinica*, *37*(4), pp. 213–218.

Mahapatra, G.P., Raman, S., Nayak, S., Gouda, S., Das, G. and Patra, J.K. (2020) 'Metagenomics Approaches in Discovery and Development of New Bioactive Compounds from Marine Actinomycetes', *Current Microbiology*, *77*(4), pp. 645–656.

Mahmoud, H. M. and Kalendar, A. A. (2016) 'Coral-associated Actinobacteria: Diversity, abundance, and biotechnological potentials', *Frontiers in Microbiology*, 7, p. 204.

Maldonado, L.A., Fenical, W., Jensen, P.R., Kauffman, C.A., Mincer, T.J., Ward, A.C., Bull, A.T. and Goodfellow, M. **(2005a)** '*Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family *Micromonosporaceae*', *International Journal of Systematic and Evolutionary Microbiology*, *55*(5), pp. 1759–1766.

Maldonado, L.A., Stach, J.E., Pathom-aree, W., Ward, A.C., Bull, A.T. and Goodfellow, M (2005b) 'Diversity of cultivable actinobacteria in geographically widespread marine sediments', *Antonie van Leeuwenhoek*, *87*(1), pp. 11–18.

Manivasagan, P., Venkatesan, J. and Kim, S.-K. (2013) 'Introduction to Marine Actinobacteria'. *Marine Microbiology: Bioactive Compounds and Biotechnological Applications*, pp. 1-19.

Manivasagan, P., Kang, K.H., Sivakumar, K., Li-Chan, E.C., Oh, H.M. and Kim, S.K. (2014) 'Marine actinobacteria: An important source of bioactive natural products', *Environmental toxicology and pharmacology*, *38*(1), pp. 172-188

Marchesi, J. R. and Ravel, J. (2015) 'The vocabulary of microbiome research: a proposal', *Microbiome*, *3*(1), pp. 1–3.

Mast, Y. and Stegmann, E. (2019) 'Actinomycetes: The Antibiotics Producers', *Antibiotics*, *8*(3), p. 105.

Meier-Kolthoff, J.P., Göker, M., Spröer, C. and Klenk, H.P. (2013) 'When should a DDH experiment be mandatory in microbial taxonomy?', *Archives of Microbiology*, *195*(6), pp. 413–418.

Menendez, E. and Carro, L. (2019) 'Actinobacteria and Their Role as Plant Probiotics', pp. 333–351.

Mincer, T.J., Jensen, P.R., Kauffman, C.A. and Fenical, W. (2002) 'Widespread and Persistent Populations of a Major New Marine Actinomycete Taxon in Ocean Sediments', *Applied and environmental microbiology*, *68*(10), pp. 5005–5011.

Mineta, K. and Gojobori, T. (2016) 'Databases of the marine metagenomics', *Gene*, *576*(2), pp. 724–728.

Missailidis, S. (2008). *Anticancer therapeutics*. John Wiley & Sons.

Mohamed, A., Nguyen, C. H. and Mamitsuka, H. (2016) 'Current status and prospects of computational resources for natural product dereplication: A review', *Briefings in Bioinformatics*, *17*(2), pp. 309–321.

Palmer, M., Steenkamp, E.T., Blom, J., Hedlund, B.P. and Venter, S.N. (2020) 'All ANIs are not created equal: implications for prokaryotic species boundaries and integration of ANIs into polyphasic taxonomy'. *International Journal of Systematic and Evolutionary Microbiology*, *70*(4), pp.2937-2948.

Pamilo, P. and Nei, M. (1988) 'Relationships between Gene Trees and Species Trees'. *Molecular biology and evolution*, *5*(5), pp.568-583.

Parker, C. T., Tindall, B. J. and Garrity, G. M. (2019) 'International code of nomenclature of Prokaryotes', *International Journal of Systematic and Evolutionary Microbiology*, *69*(1), pp. S1–S111

Prakash, O., Verma, M., Sharma, P., Kumar, M., Kumari, K., Singh, A., Kumari, H., Jit, S., Gupta, S.K., Khanna, M. and Lal, R. (2007) 'Polyphasic approach of bacterial classification — An overview of recent advances', *Indian Journal of Microbiology*, *47*(2), pp. 98–108.

Prescott, L., Harley, J. & Klein, D. (2008). Microbiology. 7th ed. New York: McGraw Hill, p. 1070.

Ramasamy, D., Mishra, A.K., Lagier, J.C., Padhmanabhan, R., Rossi, M., Sentausa, E., Raoult, D. and Fournier, P.E. (2014) 'A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species', *International Journal of Systematic and Evolutionary Microbiology*, *64*, pp. 384–391.

Ren, H., Wang, B. and Zhao, H. (2017) 'Breaking the silence: new strategies for discovering novel natural products', *Current opinion in biotechnology*, *48*, pp. 21–27.

Richter, M. and Rosselló-Móra, R. (2009) 'Shifting the genomic gold standard for the prokaryotic species definition', *Proceedings of the National Academy of Sciences of the United States of America*, *106*(45), pp. 19126–19131.

Salam, N., Jiao, J.Y., Zhang, X.T. and Li, W.J. (2020) 'Update on the classification of higher ranks in the phylum *Actinobacteria*', *International Journal of Systematic and Evolutionary Microbiology*, *70*(2), pp. 1331–1355.

Salwan, R. and Sharma, V. (2020) 'Molecular and biotechnological aspects of secondary metabolites in actinobacteria', *Microbiological Research*, *231*, p. 126374.

Sanschagrin, S. and Yergeau, E. (2014) 'Next-generation Sequencing of 16S Ribosomal RNA Gene Amplicons', *Journal of Visualized Experiment, 90,* p. e51709.

Santos, S. R. and Ochman, H. (2004) 'Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins', *Environmental Microbiology*, *6*(7), pp. 754–759.

Schloss, P. D. and Handelsman, J. (2003) 'Biotechnological prospects from metagenomics', *Current Opinion in Biotechnology*, *14*(3), pp. 303–310.

Servin, J.A., Herbold, C.W., Skophammer, R.G. and Lake, J.A. (2008) 'Evidence Excluding the Root of the Tree of Life from the Actinobacteria', *Molecular Biology and Evolution*, *25*(1), pp. 1–4.

Setubal, J. C. (2021) 'Metagenome-assembled genomes: concepts, analogies, and challenges', *Biophysical Reviews*, *13*(6), pp. 905–909.

Simon, C. and Daniel, R. (2009) 'Achievements and new knowledge unraveled by metagenomic approaches', *Applied Microbiology and Biotechnology*, *85*(2), pp. 265–276.

Singh, T.A., Passari, A.K., Jajoo, A., Bhasin, S., Gupta, V.K., Hashem, A., Alqarawi, A.A. and Abd_Allah, E.F. (2021) 'Tapping Into Actinobacterial Genomes for Natural Product Discovery', *Frontiers in Microbiology*, *12*, p. 1662.

Singh, V. and Chibale, K. (2022) 'Strategies to combat multi-drug resistance in tuberculosis'. *Accounts of chemical research*, *54*(10), pp. 2361-2376.

Stackebrandt, E. and Goebel, B. M. (1994) 'Taxonomic Note: A Place for DNA-DNA Reassociation and 16S-rRNA Sequence Analysis in the Present Species Definition in Bacteriology', *International Journal of Systematic and Evolutionary Microbiology*, *44*(4), pp. 846–849.

Stackebrandt, E. (2003) 'The richness of prokaryotic diversity: There must be a species somewhere', *Food Technology and Biotechnology*, *41*(1), pp. 17–22.

Staley, J. T. (2006) 'The bacterial species dilemma and the genomic-phylogenetic species concept.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *361*(1475), pp. 1899–909.

Sun, D.L., Jiang, X., Wu, Q.L. and Zhou, N.Y. (2013) 'Intragenomic heterogeneity of 16S-rRNA genes causes overestimation of prokaryotic diversity', *Applied and Environmental Microbiology*, *79*(19), pp. 5962–5969.

Sutcliffe, I. C., Trujillo, M. E. and Goodfellow, M. (2012) 'A call to arms for systematists: Revitalising the purpose and practises underpinning the description of novel microbial taxa', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *101*(1), pp. 13–20.

Takahashi, M., Kryukov, K. and Saitou, N. (2009) 'Estimation of bacterial species phylogeny through oligonucleotide frequency distances', *Genomics*, *93*(6), pp. 525–533.

Tindall, B.J., Rosselló-Móra, R., Busse, H.J., Ludwig, W. and Kämpfer, P. (2010) 'Notes on the characterization of prokaryote strains for taxonomic purposes', *International Journal of Systematic and Evolutionary Microbiology*, *60*(1), pp. 249–266.

Udwary, D.W., Zeigler, L., Asolkar, R.N., Singan, V., Lapidus, A., Fenical, W., Jensen, P.R. and Moore, B.S. (2007) 'Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica* '. *Proceedings of the National Academy of Sciences*, *104*(25), pp.10376-10381.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. and Swings, J. (1996) 'Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics', *Microbiological Reviews*, *60*(2), pp. 407–438.

Vandamme, P. and Sutcliffe, I. (2021) 'Out with the old and in with the new: time to rethink twentieth century chemotaxonomic practices in bacterial taxonomy', *International Journal of Systematic and Evolutionary Microbiology, 71*, p. 5127.

Ward, A. C. and Bora, N. (2006) 'Diversity and biogeography of marine actinobacteria', *Current Opinion in Microbiology*, *9*(3), pp. 279–286.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R., Stackebrandt, E.S.M.P. and Starr, M.P. (1987) 'Report of the ad hoc committee on reconciliation of approaches to bacterial systematics'. *Int J Syst Bacteriol*, *37*, pp. 463-464.

Woese, C. R. and Fox, G. E. (1977) 'Phylogenetic structure of the prokaryotic domain: the primary kingdoms.', *Proceedings of the National Academy of Sciences of the United States of America*, *74*(11), pp. 5088–5090.

Yue, Y., Huang, H., Qi, Z., Dou, H.M., Liu, X.Y., Han, T.F., Chen, Y., Song, X.J., Zhang, Y.H. and Tu, J. (2020) 'Evaluating metagenomics tools for genome binning with real metagenomic datasets and CAMI datasets'. *BMC bioinformatics*, *21*(1), pp.1-15.

Zhao, X. Q. (2011) 'Genome-based studies of marine microorganisms to maximize the diversity of natural products discovery for medical treatments', *Evidence-based Complementary and Alternative Medicine*, *2011*.

Zhi, X.Y., Zhao, W., Li, W.J. and Zhao, G.P. (2012) 'Prokaryotic systematics in the genomics era', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 101(1), pp. 21–34.

Zhu, H., Sandiford, S. K. and Van Wezel, G. P. (2014) 'Triggers and cues that activate antibiotic production by actinomycetes', *Journal of Industrial Microbiology and Biotechnology*, pp. 371–386.

CHAPTER 2:

INVESTIGATING THE CULTURABLE ACTINOBACTERIAL DIVERSITY OF MARINE SEDIMENT SAMPLES FROM GERICKE'S POINT

SUMMARY

A total of 33 presumptive actinobacterial strains was isolated from 18 different sediment samples taken from three different sites at Gericke's Point in Sedgefield, South Africa. The three sites were a rock pool (rocky site), a beach area (dry site) and a subtidal zone (ocean site). The investigation of six of the presumptive actinobacterial isolates was discontinued due to consistent contamination during cultivation. The remaining 27 isolates were successfully identified to the genus level following 16S-rRNA gene amplification, sequencing and analysis. A total of 26 *Streptomyces* strains was isolated across the three sites. One isolate belongs to the genus *Nocardia*. Phylogenetic trees based on partial 16S-rRNA gene sequences obtained from each identified strain were generated to investigate their closest phylogenetic relatives. Multiple isolates were found to have identical 16S-rRNA gene sequences indicating they could be clones. Three groups of potential clones were identified. The first group (PC1) contained *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUM1#26, the second group (PC2) contained *Streptomyces* strains GGUI10#8, GGUI#9, GGUI#10, GUUM1#11, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33 and the final group (PC3) contained *Streptomyces* strains GGUI#1 and GGUM1#2. In addition to phenotypic evidence, further analysis of partial gyrase subunit B (*gyrB*) gene sequences of the strains in each PC group indicated that the strains in PC1 and PC3 were not clones. Further analysis of the partial *gyrB* gene sequences of *Streptomyces* strains in PC2 indicated that strains GGUI#10, GUUM1#11, GGUI10#15 and GGUI#28 (PC2a) and strains GGUI10#8, GGUI#9, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 (PC2b) could be clones (identical partial *gyrB* gene sequences for the strains in each sub-group). However, the morphology of strains GGUI#10, GUUM1#11, GGUI10#15 and GGUI#28 (in PC2a) and strains GGUI#9 and GGUI#16 (in PC2b) were significantly different than other strains in PC2 suggesting they are not clones. Further study (including phylogenetic analysis) of the *gyrB* gene sequence of the *Nocardia* strain was also conducted.

TABLE OF CONTENTS

2.1. INTRODUCTION

During the golden era of antibiotic discovery (1950-1960), thousands of antibiotics were discovered. Most of these antibiotics (~60%) were isolated from species belonging to the actinomycete genus *Streptomyces* (Bérdy, 2012). After the sixties, there was a decline in novel antibiotic discovery and an increase in rediscovery of known compounds. This was due, in part, to the declining efficiency of standard culturing techniques and the continuous re-isolation of known bacterial species. Antibiotic resistance has increased the need for novel antibiotics and consequently, the need for innovative approaches to antibiotic discovery. Specifically, focus has shifted to exploring understudied habitats (such as those within the marine environment) as sources of exploitable novel bacterial diversity. The efficiency of these explorations has been enhanced by the use of selective isolation techniques (Bredholdt *et al.*, 2007).

Selective isolation involves various sample pre-treatments and the utilization of selective media containing unusual carbon sources to favour the growth of certain bacteria. As actinobacteria have proved to be the richest source of bioactive compounds, emphasis has been placed on selective isolation to increase the chances of discovering novel actinobacterial species that could produce compounds with novel chemical structures. Due to the continual rediscovery of similar and known compounds from described *Streptomyces* species, even more emphasis has been placed on the selective isolation of rarer actinobacteria genera (i.e. non-*Streptomyces* genera). That being said, some emphasis is still being placed on isolating novel *Streptomyces* species from extreme habitats, as they have proven to still be a rich source of bioactive compounds (Sivalingam *et al.*, 2019). Since actinobacterial spores are more resistant to desiccation than most bacteria, common sample pre-treatments for the selective isolation of spore-forming actinobacteria include drying and heat shock of environmental samples (Goodfellow, 2014). Most actinobacterial selective media also contain complex carbon sources, such as starch and casein, which favour the growth of these organisms (Goodfellow, 2014). Additionally, the presence of seawater in isolation media is crucial for the selective isolation of indigenous marine actinobacteria (Jensen *et al.*, 1991; Maldonado *et al.*, 2005). The composition of selective isolation media is usually based on known phenotypic characteristics of target strains. For example, in this study, some isolation media were supplemented with the antibiotic, rifampicin, to encourage the growth of *Salinispora* strains above other bacteria. This selective isolation technique is based on the antibiotic resistance profile of *Salinispora* species (Mincer *et al.*, 2005) Actinobacterial

selective isolation media are also supplemented with other antibiotics to prevent the growth of unwanted contaminants such as fungi and Gram-negative bacteria.

Due to the conserved nature of the gene, analyses of the hypervariable regions in 16S-rRNA gene sequences are still widely used in taxonomy to assess bacterial diversity and infer evolutionary relationships. However, due to the slow evolutionary rate of the 16S-rRNA gene, the resolution of these analyses has been found to be insufficient to distinguish closely-related species (Santos and Ochman, 2004). Consequently, additional gene markers have been used to analyse interspecific relationships. These include housekeeping genes such as *gyrB* (encoding DNA gyrase subunit B) and *rpoB* (encoding the beta subunit of RNA polymerase), which are conserved between bacterial species, but have higher mutation rates than the 16S-rRNA gene (Anderson and Wellington, 2001; Ogier *et al.*, 2019). Phenotypic data and data derived from the analysis of single marker genes can be utilized to identify potentially novel isolates. However, to robustly assess the novelty of an isolate, whole-genome analyses, such as dDDH and ANI need to be performed (Chun *et al.*, 2018). If the 16S-rRNA gene sequence similarity between two actinobacterial strains is above 98.8%, DDH is mandatory to establish if they belong to the same or distinct genomic species (Meier-Kolthoff *et al.*, 2013). Additionally, different levels of *gyrB* gene sequence similarity (ranging from 95.0% to 98.5%) have been proposed for species delineation of various prokaryotic taxa (Kim *et al.*, 2012a). This threshold depends on the organisms involved and the length of the gene fragments being analysed. In the genus *Streptomyces*, approximately 98.5 % *gyrB* gene sequence similarity has been shown to correspond to a DDH relatedness threshold value of 70% for species delineation (Law *et al.*, 2018).

In the work presented in this chapter, the aim was to selectively isolate a diverse range of actinobacteria from samples taken from three marine habitats at Gericke's Point, South Africa. Presumptive actinobacterial isolates were selected based on colony morphology and identified to the genus level by partial 16S-rRNA gene sequencing analysis. The partial 16S-rRNA gene sequences obtained for each confirmed actinobacterial isolate were used to determine their phylogenetic positions within their respective genera. Further investigation of specific isolates was done utilising the *gyrB* gene to distinguish clones from closely related strains.

2.2 METHODS AND MATERIALS

2.2.1 Isolation and identification of marine actinobacterial strains

2.2.1.1 Sample collection

Samples of marine sediment were collected in October 2019 from Gericke's Point, Garden Route National Park in Sedgefield, South Africa (-34.0269863, 22.7746458). Sampling was performed by members of the BTB research group. Three samples containing 100g each of sediment were collected (in duplicate) from three different sites: a rock pool (rocky site), a beach area (dry site) and a subtidal zone (ocean site). There were 18 samples in total. The sampling approach can be seen in Figure 2.1. A photographic record was kept of fauna, flora, distinguishing geological features and human impact (pollution) surrounding the sample sites. Physiochemical analyses of the sediment were conducted by Bemlab Laboratories, Strand, South Africa. The temperature of the sediment was 19°C at the time of sampling and the pH was found to be 8.0, 8.7 and 8.8 for the ocean site, rocky site and dry site sediment samples, respectively. The sediment was stored for approximately five months in sterile 50 ml conical centrifuge tubes at 4°C until sample preparation.

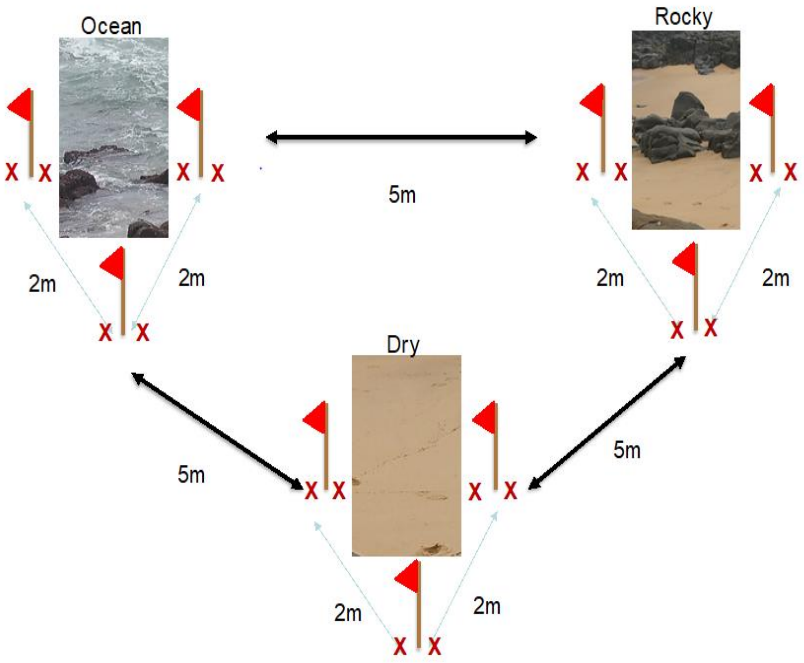

Figure 2.1 Sampling approach for sediment sample collection at Gericke's Point. Flags denote points where sediment samples were collected at each sampling sub-site. Duplicate samples were collected at each point. This image was constructed by the BTB research group.

2.2.1.2 Sediment sample preparation

The sediment samples were prepared for the isolation of actinobacteria by weighing 1g of each sample into separate sterile universal containers. Sterile Ringer's solution [7.2g NaCl, 0.37g KCl, 0.22g $CaCl_2.2H_20$, distilled water to 1L, pH 7.4] was added to the samples to make the total volume in each universal container up to 10ml. Each sample mixture was vortexed for 1 min to dislodge bacteria from sediment particles. A 10-fold dilution series of the bacterial suspension was prepared with sterile Ringer's solution down to a final dilution of 10^{-2}. Using the spread plate technique, 100µl aliquots of undiluted and diluted (10^{-1} and 10^{-2}) bacterial suspensions were spread onto prepared agar media plates in duplicate.

2.2.1.3 Isolation of actinobacterial strains

Various media were used to accommodate the growth of marine actinobacteria based on previous reports. Key aspects such as temperature, pH and salt concentration of the environment sampled were taken into consideration. All media were supplemented with 3.8% Red Sea Salt (Red Sea Fish Pharm Ltd, Tel Aviv, Israel).The media used were M1 medium (Jensen *et al.*, 2005), International *Streptomyces* Project medium No. 2 containing 0.2% calcium carbonate (ISP2-C), diluted ISP2-C (1-in-10 dilution) and International *Streptomyces* Project medium No. 5 (ISP5) (Shirling and Gottlieb, 1966). Penicillin (100µg/ml), nalidixic acid (25µg/ml) and cycloheximide (100µg/ml) were used to supplement the actinobacterial isolation media to limit the growth of Gram-negative bacteria (penicillin and nalidixic acid) and fungi (cyclohexamide). Additionally, rifampicin (15µg/ml) was used in some media in an attempt to selectively isolate *Salinispora* strains. All isolation plates were incubated at room temperature (~22 °C).

Due to the COVID-19 pandemic and South Africa's extended hard lockdown in 2020, all bacterial isolation plates were incubated for a period of six months before they could be inspected. Isolation plates are usually incubated for a minimum of one month with weekly inspections. As weekly checking was not possible, some plates had become overgrown with moulds before they could be inspected and no actinobacterial colonies could be isolated from them. Once all the plates with fungal overgrowth were discarded, the remaining plates were inspected for actinobacterial colonies.

Presumptive marine actinobacterial colonies were selected based on certain morphological features. These included colonies with leathery surfaces, matt looking colonies and colonies

that were slightly sunken into the surface of the agar and which displayed aerial mycelium that had developed into spores. Presumptive marine actinobacterial colonies were sub-cultured onto fresh agar plates (excluding antibiotics) of the same medium from which they were isolated and each new sub-cultured colony was given a strain name and numbered chronologically e.g. GGUI#1, GGUM1#2, GGUI5#3. The strain names were derived using a standardized naming system set out within the larger project of which this study was a part. 'GG' indicates that a strain was isolated from Gericke's Point in Garden Route National Park and 'U' indicates the soil samples were untreated (no physical or mechanical pre-treatment was used in the part of the study described in this thesis). This is followed by the media type on which the strain was isolated (I indicates ISP2-C, I5 indicates ISP5, I10 indicates diluted ISP2-C and M1 indicates M1 medium) and the isolate number. For example, GGUI#1 denotes isolate No.1 in this study that was isolated on ISP2-C medium from an untreated sediment sample from Gericke's Point. The site and sample from which the strain was isolated were recorded separately. Each strain was also streaked onto International *Streptomyces* Project medium No. 4 (Shirling and Gottlieb, 1966) agar plates and inspected after being incubated for 14 days at 30°C (all isolates were confirmed to grow at 30°C during the investigations detailed in Chapter 4) to record phenotypic characteristics such as substrate mycelium colour, spore mass colour and whether the strain produced a diffusible pigment.

2.2.1.4 Cultivation of actinobacterial strains

Liquid cultures were prepared by inoculating actinobacterial strains into 20ml of liquid medium in 250-ml Erlenmeyer flasks and growing them on an orbital shaker (110 rpm) at room temperature (~22 °C) until sufficient cell mass was observed. These cultures were then inspected for contamination by Gram staining and streaking for single colonies onto agar plates of the same media type as that in which the strain was grown. Additional inspection had to be done when it was observed that contamination by non-actinobacteria went undetected using these techniques (i.e. sometimes apparently pure actinobacterial cultures yielded high-quality DNA-sequencing chromatograms of the 16S-rRNA gene, but the sequence belonged to a Gram-negative bacterium rather than an actinobacterium). To overcome the problem of "hidden" contamination, a colony from the single colony plate was suspended in sterile H_20, a dilution series was set-up (up to 10^{-5} dilution) using sterile H_20 as the diluent and 100µl of each dilution was spread plated on appropriate agar plates. Plates containing the higher dilutions revealed whether a contaminant had gone undetected, as distinct actinobacterial and non-actinobacterial colonies could be seen. When a contaminant was observed, a single colony of

the actinobacterial strain was used to streak for single colonies again and the dilution series was repeated to separate the actinobacterial and contaminant strains. The actinobacterial strain was then re-inoculated into liquid medium. Once contaminated cultures had been purified (i.e. no contamination was detected using the described inspections), glycerol stocks of each isolate were prepared by adding 1000µl of broth culture to 300µl of 50% (v/v) sterile glycerol in a sterile microcentrifuge tube. Stock cultures were stored at either -20°C or -80°C for long-term storage.

2.2.1.5 Genomic DNA extraction

A benchtop centrifuge (Eppendorf Centrifuge 5418) was used to harvest cells from the liquid cultures (14000 rpm for 1 min per run) to obtain ~200µl of wet cell mass for genomic DNA extraction. Genomic DNA (gDNA) was extracted using two different methods. The first is based on a modified version of Wang et al.'s phenol extraction method (1996). An increased lysozyme concentration of 25mg/ml was used and the cells were incubated in lysis buffer at 37°C overnight rather than for 30 minutes. The pelleted gDNA was dissolved in varying amounts of 10 mM Tris- HCl, 1 mM EDTA (TE) buffer (pH 8.0) depending on pellet sizes and stored at 4°C. Alternatively, gDNA was extracted using the ZR Soil Microbe DNA miniprep Kit™ (Zymo Research, Irvine, California) according to the manufacturer's protocol for extraction from bacterial cells. Instead of lysing via bead beater, the cells were placed in BashingBead™ Lysis Tubes (beads of 0.1 and 0.5 mm in diameter; Zymo Research, Irvine, California) and vortexed for 5 min. DNA concentrations were measured using a Nanodrop™ spectrophotometer, model ND-1000 UV-VIS.

2.2.1.6 16S-rRNA and gyrB gene amplification

Extracted gDNA was amplified using the polymerase chain reaction (PCR). The universal forward primer F1 and the universal reverse primer R5 were used to amplify 16S-rRNA gene sequences (Cook & Meyers, 2003), while the forward primer GgyrB-F2 (GF2) (Le Roes *et al.*, 2008) and three different reverse primers (Everest and Meyers, 2009; Kirby *et al.*, 2010) were used to amplify partial *gyrB* gene sequences (Table 2.1). PCR reactions (50 µl) contained 2 mM $MgCl_2$, 600 µM dNTP, 0.5 µM of each primer, 100 – 500 ng/µl template DNA and 0.5 units Super-Therm *Taq* polymerase (JMR Holdings Inc, USA). Bovine serum albumin (BSA; 200 ng/µl final concentration) was added to the PCR reactions as it has been shown to reduce the effect of substances that inhibit enzyme activity and to prevent adhesion

of enzymes to reaction tubes (Kreader, 1996). A standard PCR programme was followed for all PCR reactions: an initial denaturation at 96 ^{o}C for 2 min, 30 cycles of denaturation at 96 ^{o}C for 30s, annealing for 20s and extension at 72 ^{o}C for 30s, with a final extension at 72 ^{o}C for 5 min. PCR programmes were performed on a Bio-Rad T100^{TM} Thermal Cycler PCR machine. PCR products were electrophoresed on 0.8% (w/v) agarose gels containing ethidium bromide (600 ng/ml) at 90V. A molecular marker of *Pst*I-digested bacteriophage lambda DNA was used and the agarose gel was visualized using a Bio-Rad ChemiDoc^{TM} XRS+ Molecular Imager® illuminometer.

Table 2.1 PCR primers used to amplify and sequence 16S-rRNA and *gyrB* gene sequences and their respective annealing temperatures. "F" denotes the forward primers and "R" denotes the reverse primers.

Primer	Primer sequence [R=A/G, S=C/G ,W=A/T, Y=C/T, H=A/C/T, V=A/C/G, N=any base, I= Inosine]	Annealing Temperature (^{o}C)	Reference
F1	5'-AGAGTTTGATCITGGCTCAG-3'	56	Cook & Meyers (2003)
R5	5'-ACGGITACCTTGTTACGACTT-3'	56	Cook & Meyers (2003)
GgyrB-F2	5'-CAGTTCGAGGGHCAGACSAAGAC-3'	56	Le Roes *et al.* (2008)
7GgyrB-R	5'-CCGTCVACRTCRGCRTCSGCCATS-3'	59	Everest and Meyers (2009)
KgyrB-R	5'-CGATCCGGGCCTTCTCGACGTTCAG-3'	62	Kirby *et al.* (2010)
KgyrB-R1892	5'-CCSAGRCCCTTGWAGCGCTGG-3'	59	Kirby *et al.* (2010)
518F	5'-CCAGCAGCCGCGGTAATACG-3'	55	Faisal *et al.* (2017)
800R	5'-TACCAGGGTATCTAATCC-3'	50	Faisal *et al.* (2017)

2.2.1.7 Identification of isolates: DNA sequencing and analysis

PCR products were purified using the ISOLATE II PCR and Gel Kit (Bioline) according to the manufacturer's protocol. Sequencing was performed as a service by Macrogen Europe Inc. (Amsterdam-Zuidoost, Netherlands) using the universal primers 518F and 800R for the 16S-rRNA gene, and GF2 for the *gyrB* gene (Table 2.1). Sequence chromatograms were analysed and edited using Chromas version 2.33 (Technelysium Pty. Ltd., Australia). Actinobacterial isolates were identified to the genus level by comparing the edited 16S-rRNA partial gene sequences to the curated database (Kim *et al.*, 2012b) on the EzBioCloud platform (https://www.ezbiocloud.net/). This also identified the closest phylogenetic relatives of each isolate. The Basic Local Alignment Search Tool (BLAST; Altschul *et al.*, 1990) was used to compare edited *gyrB* partial gene sequences of selected strains against the GenBank nucleotide database (Sayers *et al.*, 2019). The similarity of the 16S-rRNA and *gyrB* gene

sequences between isolates was assessed by aligning the sequences using DNAMAN version 4.13 (Lynnon Biosoft).

2.2.2 Phylogenetic analyses

2.2.2.1 16S-rRNA gene phylogenetic analysis

The 16S-rRNA gene sequences of the type strains of the top five species with the highest nucleotide identity (top 5 hits) were downloaded from the EzBioCloud database for strains identified to belong to the genus *Streptomyces.* For strains having identical top hits, the next five top hits were selected to avoid repetition when constructing phylogenetic trees. For the strain that was identified as belonging to the genus *Nocardia*, the 16S-rRNA gene sequences of the type strains of the top 30 hits were selected from the EzBioCloud database. The software package Molecular Evolutionary Genetics Analysis version 7.0 (MEGA7; Kumar *et al.*, 2016) was used to align the 16S-rRNA gene sequences using the MUSCLE multiple sequence alignment algorithm (Edgar, 2004). The alignments for each genus were edited to ensure all sequences were the same length and to remove columns containing ambiguous bases, insertions or deletions. The edited alignments were used to construct Maximum Likelihood (Felsenstein, 1981), Neighbour-Joining (Saitou and Nei, 1987) and Maximum Parsimony (Fitch, 1971) trees using the default settings in MEGA7. The Kimura two-parameter model (Kimura, 1980) was used as the substitution model for the maximum-likelihood and neighbour-joining trees. Bootstrapping for all trees was carried out using 1000 replicates (Felsenstein, 1985).

2.2.2.2 Gyrase subunit B phylogenetic analysis

The *gyrB* gene sequences of the type strains of the most closely related species (with validly published names) for which whole-genome or partial *gyrB* sequences were available in the GenBank database were downloaded for the isolated *Streptomyces* strains identified to be potential clones (i.e. strains that had identical partial 16S-rRNA gene sequences). This was also done for the 30 most closely related species of the marine *Nocardia* strain. A combination of results obtained from the EzBioCloud database searches (16S-rRNA gene top hits) and the *gyrB blastn* analysis were used to select the most closely related species for each strain. The *gyrB* gene sequences were aligned, edited and used for the generation of phylogenetic trees as described in section 2.2.2.1.

2.3 RESULTS AND DISCUSSION

2.3.1 Isolation and cultivation of actinobacterial strains

Due to the COVID-19 pandemic and the abnormally long incubation period of the isolation plates before inspection, about two thirds of the 864 plates that had been inoculated contained fungal overgrowth and had to be discarded. The discarded plates were evenly distributed among all sub-sites. The remaining 298 plates were used to select a total of 33 presumptive actinobacterial isolates across 18 environmental samples. Two isolates were selected from the ocean site samples, five were selected from the dry site samples and 26 were selected from the rocky site samples. Twelve isolates were selected from the ISP2-C agar plates, five were selected from the ISP2-C (1:10) agar plates, six were selected from the M1 agar plates and ten were selected from the ISP5 agar plates. The investigation of three isolates from the rocky site, namely GGUI#13, GGUI#14 and GGUI5#17 were discontinued due to consistent fungal contamination during each sub-culturing attempt. Notably, strain GGUI#13 was an orange colony with black spore mass (Figure A1). These morphological features are typical of bacteria belonging to the actinobacterial genera *Micromonospora* and *Salinispora (*Jensen *et al.*, 2015).

Other isolates that displayed interesting phenotypic characteristics on ISP4 medium include GGUI#1, GGUI10#6, GGUI#7, GGUI10#20, GGUI#25 and GGUI#28 (Figure 2.2). Strain GGUI#1 produced a pale pink diffusible pigment. Strain GGUI10#6 displayed dark brown to black wrinkly colonies with a white spore mass. Strain GGUI#7 displayed yellow colonies that were sunken into the agar with a dark grey and white spore mass. Strain GGUI10#20 displayed colonies with a pale pink spore mass and a ‘branched’ structure. Strain GGUI#25 displayed wrinkled and ‘volcano like’ dark brown colonies with a grey spore mass. Individual colonies of GGUI#28 also displayed a wrinkly morphology, but were bright orange in colour with a white spore mass. A full record of all 30 pure-culture strains and their phenotypic characteristics can be found in the appendix (Table A1).

2.3.2. Identification of isolates: 16S-rRNA and *gyrB* gene analyses

The remaining 30 presumptive actinobacterial isolates were cultured in liquid media to extract, amplify and sequence their DNA for genus identification. Twenty seven (27) of the 30 isolates were successfully identified to the genus level and confirmed as actinobacteria. The identification of the remaining three isolates (GGUI5#21, GGUI#22 and GGUI5#32)

was discontinued due to consistent contamination by certain Gram-negative bacteria that was only detected after DNA sequencing took place. This was an ongoing issue throughout this study and is detailed in section 2.3.3. Genus identifications were based on partial 16S-rRNA gene sequences ranging from approximately 200–900 base pairs. A summary of all strains identified to the genus level, together with their respective top hits can be found in Table 2.2. All isolates except GGUI10#20, were identified as belonging to the genus *Streptomyces.* Isolate GGUI10#20 (highlighted in pink) was identified as belonging to the genus *Nocardia*.

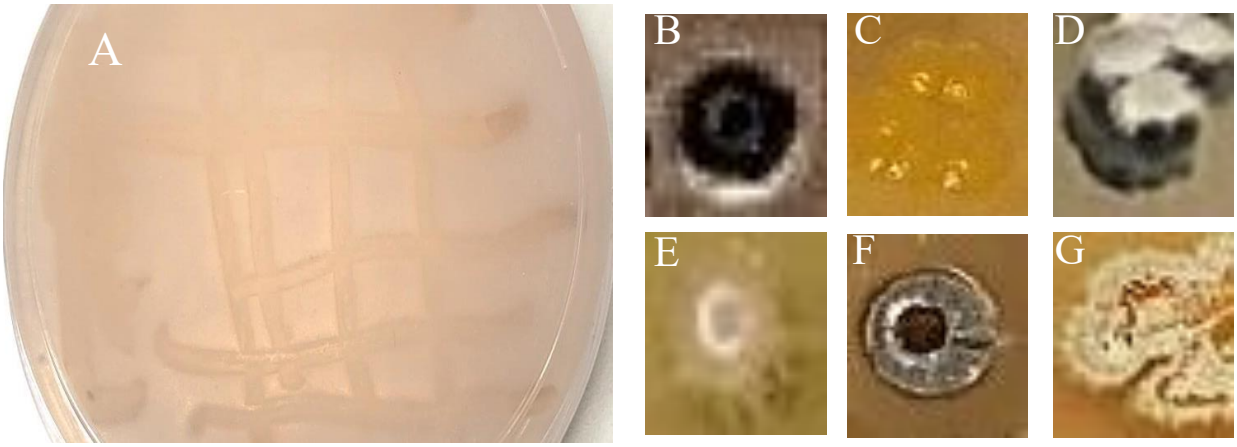

Figure 2.2 Colonies of strains displaying interesting phenotypic features when cultivated on ISP4 medium for 14 days at 30 ^{o}C. (A) Strain GGUI#1 (B) Strain GGUI10#6 (C-D) Strain GGUI#7 (E) Strain GGUI10#20 (F) Strain GGUI#25. (G) Strain GGUI#28

Isolates identified as belonging to the genus *Streptomyces* that had identical 16S-rRNA gene sequences when aligned in DNAMAN, were grouped as potential clones and underwent further analyses using *gyrB* gene sequences. Three potential clone groups were identified: one group included all the isolates with the type strain of *Streptomyces chumphonensis* as their top hit (PC1), another group included all the isolates with the type strains of *Streptomyces bacillarus* and *Streptomyces griseus* subsp. *griseus* as their top hits with 100% 16S-rRNA gene sequence similarity (PC2) and the last group contained the two isolates which had the type strain of *Streptomyces microflavus* as their top hit (PC3). A partial *gyrB* gene sequence was obtained for each strain in potential clone group to determine if the strains had identical sequences (i.e. to determine whether analysis of partial *gyrB* gene sequences indicated that these strains were indeed clones). The full investigation of these strains, including phenotypic and phylogenetic evidence, is detailed in section 2.3.5. Strain GGUI10#20 also underwent further analysis using the *gyrB* gene to gain more insight into the phylogenetic position of this isolate within the genus *Nocardia*.

Table 2.2. Genus identification of 27 actinobacterial strains isolated from the ocean, dry and rocky site sediment samples collected from Gericke's Point based on EzBioCloud analyses of partial 16S-rRNA gene sequences. Further study of presumptive clones (as well as isolate GGUI10#20) was based on *blastn* analyses of partial *gyrB* gene sequences using the GenBank database. The partial 16S-rRNA gene sequences were generated using the 518F primer, except in the case of isolate GGUM1#5 where the 800R primer was used. The partial *gyrB* gene sequences were generated using the GF2 primer. Sequence length indicates the number of base pairs (bp) in the query sequence. The full-length actinobacterial 16S-rRNA and *gyrB* gene sequences are ~1.5kb and ~2.2kb respectively. Percentage (%) identity indicates the degree of similarity between the query sequence and the top hit strain gene sequence. Asterisks (*) indicate isolates with more than one top hit (i.e. multiple sequences from different species were found to have identical percentage similarities to the query sequence). ND indicates that the *gyrB* sequence for that isolate was not determined.

Isolate	Top Hit (16S-rRNA)	% identity	Sequence length (bp)	Top Hit (*gyrB*)	% identity	Sequence length (bp)
			OCEAN SITE			
GGUI#1	*Streptomyces microflavus**	99.88	840	*Streptomyces griseus* subsp. *griseus*	98.21	616
GGUI10#6	*Streptomyces synnematoformans*	98.91	922	ND		
			DRY SITE			
GGUM1#2	*Streptomyces microflavus**	99.89	887	*Streptomyces globisporus*	98.57	631
GGUI#3	*Streptomyces apocyni*	99.52	841	ND		
GGUM1#4	*Streptomyces flavovirens**	99.39	655	ND		
GGUI10#15	*Streptomyces bacillaris **	100	926	*Streptomyces cavourensis*	98.51	672
GGUM1#29	*Streptomyces violaceus **	99.57	935	ND		
			ROCKY SITE			
GGUM1#5^	*Streptomyces olivaceus**	92.19	193	ND		
GGUI#7	*Streptomyces umbrinus**	99.32	591	ND		
GGUI10#8	*Streptomyces griseus* subsp. *griseus**	100	558	*Streptomyces cavourensis*	98.45	645
GGUI#9	*Streptomyces griseus* subsp. *griseus**	100	662	*Streptomyces cavourensis*	98.38	618
GGUI#10	*Streptomyces bacillaris**	100	887	*Streptomyces cavourensis*	98.42	631
GGUM1#11	*Streptomyces bacillaris **	100	913	*Streptomyces cavourensis*	98.50	667
GGUI#12	*Streptomyces pratensis*	99.77	928	ND		
GGUI#16	*Streptomyces bacillaris **	100	892	*Streptomyces cavourensis*	98.37	678
GGUI5#18	*Streptomyces bacillaris **	100	922	*Streptomyces cavourensis*	98.44	640
GGUI10#19	*Streptomyces chumphonensis*	100	597	*Streptomyces roseoverticillatas*	92.83	349
GGUI10#20	***Nocardia otitidiscaviarum***	**99.43**	**883**	***Nocardia otitidiscaviarum***	**99.52**	**414**
GGUI#23	*Streptomyces bacillaris **	99.85	646	ND		
GGUI5#24	*Streptomyces chumphonensis*	100	573	*Streptomyces roseoverticillatas*	92.98	301
GGUI5#25	*Streptomyces kurssanovii*	99.44	885	ND		
GGUM1#26	*Streptomyces chumphonensis*	99.55	895	*Streptomyces laurentii*	92.06	353
GGUI5#27	*Streptomyces bacillaris **	100	929	*Streptomyces cavourensis*	98.48	656
GGUI#28	*Streptomyces griseus* subsp. *griseus**	100	614	*Streptomyces cavourensis*	98.38	680
GGUI5#30	*Streptomyces bacillaris **	100	800	*Streptomyces cavourensis*	98.42	634
GGUI5#31	*Streptomyces bacillaris **	100	855	*Streptomyces cavourensis*	98.50	668
GGUI5#33	*Streptomyces bacillaris **	100	936	*Streptomyces cavourensis*	98.51	672

^Sequence was of poor quality resulting in shorter sequence length. This sequence was not included in any subsequent phylogenetic analyses.

2.3.3 Contamination

As mentioned previously, there was a consistent contamination problem involving certain Gram-negative bacteria that went undetected using standard screening techniques (i.e. Gram staining and streaking for single colonies) and was only discovered after sequencing of the 16S-rRNA gene when sequence analysis showed that the sequence belonged to a non-actinobacterium. These contaminants were most closely related to *Pseudomonas moraviensis*, *Pantoea septica*, *Delftia tsuruhatensis* and various species of the genus *Alcanivorax*. Notably, all contaminants except one were found among actinobacterial strains isolated from the rocky site. One Gram-positive contaminant most closely related to *Staphylococcus saprophyticus* subsp. *saprophyticus* was also encountered in this study. More information on the various contaminants and the isolates affected is shown in the appendix (Table A2).

Discovering this contamination was interesting as most of the time there was no mixed peak pattern observed on the 16S-rRNA sequence chromatograms, which indicated that these organisms existed as "pure" contaminants, yet no colonies of said contaminants were visible on agar plates streaked for single colonies or Gram-stained microscope slides from liquid cultures. It was concluded that this may have occurred due to insufficient lysis of actinobacterial cells during DNA extraction. Thus, if contaminants were present (even in small amounts) among the harvested cells and their cells lysed more easily than the actinobacterial strains, the DNA used in the PCR reaction would most likely belong to the contaminant. This contamination problem was resolved (for the most part) after including an additional screening step before harvesting liquid cultures and using BashingBead™ Lysis Tubes (beads of 0.1 and 0.5 mm in diameter; Zymo Research, Irvine, California) to improve mechanical lysis of actinobacterial cells before DNA extraction. Strains GGUI5#21, GGUI#22 and GGUI5#32 (contaminated by a Gram-negative bacterium most closely related to *Alcanivorax*) could not be identified as they could not be purified from the mixed cultures. Presumptive actinobacterial colonies of strains GGUI5#21 and GGUI5#32 were stab inoculated on agar media and screened for antibacterial activity (described in Chapter 4) before the contamination was detected. No contamination was seen on the stab-inoculated plates. Strain GGUI#22 could not be screened as there was difficulty in culturing the bacterium on agar and in liquid media (slow to no growth) after the contamination was detected.

2.3.4 Phylogenetic analyses

2.3.4.1 Genus *Nocardia*

According to the List of Prokaryotic names with Standing in Nomenclature (Parte, 2018), the genus *Nocardia* is comprised of 123 species with validly published names (https://lpsn.dsmz.de/genus/nocardia). The genus was described by Trevisan (1889) and assigned to the *Nocardiaceae* family (Castellani and Chalmer, 1919) within the actinobacterial order *Mycobacteriales*. *Nocardia* strains are usually present as weakly-staining Gram-positive, filamentous branching bacilli that fragment into coccobacillary forms (Hoza *et al.*, 2017). *Nocardia* strains display colonies that can be smooth, wrinkly, grainy or heaped with an orange, pink, red, or yellow colour, which is caused by carotenoid-like pigments (Goodfellow and Maldonado, 2015). Some *Nocardia* strains also produce brown or yellowish diffusible pigments. *Nocardia* strain GGUI10#20 displayed irregularly shaped colonies that produced a pale pink spore mass and a yellowish brown diffusible pigment (Table A1).

Species belonging to the genus *Nocardia* have been isolated from both soil and aquatic (freshwater, wastewater and marine) habitats (Wright *et al.*, 2021). *Nocardia* strain GGUI10#20 was isolated from the rocky site. Some members of the genus *Nocardia* are also known to cause nocardiosis, an opportunistic infection that most commonly presents as pulmonary disease, but can also affect the brain and skin. Consequently, multiple species of *Nocardia* have also been isolated from human clinical samples. The most common infection-causing *Nocardia* species are *Nocardia asteroides* and *Nocardia brasiliensis*. Although nocardiosis primarily affects humans, multiple *Nocardia* species have been shown to cause infection in marine mammals (Kudo *et al.*,1988; Leger *et al.*, 2009). This includes *Nocardia otitidiscaviarum* and *Nocardia seriolae*, the type strains of which were revealed to be the first and second closest relatives of *Nocardia* strain GGUI10#20, respectively (according to the EzBioCloud database). Marine strains of *N. otitidiscaviarum* have been reported to potentially play a role in the restoration of oil-contaminated marine environments (Vyas and Dave, 2011).

2.3.4.1.1 *Nocardia* 16S-rRNA gene phylogenetic analysis

A phylogenetic tree including the 16S-rRNA gene sequences of 30 *Nocardia* type strains was generated in order to determine the phylogenetic position of *Nocardia* strain GGUI10#20 within the genus *Nocardia* (Figure 2.3). *Nocardia* strain GGUI10#20 can be seen clustering closely with *N. otitidiscaviarum* NBRC 14405^{T} with very strong bootstrap support (99%). This association was also seen in the neighbour joining and maximum parsimony trees. As mentioned in section 2.3.4.1, *N. otitidiscaviarum* was the top hit for *Nocardia* strain GGUI10#20 with 99.43% 16S-rRNA gene sequence similarity over 883 nucleotides (Table 2.2). The full actinobacterial 16S-rRNA gene sequence is ~1500 nucleotides. Conville *et al.* (2018) reported that the genus *Nocardia* has a high 16S-rRNA gene sequence similarity between species that have been shown to be distinct by DDH. Sometimes only a few base differences separate closely related species (Roth *et al.*, 2003). For example, the type strains of the species *Nocardia brevicatena* and *Nocardia paucivorans* have 16S-rRNA gene sequences with 99.5% similarity over 1352 nucleotides. Consequently, the Clinical and Laboratory Standards Institute has recommended a 16S-rRNA gene sequence similarity $\geq$ 99.6% for identification of *Nocardia* isolates to the species level (Petti *et al.*, 2008). The degree of divergence of *gyrB* gene sequences was found to be approximately 3.6 times greater than that of the 16S-rRNA gene in *Nocardia* species (Takeda *et al.*, 2010) , indicating that it would be more effective to use the *gyrB* gene to identify *Nocardia* species isolates to the species level. For this reason, a *gyrB* gene analysis of *Nocardia* strain GGUI10#20 was also included in this study.

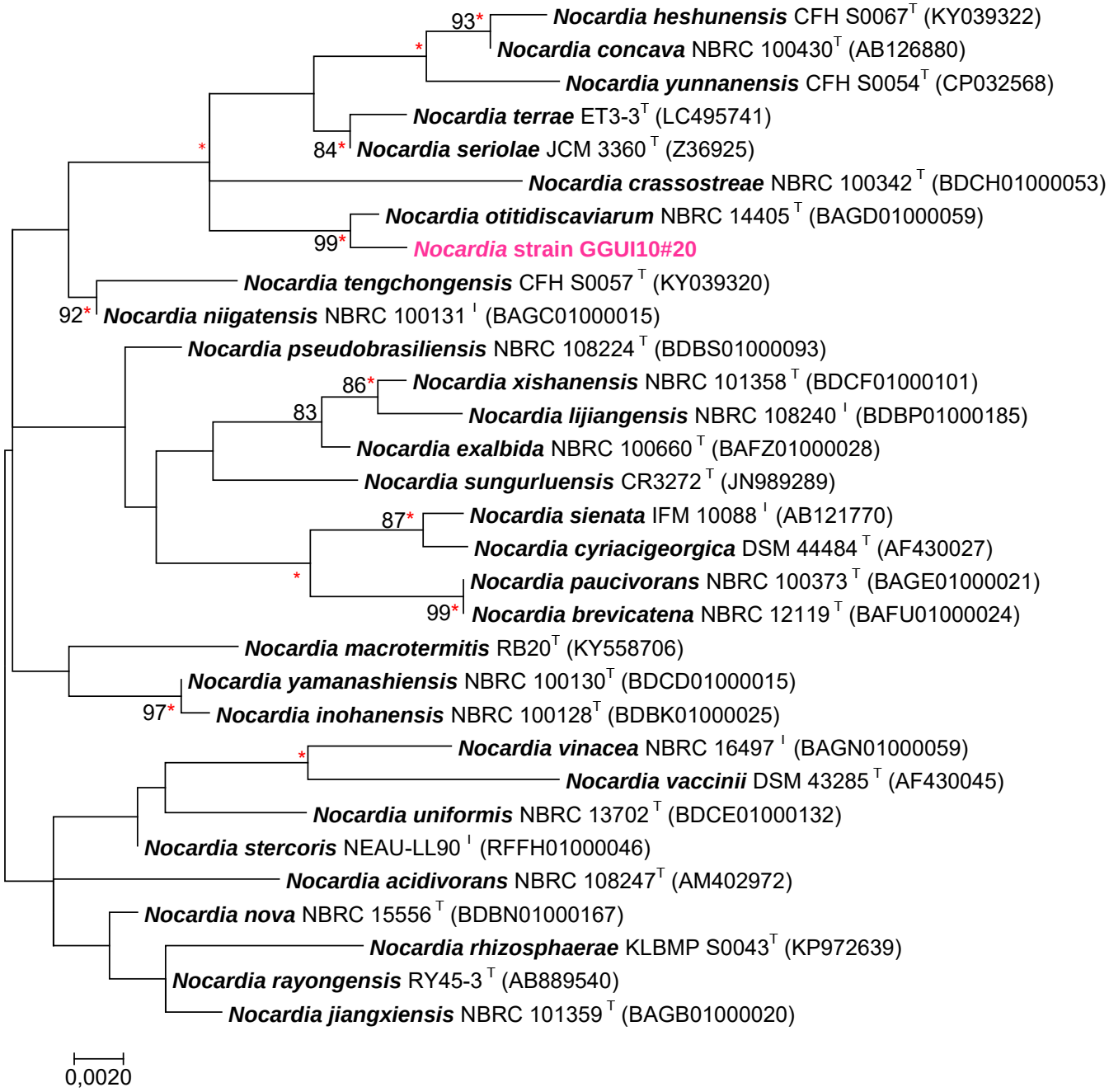

Figure 2.3. Maximum-likelihood phylogenetic tree comparing *Nocardia* strain GGUI10#20 to the type strains of the *Nocardia* species to which it is most closely related based on the top hits obtained from the EzBioCloud database. The tree is based on the alignment of 31 879-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 2 nucleotide substitutions per 1000 nucleotides.

2.3.4.1.2 *Nocardia gyrB* gene phylogenetic analysis

The partial *gyrB* gene sequences of the type strains of the top 30 species to which *Nocardia* strain GGUI10#20 is most closely related and for which *gyrB* gene sequences are available were aligned against the 414 nt partial *gyrB* gene sequence obtained for strain GGUI10#20. The final sequence alignment in MEGA was based on 402-nt *gyrB* gene sequences (approximately 18% of the full actinobacterial *gyrB* gene). *Nocardia* strain GGUI10#20 grouped with *N. otitidiscaviarum* NBRC 14405^{T} with very strong bootstrap support (98%) (Figure 2.4). This association was also seen in the neighbour joining and maximum parsimony trees. This is in agreement with the *blastn* analysis, which revealed *N. otitidiscaviarum* as the top hit of strain, GGUI10#20 with 99.52% *gyrB* gene sequence similarity over 414 nucleotides (Table 2.2). Therefore, it can be noted that *Nocardia* strain GGUI10#20 clustered closely with the type strain of its top hit, *N. otitidiscaviarum*, in both the 16S-rRNA and *gyrB* gene phylogenetic trees. It has been shown that, when compared to other *Nocardia* species, *N. otitidiscaviarum* has <93.5 % *gyrB* gene sequence similarity (Takeda *et al.*, 2010). This makes it is easier to distinguish between *N. otitidiscaviarum* and other *Nocardia* species when using *gyrB* gene sequence information. Similarly, the *blastn* results in this study showed that, after *N. otitidiscaviarum*, the next closest match (*Nocardia farcinica* strain IFM 10152^{T}) had only 94.2% *gyrB* gene sequence similarity when compared to the partial *gyrB* sequence obtained from *Nocardia* strain GGUI10#20. Although it is evident that *Nocardia* strain GGUI10#20 could belong to the species *N. otitidiscaviarum*, it is important to remember these inferences are based on only a small portion of the *gyrB* gene. Since a genome sequence is available for *N. otitidiscaviarum* NBRC 14405^{T}, it is recommended that the genome of strain GGUI10#20 be sequenced and whole-genome analyses (i.e. ANI and dDDH) be used to establish whether it could represent a new species in the genus *Nocardia*.

In the 16S-rRNA phylogenetic tree (Figure 2.3), *Nocardia* strain GGUI10#20 is shown to form part of a bigger cluster containing the type strains of the species *Nocardia heshuensis*, *Nocardia concava*, *Nocardia yunnanensis*, *Nocardia terrae*, *N. seriolae*, and *Nocardia crassostreae*. Of these strains, *N. concava*, *N. seriolae* and *N. crassostreae* were isolated as pathogens from a human, a fish and an oyster, respectively (Kudo *et al.*, 1988; Friedman *et al.*, 1998; Kagayama *et al.*, 2005). In the *gyrB* phylogenetic tree (Figure 2.4), the strains in the aforementioned cluster were shown to be relatively distantly related to *Nocardia* strain

GGUI10#20. The cluster to which strain GGUI10#20 is mostly closely related in the *gyrB* gene phylogenetic tree contains the type strains of the species *Nocardia africana, N. nova, Nocardia aobensis, Nocardia kruczakiae, Nocardia cerradoensis* and *Nocardia elegans*. All of these strains, besides the type strain of *N. cerradoensis,* were isolated as human pathogens. Besides *N. nova*, none of the type strains of these species were included in the 16S-rRNA gene phylogenetic analysis, as they were not among the 30 strains with the highest nucleotide identity against the partial 16S-rRNA gene sequence of *Nocardia* strain GGUI10#20 in the EzBioCloud analysis. Similar to the results found by Takeda *et al.* (2010), these data show that some discrepancies do exist in the phylogenetic positions of *Nocardia* species based on 16S-rRNA and *gyrB* gene sequences. Therefore, utilizing a larger set of genes would provide a more robust analysis for inferring evolutionary relationships between these organisms.

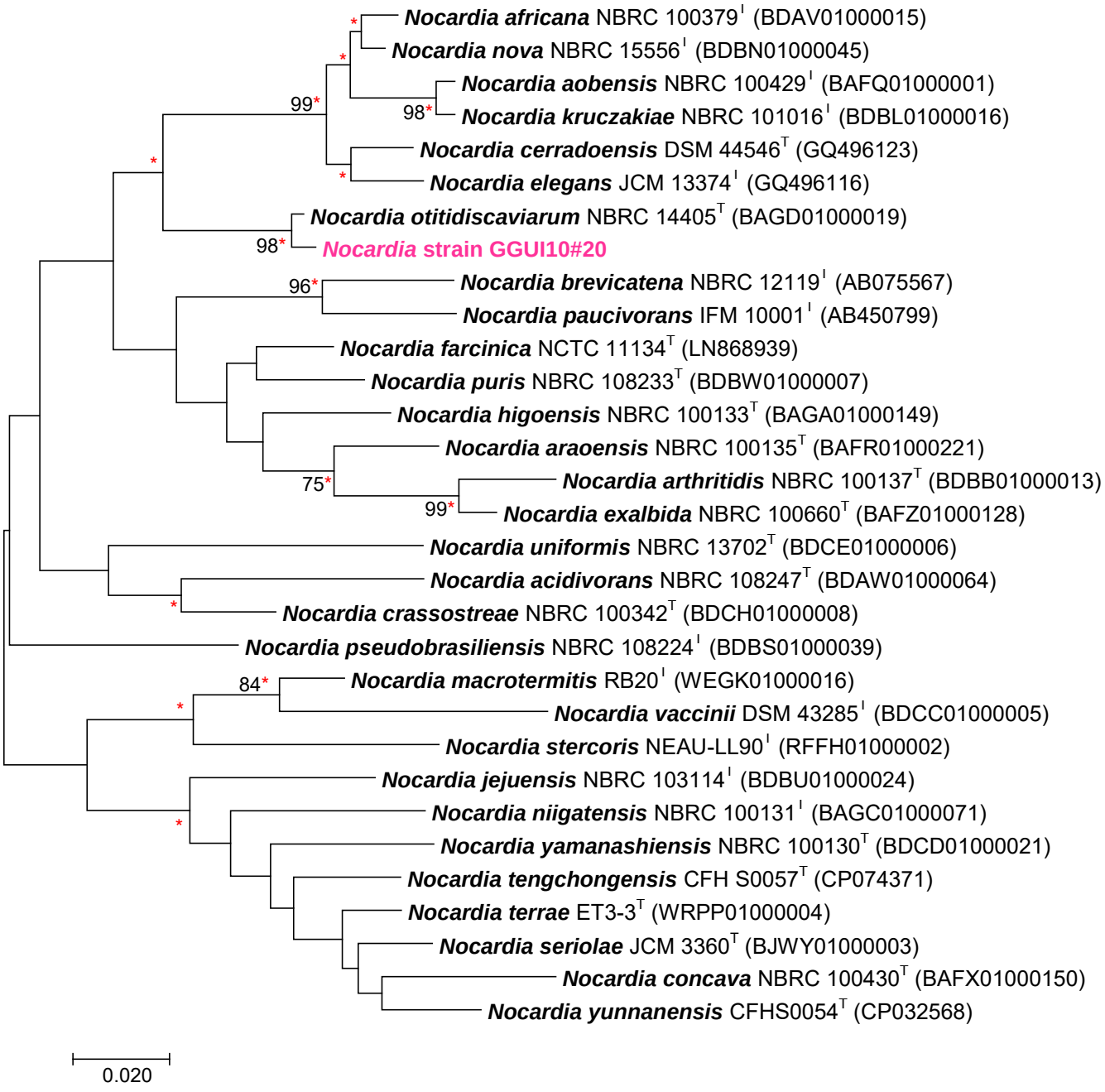

Figure 2.4 Maximum-likelihood phylogenetic tree comparing *Nocardia* strain GGUI10#20 to the type strains of the *Nocardia* species to which it is most closely related and for which *gyrB* sequences are available in the GenBank database. The tree is based on the alignment of 31 402-nt *gyrB* gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 20 nucleotide substitutions per 1000 nucleotides.

2.3.4.2 Genus *Streptomyces*

The genus *Streptomyces* is comprised of 696 species with validly published names (https://lpsn.dsmz.de/genus/streptomyces), making it the largest genus within the phylum *Actinobacteria*. The type species of this genus, *Streptomyces albus*, was first described as *Streptotrix alba* by Rossi Doria in 1891. The genus name was then emended to *Actinomyces* before being proposed as *Streptomyces* in 1943 (Waksman and Henrici, 1943). Members of the genus *Streptomyces* are Gram-positive and form extensively branched substrate and aerial mycelia. Often, colonies are initially smooth, but later develop spore mass that may appear chalky, 'fluffy' or granular (Kämpfer, 2015). *Streptomyces* species can also produce a variety of pigments which influence the colour of these colonies. Some species also produce diffusible pigments. Members of the genus *Streptomyces* are widely distributed in all environments (aquatic and terrestrial). Recently, it was shown that highly diverse populations of *Streptomyces* were even present in atmospheric precipitations (Sarmiento-Vizcaíno *et al.*, 2018). In this study, *Streptomyces* strains were isolated from all three sites, with most strains being isolated from the rocky site.

Phylogenies of the *Streptomyces* isolates were constructed based on their top hits obtained from the EzBioCloud database and the lengths of their partial 16S-rRNA gene sequences. Closely related isolates with similar top hits and/or isolates with similar lengths of their partial 16S-rRNA gene sequences were grouped in different trees. *Streptomyces* strain GGUM1#5 was not included in any phylogenetic analysis as the 16S-rRNA gene sequence obtained for this isolate was only 193 nucleotides in length. The species *Kitasatospora papulosa* NRRL B-16504^{T} (JNYQ01000038) was included in multiple analyses as it formed part of the top hits obtained on EzBioCloud for many of the *Streptomyces* isolates. *Kitasatospora* was once thought to be synonymous with *Streptomyces* (Wellington *et al.*, 1992), however phylogenetic analysis showed that members of the genus *Kitasatospora* always formed a stable monophyletic clade outside clades comprising of *Streptomyces* species (Zhang *et al.*, 1997). Therefore, the genus *Kitasatospora* has since been described as a sister taxon of the genus *Streptomyces* or a lineage that originated from within *Streptomyces*.

2.3.4.2.1 *Streptomyces* phylogenetic analysis: 16S-rRNA gene (~700 -900 nt)

Figure 2.5 shows the phylogenetic positions of *Streptomyces* strains GGUI10#6 and GGUM1#29 based on a final 16S-rRNA gene sequence alignment of 917 nucleotides. These strains were isolated from sediment from the ocean (highlighted in blue) and dry (highlighted in green) sites. In Figure 2.5, it is shown that *Streptomyces* strain GGUI10#6 clustered with the type strains of *Streptomyces synnematoformans* and *Streptomyces aculeolatus*. This association had very high bootstrap support (99%) but was not conserved in the neighbour joining and maximum parsimony trees. These results are in accordance with those obtained from EzBioCloud where the type strains of *S. synnematoformans* and *S. aculeolatus* were the top two hits of *Streptomyces* strain GGUI10#6 with 98.91% and 98.26% 16S-rRNA gene sequence similarity, respectively. These values were calculated over a length of 922 nucleotides.

Streptomyces strain GGUM1#29 grouped within a clade containing the type strains of *Streptomyces violarus*, *Streptomyces violaceus*, *Streptomyces purpurascens*, *Streptomyces koyangensis* and *Streptomyces daghestanicus* (Figure 2.5). This clade was highly supported (bootstrap value of 98%) and was conserved in the neighbor joining and maximum parsimony trees. These results are also in accordance with the EzBioCloud results where the aforementioned *Streptomyces* type strains were revealed to be the top five hits of *Streptomyces* strain GGUM1#29 with 99.57% 16S-rRNA gene sequence similarity over 935 nucleotides. Whole-genome analyses could be used to further investigate if GGUM1#29 belongs to the same species as one of its closest relatives. This can only be done for strain GGUI10#6 once a full genome sequence becomes available for the type strain of *S. synnematoformans* in the GenBank database.

Figure 2.6 shows the phylogenetic positions of *Streptomyces* strains GGUI#1, GGUM1#2, GGUI#10, GGUM1#11, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33, based on a final 16S-rRNA gene sequence alignment of 797 nucleotides. These strains were isolated from sediment from the ocean, dry and rocky (highlighted in orange) sites.

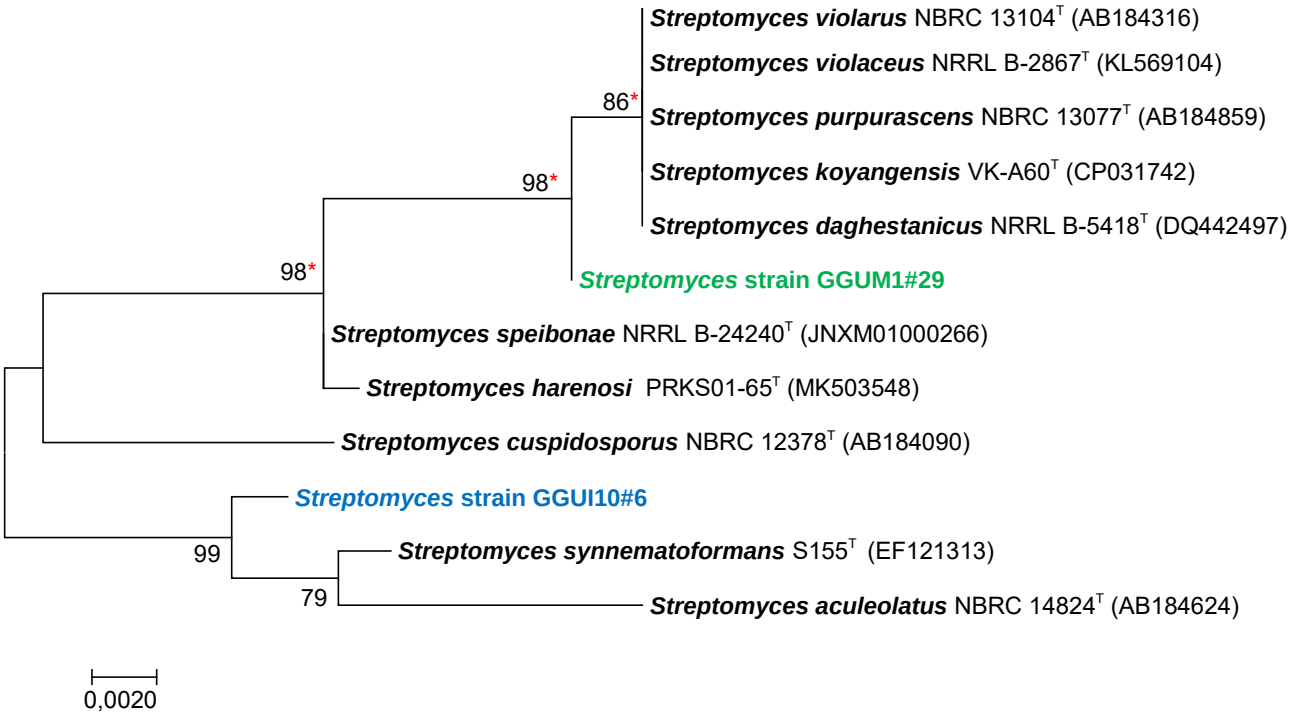

Figure 2.5. Maximum-likelihood phylogenetic tree comparing two *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the ocean (blue) and dry (green) sites. The tree is based on the alignment of 12 917-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 2 nucleotide substitutions per 1000 nucleotides.

In Figure 2.6, it is seen that *Streptomyces* strains GGUI#10, GGUM1#11, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 clustered with each other and the type strains of *Streptomyces bacillaris*, *Streptomyces cavourensis*, *Streptomyces lunaelactis*, *Streptomyces rhizosphaericola*, *Streptomyces crystallinus*, *Streptomyces araujoniae*, *Streptomyces flavovirens* and *Streptomyces nitrosporeus*. This clade was conserved in the neighbour joining and maximum parsimony trees. These results agree with the EzBioCloud results as the type strains of *S. bacillaris* and *S. cavourensis* were the top two hits of *Streptomyces* strains GGUI#10, GGUM1#11, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 with 100% 16S-rRNA gene sequence similarity over ~900 nucleotides. The type strain of *S. rhizosphaericola* was an additional top hit of strains GGUI5#30 and GGUI5#31 with 100% 16S-rRNA gene sequence similarity over 800 and 855 nucleotides respectively. All of the

Streptomyces isolates shown in Figure 2.6, except GGUI#12, are members of the PC2 group of potential clones. The top two hits of *Streptomyces* strain GGUI#12 were the type strains of *Streptomyces pratensis* and *S. nitrosporeus* with 99.77% and 99.46% 16S-rRNA gene sequence similarity, respectively, over a length of 928 nucleotides. *S. bacillaris* and *S. cavourensis* did, however, form part of the top five hits of *Streptomyces* strain GGUI#12 with 99.24% 16S-rRNA gene sequence similarity. *S. lunaelactis* also formed part of the top ten hits of *Streptomyces* strains GGUI#10, GGUM1#11, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 with a 16S-rRNA gene sequence similarity ranging from 99.13-99.89%. Interestingly, *S. crystallinus*, *S. araujoniae* and *S. flavovirens* were in the lower half of the top 30 hits of each aforementioned isolate. Based on the EzBioCloud results, these *Streptomyces* isolates were shown to be more closely related to *Streptomyces globisporus*, *Streptomyces fulvorobeus* and *Streptomyces griseus* subsp. *griseus*. The 16S-rRNA gene sequence similarity between *Streptomyces* strain GGUI#12 and strains belonging to PC2 displayed in Figure 2.6 ranged from 99.2% to 99.4% over 928 nucleotides.

Streptomyces strains GGUI#1 and GGUM1#2 clustered closely together with weak bootstrap support (67%) and identical branch lengths (Figure 2.6). This association was also seen in the neighbour joining and maximum parsimony trees. This result is not surprising as the 16S-rRNA gene sequences obtained for GGUI#1 and GGUM1#2 were identical (PC3 group of potential clones). *Streptomyces* strains GGUI#1 and GGUM1#2 also formed part of a larger clade containing the type strain of *S. microflavus* and other top hits obtained from the EzBioCloud database. The top 15 hits of these strains all had 99.89% 16S-rRNA gene sequence similarity over 840 nucleotides for strain GGUI#1 and 887 nucleotides for strain GGUM1#2.

Identical branch lengths indicate no nucleotide substitutions (i.e. identical partial 16S-rRNA gene sequences). Since this can be seen in most of the clades in Figure 2.6, it is clear there is a lack of sequence variation within the section of the 16S-rRNA gene included in this analysis. This makes it harder to infer accurate phylogenetic relationships. This trend is also seen in Figure 2.9 and Figure 2.10. A phylogenomic analysis would therefore provide more robust results, as it would be based on a collection of genes rather than a single, highly-conserved gene.

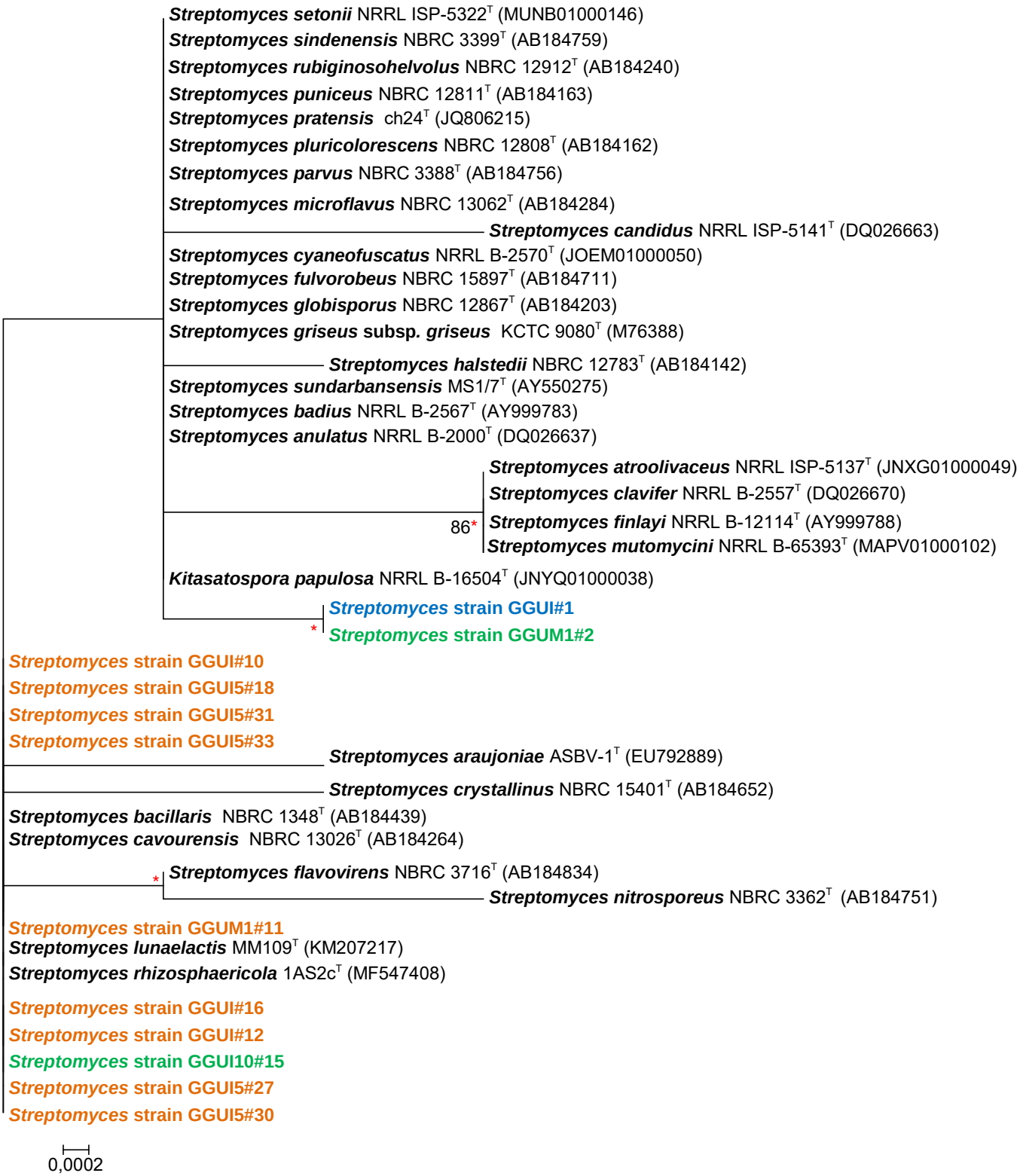

Figure 2.6. Maximum-likelihood phylogenetic tree comparing 12 *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the ocean (blue), dry (green) and rocky (orange) sites. The tree is based on the alignment of 42 797-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 0.2 nucleotide substitutions per 1000 nucleotides.

Figure 2.7 shows the phylogenetic positions of *Streptomyces* strains GGUI#3, GGUI5#25 and GGUM1#26 based on a final 16S-rRNA gene sequence alignment of 781 nucleotides. Figure 2.7 shows *Streptomyces* strain GGUI#3 clustering with *Streptomyces apocyni* TRM 66233^{T} (MN565731) with very high bootstrap support (98%). This association was conserved in the neighbor joining and maximum parsimony trees. This is in agreement with the EzBioCloud results, as the type strain of *S. apocyni* was also the top hit for *Streptomyces* strain GGUI#3 with 99.52% 16S-rRNA gene sequence similarity over 841 nucleotides.

In Figure 2.7, *Streptomyces* strain GGUI5#25 formed a branch between the type strain of *S. lunaelactis* and a cluster containing its top two hits from the EzBioCloud database, the type strains of *Streptomyces kurssanovii* and *Streptomyces xantholiticus* (both 99.44% 16S-rRNA gene sequence similarity over 885 nucleotides) as well as its next closest relative, the type strain of *Streptomyces peucetius* (99.32% 16S-rRNA gene sequence similarity). The type strain of *S. lunaelactis* was 5th on the list of top hits of strain GGUI5#25 with a 16S-rRNA gene sequence similarity of 99.32%. The aforementioned clade was supported by a low bootstrap value of 50% and was not conserved in the neighbour joining and maximum parsimony trees. Whole-genome analyses could be used to further investigate whether *Streptomyces* strains GGUI#3 and GGUI5#25 belong to the same species as one of their closest relatives.

Streptomyces strain GGUM1#26 is shown to group closely with the type strain of *S. chumphonensis* (Figure 2.7). This species was also the top hit in the EzBioCloud analysis with 99.55% 16S-rRNA gene sequence similarity over 895 nucleotides (Table 2.2). The aforementioned association was supported by a very high bootstrap value of 96% and was also seen in the neighbour joining and maximum parsimony trees. *Streptomyces* strain GGUM1#26 also formed part of a bigger monophyletic clade including the type strains of its other top hits *Streptomyces xinghaiensis*, *Streptomyces alkaliterrae* and *Streptomyces diastatochromogenes* with 99.22%, 99.11% and 98.88% 16S-rRNA gene sequence similarity, respectively. This clade is supported by a very high bootstrap value of 99% but was not seen in the neighbour joining and maximum parsimony trees.

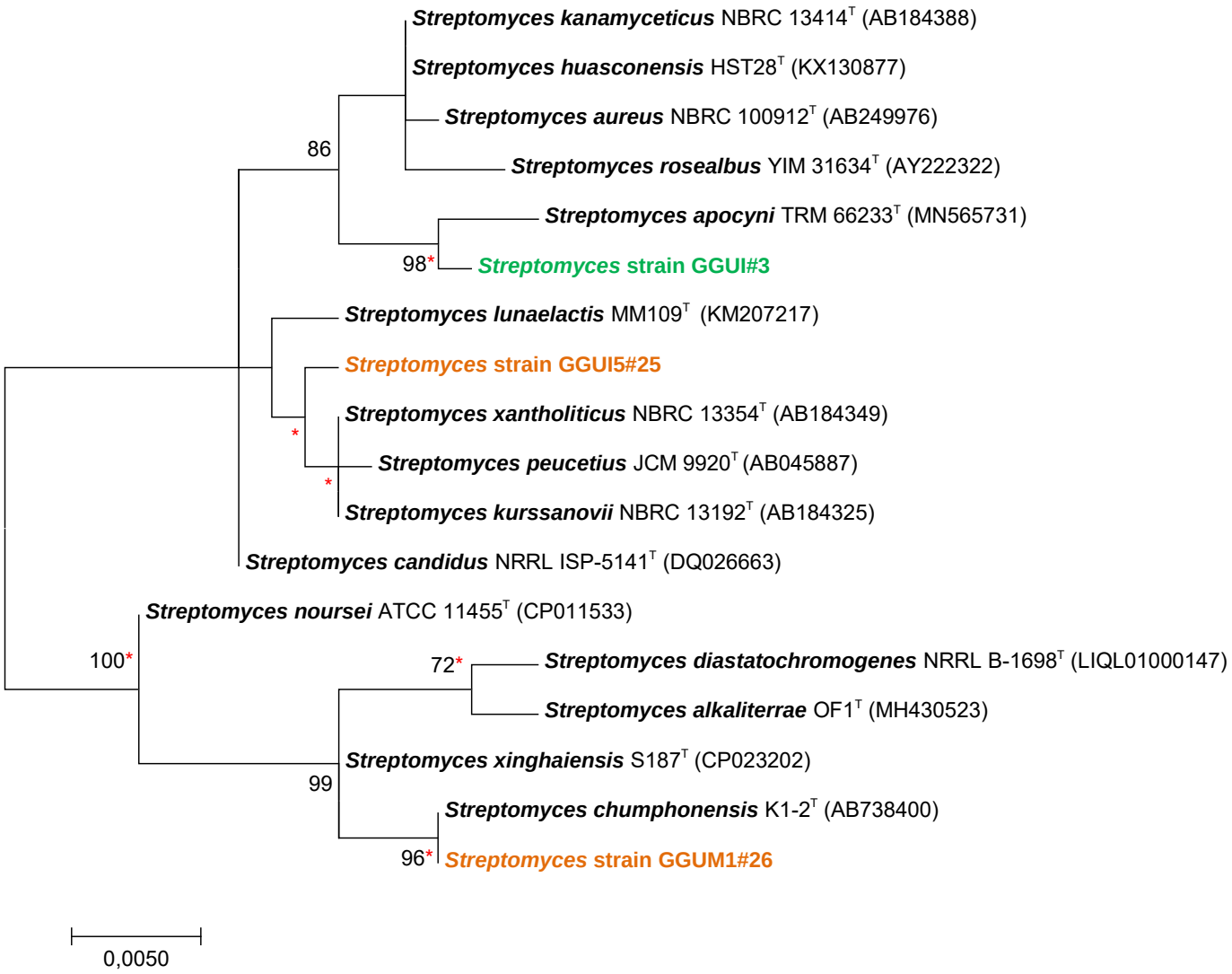

Figure 2.7. Maximum-likelihood phylogenetic tree comparing three *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the dry (green) and rocky sites (orange). The tree is based on the alignment of 18 781-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 5 nucleotide substitutions per 1000 nucleotides.

2.3.4.2.2 *Streptomyces* phylogenetic analysis: 16S-rRNA gene (~500 nt)

Figure 2.8 shows the phylogenetic positions of *Streptomyces* strains GGUI#7, GGUI10#19, and GGUI5#24, based on a final 16S-rRNA gene sequence alignment of 573 nucleotides. In Figure 2.8, *Streptomyces* strain GGUI#7 is seen grouping within a clade containing the type strains of *Streptomyces umbrinus*, *Streptomyces phaeochromogenes*, *Streptomyces ederensis*, *Streptomyces dioscori* and *Streptomyces bottropensis*. This clade has very strong bootstrap support (100%) and was conserved in the neighbor joining and maximum parsimony trees. This indicates that this is reliable representation of the phylogenetic relationships between these strains, though utilizing a larger set of genes would provide a more robust analysis. These results are also in agreement with the EzBioCloud results where it was revealed that the type strains of *S. umbrinus*, *S. phaeochromogenes* and *S. ederensis* were the top three hits of *Streptomyces* strain GGUI#7 with 99.32% 16S-rRNA gene sequence similarity over 591 nucleotides. *S. dioscori* and *S. bottropensis* were the next closest relatives of strain GGUI#7 with 99.15% 16S-rRNA gene sequence similarity. Whole-genome sequencing and analyses could be performed to gain definitive data on the relationship between strain GGUI#7 and its closest relatives.

Streptomyces strains GGUI10#19 and GGUI5#24 have identical 16S-rRNA gene sequences and form part of the PC1 potential clone group. In Figure 2.8, they can be seen clustering with the type strain of *S. chumphonensis* which, according to the EzBioCloud results, is their closest relative (100% 16S-rRNA gene sequence similarity). This similarity was calculated over 573 nucleotides for strain GGUI5#24 and 597 nucleotides for strain GGUI10#19. The aforementioned clade was supported by a very high bootstrap value of 95% and grouped within a larger clade containing the other top hits of *Streptomyces* strains GGUI10#19 and GGUI5#24. Analysis of the partial *gyrB* gene sequences of these strains and the other member of PC1, *Streptomyces* strain GGUIM1#26, was done to further investigate the relationship between them and their closest relatives (section 2.3.5.1).

Figure 2.9 shows the phylogenetic positions of *Streptomyces* strains GGUM1#4, GGUI10#8, GGUI#9, GGUI#23 and GGUI5#28, based on a final 16S-rRNA gene sequence alignment of 556 nucleotides. *Streptomyces* strains GGUM1#4, GGUI10#8, GGUI#9, GGUI#23, and GGUI5#28 clustered in a large clade with all their closest relatives (according to the EzBioCloud results). This clade is supported by a high bootstrap value (89%) and was conserved in the neighbour joining and maximum parsimony trees. The clustering in the

aforementioned clade is similar to the clades seen in Figure 2.6. However, since the tree in Figure 2.9 was based on sequences containing less than a third of the full 16S-rRNA gene, the sequence variation between the strains in this analysis is very low. This has caused the branch length of most strains in Figure 2.9 to be identical.

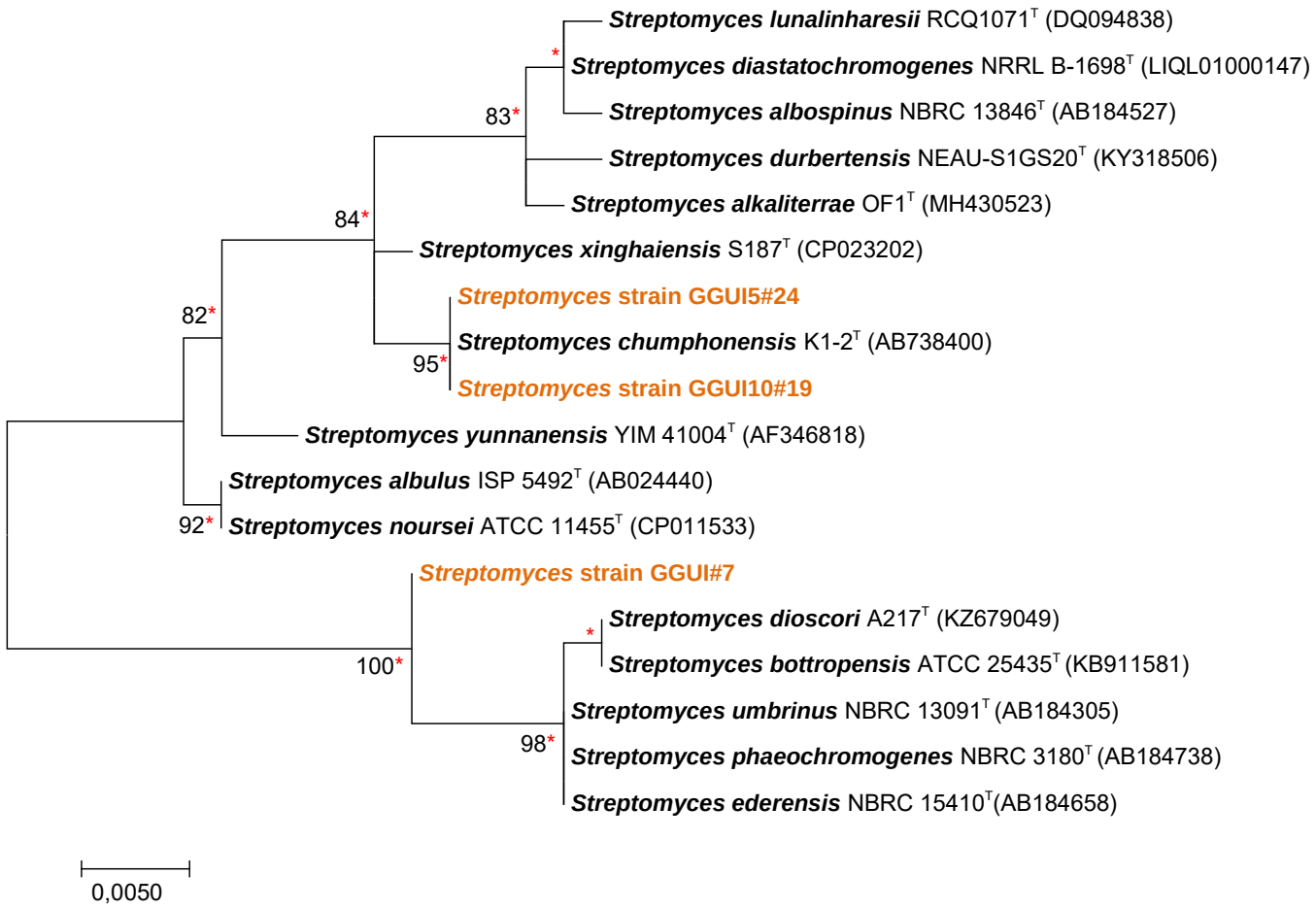

Figure 2.8. Maximum-likelihood phylogenetic tree comparing three *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the dry rocky site. The tree is based on the alignment of 18 573-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 5 nucleotide substitutions per 1000 nucleotides

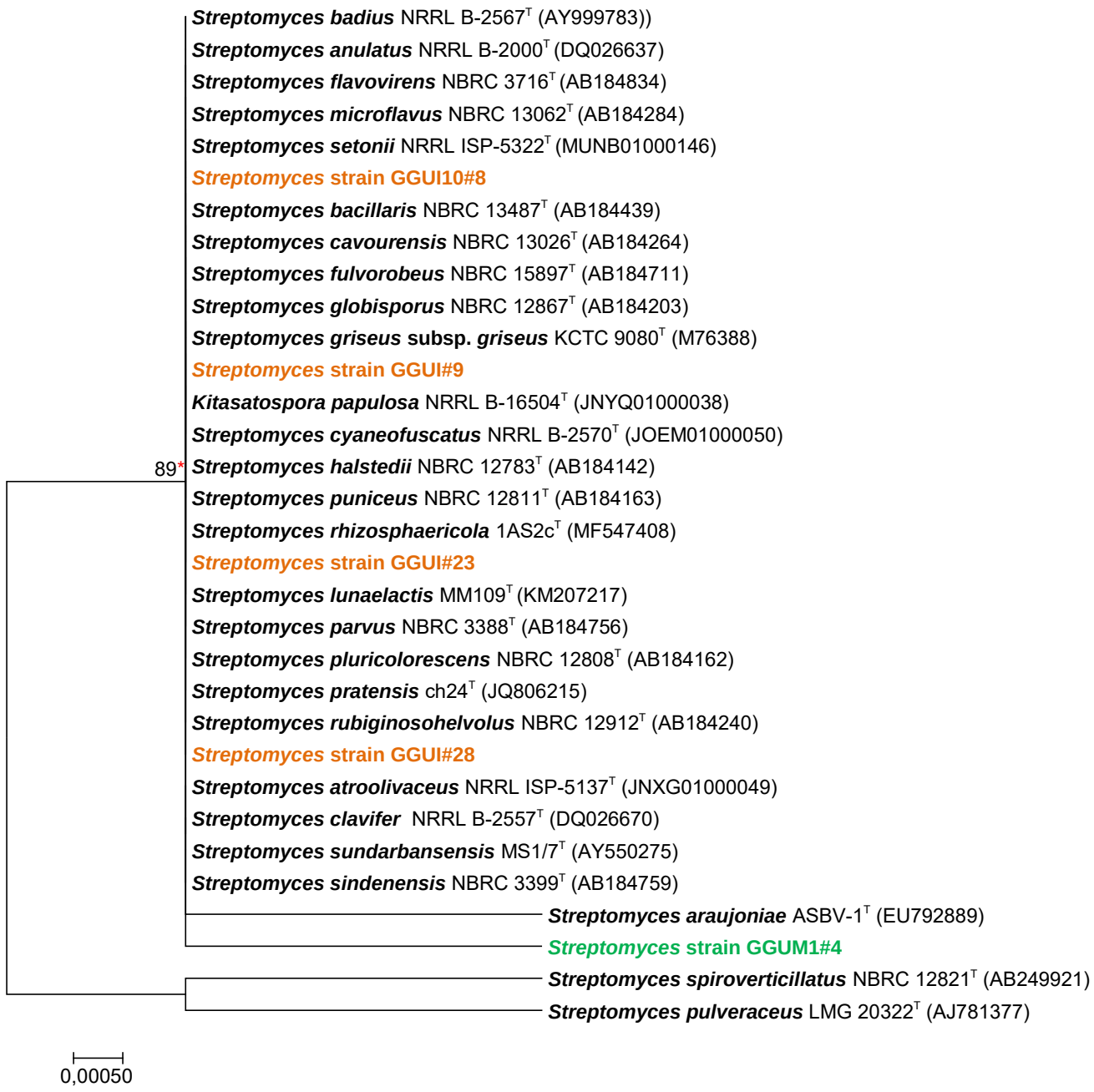

Figure 2.9. Maximum-likelihood phylogenetic tree comparing five *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the dry (green) and rocky sites (orange). The tree is based on the alignment of 32 556-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 0.5 nucleotide substitutions per 1000 nucleotides.

Strains GGUI10#8, GGUI#9 and GGUI5#28 form part of the PC2 potential clone group and have *S. griseus* subsp. *griseus*, *S. globisporus*, *S. bacillaris*, *S. cavourensis*, *S. fulvorobeus* and *S. rhizosphaericola* as their top six hits with 100% 16S-rRNA gene sequence similarity over 558, 662 and 614 nucleotides, respectively. Although the top three hits of *Streptomyces* strain GGUI#23 are also *S. bacillaris*, *S. cavourensis* and *S. rhizosphaericola* (99.85% 16S-rRNA gene sequence similarity over 646 nucleotides), this strain does not form part of PC2 because its 16S-rRNA gene sequence is not identical to those of the other isolates in this group. The 16S-rRNA gene sequence similarity between *Streptomyces* strain GGUI#23 and the strains in PC2 ranged from 99.4% to 99.8%. *Streptomyces* strains GGUM1#4 was the only isolate to display a branch length not identical to the other isolates in Figure 2.9. According to the EzBioCloud results, *Streptomyces* strain GGUM1#4 has 19 top hits all with 99.39% 16S-rRNA gene sequence similarity over 655 nucleotides. All of these species are present in the aforementioned clade. The 16S-rRNA gene sequence similarity between *Streptomyces* strain GGUM1#4 and the other marine isolates in Figure 2.9 ranged from 99.1% to 99.5% over 655 nucleotides.

Figure 2.10 displays the positions of all isolates shown in Figure 2.6 and Figure 2.9 in one phylogenetic tree. This was done to display all closely related isolates (including all strains in PC2 and PC3) in one figure even though this meant using shorter sequences than were used in Figure 2.6. The same was done for all strains in PC1 by combining all isolates shown in Figure 2.7 and Figure 2.8 in one phylogenetic tree (Figure 2.11). Both trees confirm the results in previous figures. All isolates displayed in Figure 2.10 are shown to be clustered together. Furthermore, the branch length for each isolate, besides *Streptomyces* strain GGUM1#4, was identical. In Figure 2.11, *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUIM1#26 (PC1) clustered together with identical branch lengths. This association was supported by a high bootstrap value of 87% and was conserved in the neighbour joining and maximum parsimony trees.

As mentioned in section 2.1, the 16S-rRNA gene has a slow evolutionary rate which complicates comparison between closely related strains. This is particularly true in the case of the genus *Streptomyces* as closely related *Streptomyces* species have been shown to exhibit highly similar or identical 16S-rRNA gene sequences (Law *et al.*, 2018). There are three variable regions within the 16S-rRNA gene considered to be diagnostic for the *Streptomyces* genus. These include the nucleotide positions 158-203 (gamma region), 982-998 (alpha

region) and 1102-1122 (beta region) (Stackebrandt *et al.*, 1991). The gamma region is the most variable and, thus, is important for species discrimination within the genus *Streptomyces*. The 16S-rRNA sequences obtained in this study fall within the region containing the nucleotide positions 511-1447; therefore, the gamma region could not be compared between the isolates and their closest relatives. This could explain the low resolution observed in the phylogenetic analyses. The resolution could be increased by utilising the full 16S-rRNA gene sequence. However, based on the literature, it is clear that the 16S-rRNA gene overall provides low resolution as a phylogenetic marker for *Streptomyces* species. Thus, methods with higher resolution (such as phylogenomic analyses) are crucial to increase the robustness of future classifications and evaluate the evolutionary relationships among closely related *Streptomyces* species.

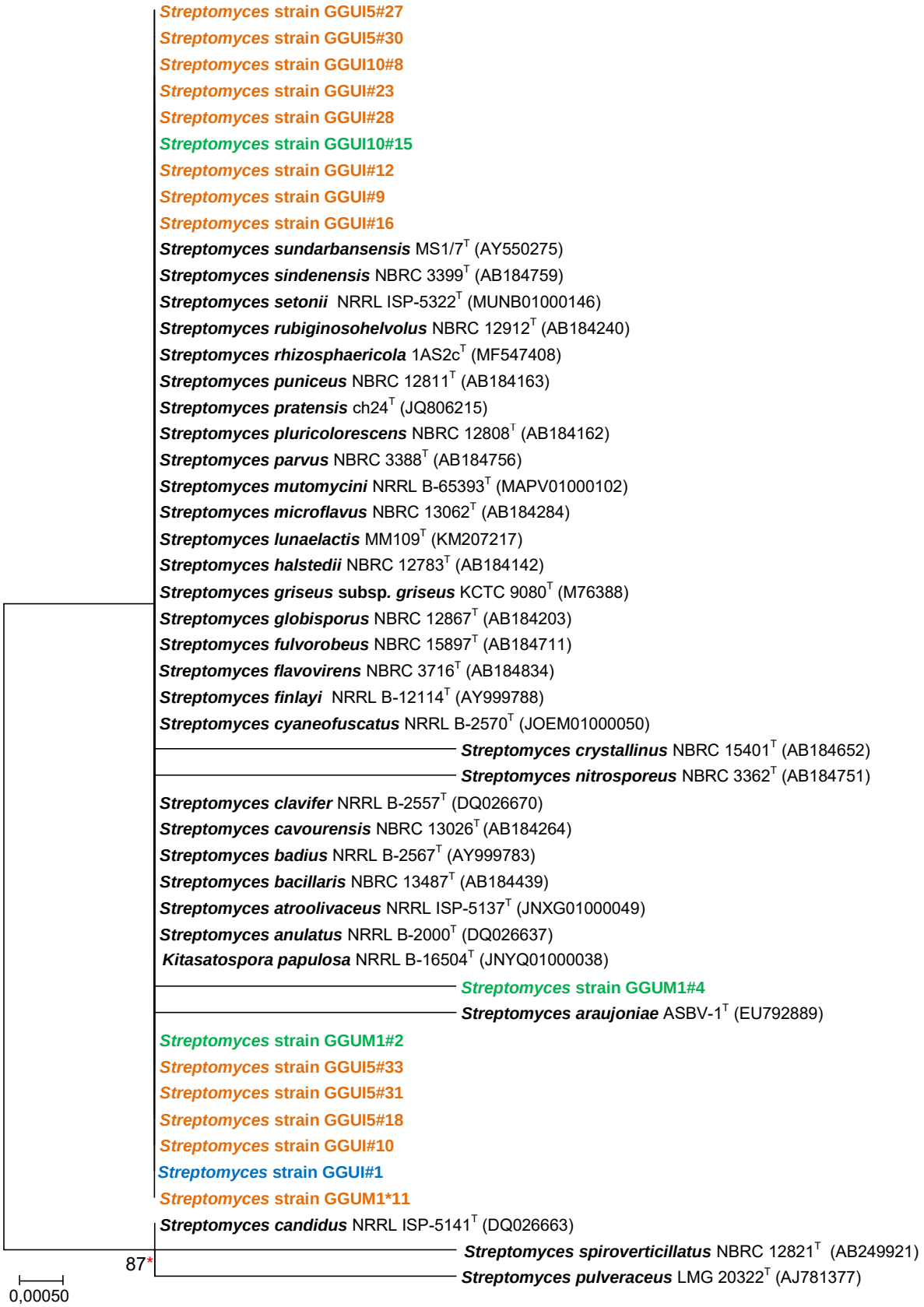

Figure 2.10. Maximum-likelihood phylogenetic tree comparing 17 *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the ocean (blue), dry (green) and rocky sites (orange). The tree is based on the alignment of 49 556-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 0.5 nucleotide substitutions per 1000 nucleotides.

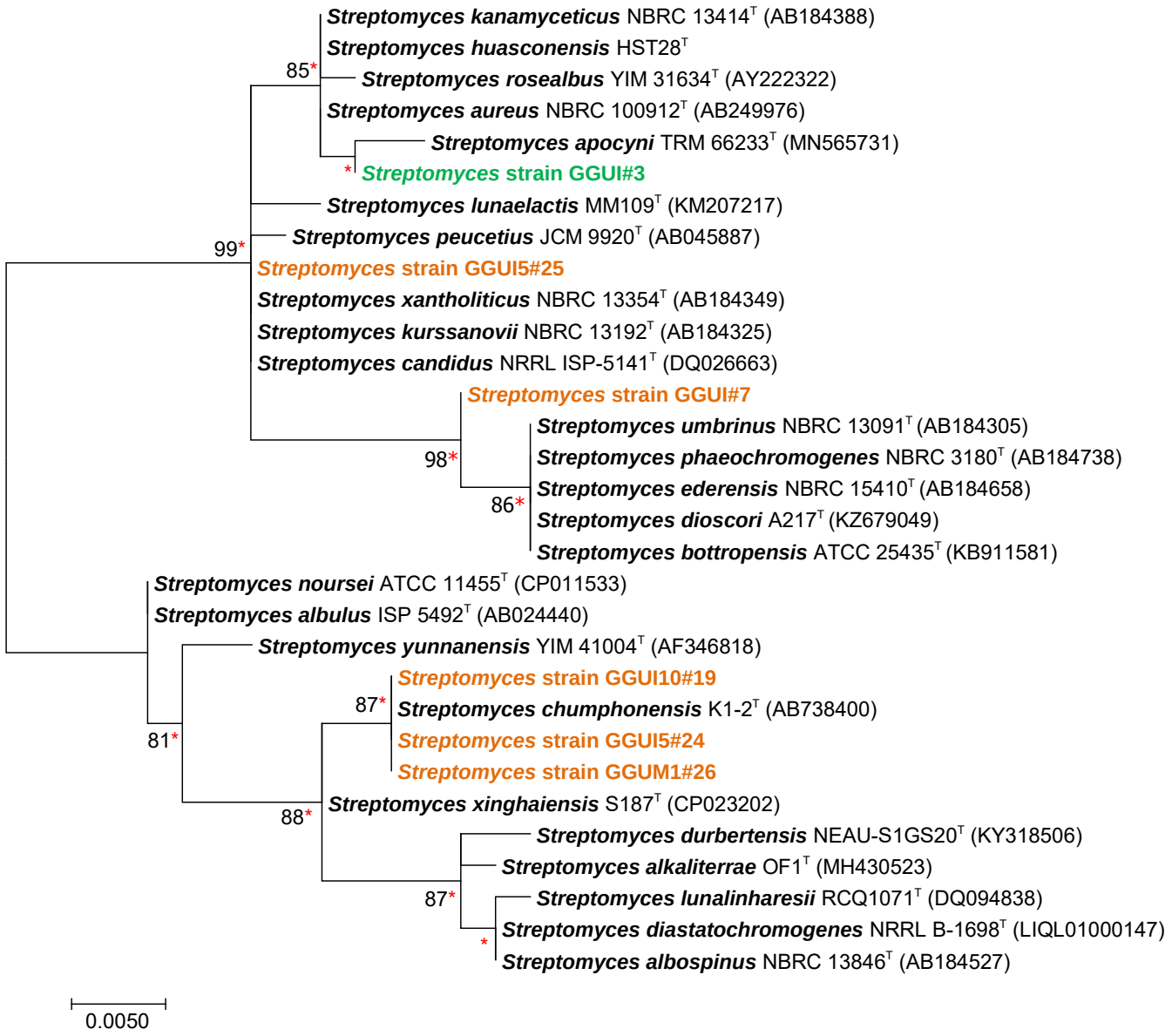

Figure 2.11. Maximum-likelihood phylogenetic tree comparing six *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related based on the top hits obtained from the EzBioCloud database. These strains were isolated from the dry (green) and rocky sites (orange). The tree is based on the alignment of 31 543-nt 16S-rRNA gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 5 nucleotide substitutions per 1000 nucleotides

2.3.5 Analysis of potential clones

2.3.5.1 *Streptomyces chumphonensis* group (PC1)

The isolates belonging to PC1 are *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUM1#26. To further investigate the relationship between these strains, partial *gyrB* sequences were obtained of 349, 301 and 353 nucleotides in length, respectively (Table 2.2). This is only ~13% of the full-length *gyrB* gene. The sequences were aligned, edited so that they were the same length and a homology matrix was generated to assess the *gyrB* sequence similarity between the strains of PC1 (Table 2.3). A phylogenetic tree was also built using the partial *gyrB* gene sequences to analyse the phylogeny of the strains in PC1 in relation to their closest relatives (Figure 2.8).

Table 2.3. Homology matrix comparing 290-nt *gyrB* gene sequences of isolates in PC1.

Isolate no.	#19	#24	#26
#19	100%		
#24	99.3%	100%	
#26	99.7%	99.7%	100%

Table 2.3 shows that when comparing the *gyrB* sequences obtained for strains GGUI10#19, GGUI5#24 and GGUM1#26, each strain had at least one variable nucleotide position indicating that none of these strains could be clones of each other. If these strains were clones, one would expect the *gyrB* gene sequences to be identical over any length compared. There is phenotypic evidence to support these results as the aforementioned strains all display slightly different colony characteristics (Table A1, Figure A2). In accordance with the results in Table 2.3, strains GGUI10#19, GGUI5#24 and GGUM1#26 are shown to cluster together (different branch lengths) with strains GGUI10#19 and GGUM1#24 appearing to be least closely related to each other (Figure 2.12). The aforementioned clade has very high bootstrap support (97%) and was conserved in the neighbour joining and maximum parsimony trees. In Figure 2.12, strains GGUI10#19, GGUI5#24 and GGUM1#26 are also shown to group with the type strain of *S. chumphonensis* with moderate bootstrap support (82%). This association is seen in the neighbour joining and maximum parsimony trees. This is in agreement with the 16S-rRNA gene EzBioCloud results, but not the *blastn* results. The *blastn* analysis revealed the top hit of strains GGUI10#19 and GGUI5#24 to be the type strain of *Streptomyces roseoverticillatus* with 92.83% and 92.98% *gyrB* gene sequence similarity, respectively, and

the top hit of strain GGUM1#26 to be the type strain of *Streptomyces laurentii* with 92.06% *gyrB* gene sequence similarity (Table 2.2). The *gyrB* sequence of *S. chumphonensis* K1-2^{T} (JACXYU010000015) is publicly available in the GenBank database. However, it forms part of a whole genome shotgun sequence meaning it is not included in the database used for a standard *blastn* analysis (i.e. the nucleotide collection database (nr/nt)). When using the whole-genome shotgun contig database (wgs) in the *blastn* analysis, the top hit for strains GGUI10#19, GGUI5#24 and GGUM1#26 was revealed to be *S. chumphonensis* K1-2^{T} with 93.52%, 94.31% and 94.08% *gyrB* gene sequence similarity, respectively.

Although it has been revealed that the strains in PC1 are not clones, they are still very closely related. Furthermore, they were isolated from the same sediment sample from the rocky site. For these reasons, it is possible that isolates GGUI10#19, GGUI5#24 and GGUM1#26 are different strains of the same species. The *gyrB* gene sequence similarity between the strains in PC1 and their closest phylogenetic relative, *S. chumphonensis* K1-2^{T}, is relatively low (~94%). Based on existing criteria, this suggests the strains in this group could represent a novel species. However, since only a small portion of the *gyrB* gene was used in this analysis and there is also a small possibility that these results could be a consequence of horizontal gene transfer, only whole-genome comparison would provide a definitive answer on whether the strains in PC1 and *S. chumphonensis* belong to the same species.

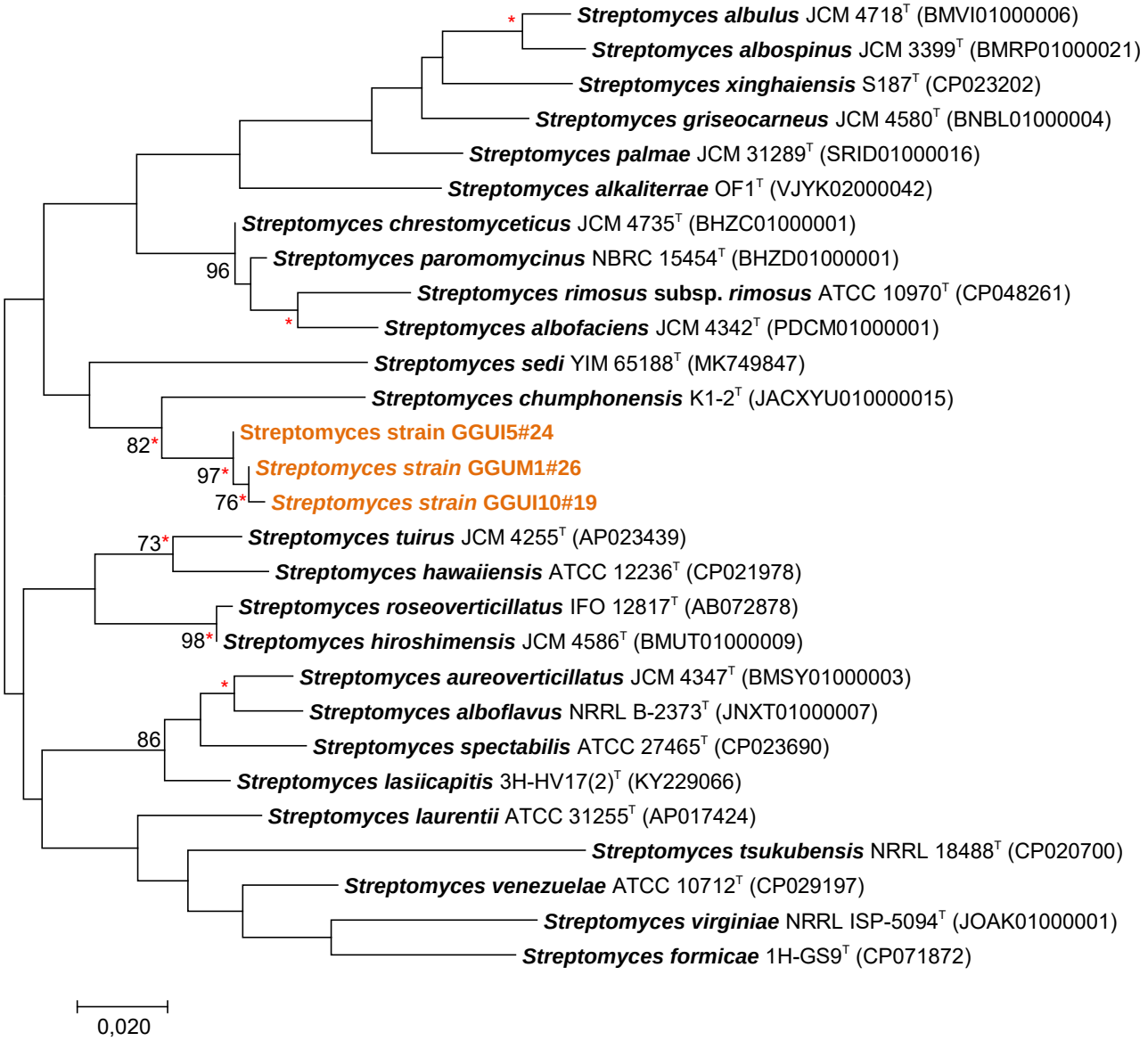

Figure 2.12 Maximum-likelihood phylogenetic tree comparing three *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related and for which *gyrB* sequences are available. The *Streptomyces* species were chosen based on the top hits obtained from the EzBioCloud and GenBank databases. The tree is based on the alignment of 28 289-nt *gyrB* gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 20 nucleotide substitutions per 1000 nucleotides.

2.3.5.2 *Streptomyces bacillaris* group (PC2)

The isolates belonging to PC2 are *Streptomyces* strains GGUI10#8, GGUI#9, GGUI#10, GUUM1#11, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33. Similar to the strains in PC1, a phylogenetic tree (Figure 2.13) and a homology matrix (Table 2.4) was generated to compare the *gyrB* gene sequences obtained for these strains. The *gyrB* sequences obtained for the strains in PC2 were 618 to 680 nucleotides in length (Table 2.2).

Table 2.4. Homology matrix comparing 612-nt *gyrB* gene sequences of isolates in PC2.

Isolate no.	#8	#9	#10	#11	#15	#16	#18	#27	#28	#30	#31	#33
#8	100%											
#9	100%	100%										
#10	99.7%	99.7%	100%									
#11	99.7%	99.7%	100%	100%								
#15	99.7%	99.7%	100%	100%	100%							
#16	100%	100%	99.7%	99.7%	99.7%	100%						
#18	100%	100%	99.7%	99.7%	99.7%	100%	100%					
#27	100%	100%	99.7%	99.7%	99.7%	100%	100%	100%				
#28	99.7%	99.7%	100%	100%	100%	99.7%	99.7%	99.7%	100%			
#30	100%	100%	99.7%	99.7%	99.7%	100%	100%	100%	99.7%	100%		
#31	100%	100%	99.7%	99.7%	99.7%	100%	100%	100%	99.7%	100%	100%	
#33	100%	100%	99.7%	99.7%	99.7%	100%	100%	100%	99.7%	100%	100%	100%

Both Table 2.4 and Figure 2.13 indicate that PC2 can be further divided into two separate groups of potential clones. The first group contains GGUI#10, GUUM1#11, GGUI10#15 and GGUI#28 which have 100% *gyrB* gene sequence similarity (Table 2.4). These strains also clustered together (with identical branch lengths) within a larger clade containing the other strains in PC2 (Figure 2.13). The clade containing the strains GGUI#10, GUUM1#11, GGUI10#15 and GGUI#28 has moderate bootstrap support (84%) and was also seen in the neighbour joining and maximum parsimony trees. These results provide more evidence that the strains GGUI#10, GUUM1#11, GGUI10#15 and GGUI#28 could be clones, as their partial *gyrB* gene sequences are identical. However, it is important to note that the *gyrB* sequences obtained for these strains make up only ~27% of the full-length actinobacterial *gyrB* gene and more sequence variation might be observed when comparing longer or full-length *gyrB* gene sequences. Furthermore, the phenotypic differences between these strains cannot be ignored. For example, strain GGUI#28 has bright orange substrate mycelium and

produces a faint brown pigment whereas the rest of the strains have brownish substrate mycelium and produce pigments in various darker shades of brown. Strain GGUI10#15 has grey/cream spore mass whereas the other strains have white/cream spore mass. Details on the different phenotypic features of the other strains can be found in Table A1 and Figure A3 contains visual evidence of these findings. Similar to the strains in PC1, strains GGUI#10 and GUUM1#11 were also isolated from the same sediment sample taken from the rocky site. These strains were also the least phenotypically unique and therefore, they could be clones or, if the analysis of longer *gyrB* gene sequences were to reveal sequence differences, they could be different strains of the same species. *Streptomyces* strain GGUI10#15 was isolated from another site (dry), though this alone would not rule out the possibility of it also being of the same species as isolates GGUI#10, GUUM1#11 and GGUI#28, as the sites were only 5m apart.

The other group within PC2 consists of *Streptomyces* strains GGUI10#8, GGUI#9, GGUI#16, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33. Similar to the first group, these strains exhibited 100% *gyrB* gene sequence similarity (Table 2.4) and clustered together with identical branch lengths. This cluster had moderate bootstrap support (83%) and was also conserved in the neighbour joining and maximum parsimony trees. Phenotypically, strains GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 are very similar (almost identical) (Table A1) and since they were all isolated from the same sediment sample from the rocky site, these strains are likely to be clones. However, more sequence information would be able to provide a definitive classification. The same can be said about strains GGUI10#8 and GGUI5#18, as they were also isolated from the same sediment sample from the rocky site and have very similar phenotypic characteristics. Strains GGUI#9 and GGUI#16 are similar to strain GGUI#28 as they both display phenotypic characteristics that are unique among all the strains in PC2. Strain GGUI#9 displays dark brown wrinkly colonies with a white spore mass and also produces a dark brown diffusible pigment, while strain GGUI#16 displays pale beige wrinkly colonies with little to no spore mass (Figure A3).

In Figure 2.13, the *Streptomyces* isolates belonging to PC2 are shown to cluster with the type strains of *S. rhizosphaericola* and *S. cavourensis* with very high bootstrap support (91%). This clade is conserved in the neighbour joining and maximum parsimony trees. These results are in accordance with the results obtained from the EzBioCloud and *blastn* sequence analyses. When using the nr/nt database, the type strain of *S. cavourensis* was the top hit for

all the strains in PC2 with *gyrB* sequence similarities ranging from 98.23 – 98.51% over 618 to 680 nucleotides (Table 2.2). When using the wgs database, the type strain of *S. rhizosphaericola* was found to have the next highest *gyrB* sequence similarity (after *S. cavourensis)* among the strains included in the PC2 phylogenetic analysis. In Figure 2.13, the branch lengths of the *Streptomyces* isolates and their closest relatives are also no longer identical (as was the case in the 16S-rRNA gene tree; Figure 2.10). This illustrates the benefit of using genes with higher mutation rates than the 16S-rRNA gene to increase the resolution of phylogenetic analyses. The robustness of this analysis could also be increased if a longer section of the *gyrB* gene was used.

2.3.5.3 *Streptomyces microflavus* group (PC3)

Streptomyces strains GGUI#1 and GGUM1#2 were labelled as potential clones because their partial 16S-rRNA gene sequences were identical. However, analysis of the partial *gyrB* gene sequences obtained for these strains showed otherwise. The *gyrB* gene sequence similarity of strains GGUI#1 and GGUM1#2 was 97.40% over 616 nucleotides. The phenotypic characteristics of strains GGUI#1 and GGUM1#2 are also significantly different (Table A1). Furthermore, the *gyrB* phylogenetic analysis of these strains revealed that they group in separate clusters within a larger monophyletic clade (Figure 2.13). Strain GGUI#1 grouped with the type strains of *S. microflavus and S. cyaneofuscatus* with low bootstrap support (49%). Strain GGUM1#2 grouped with the type strains of *S. sindenensis*, *S. anulatus* and *S. griseus* subsp. *griseus* with very low bootstrap support (13%). Only the clade containing strain GGUI#1 was conserved in the neighbour joining and maximum parsimony trees. The phylogenetic results were not in agreement with the *blastn* results, as both strains GGUI#1 and GGUM1#2 grouped away from their top hits, *S. griseus* subsp. *griseus* and *S. globisporus* respectively (Table 2.2). This again, is due to nr/nt database bias, as when the wgs database was used in the *blastn* analysis, the top hits (among the strains used in the PC3 phylogenetic analysis) for strains GGUI#1 and GGUM1#2 were the type strains of *S. cyaneofuscatus* (98.22% *gyrB* gene sequence similarity over 616 nucleotides) and *S. sindenensis* (98.89% *gyrB* gene sequence similarity over 631 nucleotides), respectively.

Overall, the database bias shown in this section illustrates how taxonomic inferences can also be affected when relying on particular reference databases. In future, it would be important to make sure that analyses are more inclusive of the vast amount of sequence data currently being made available.

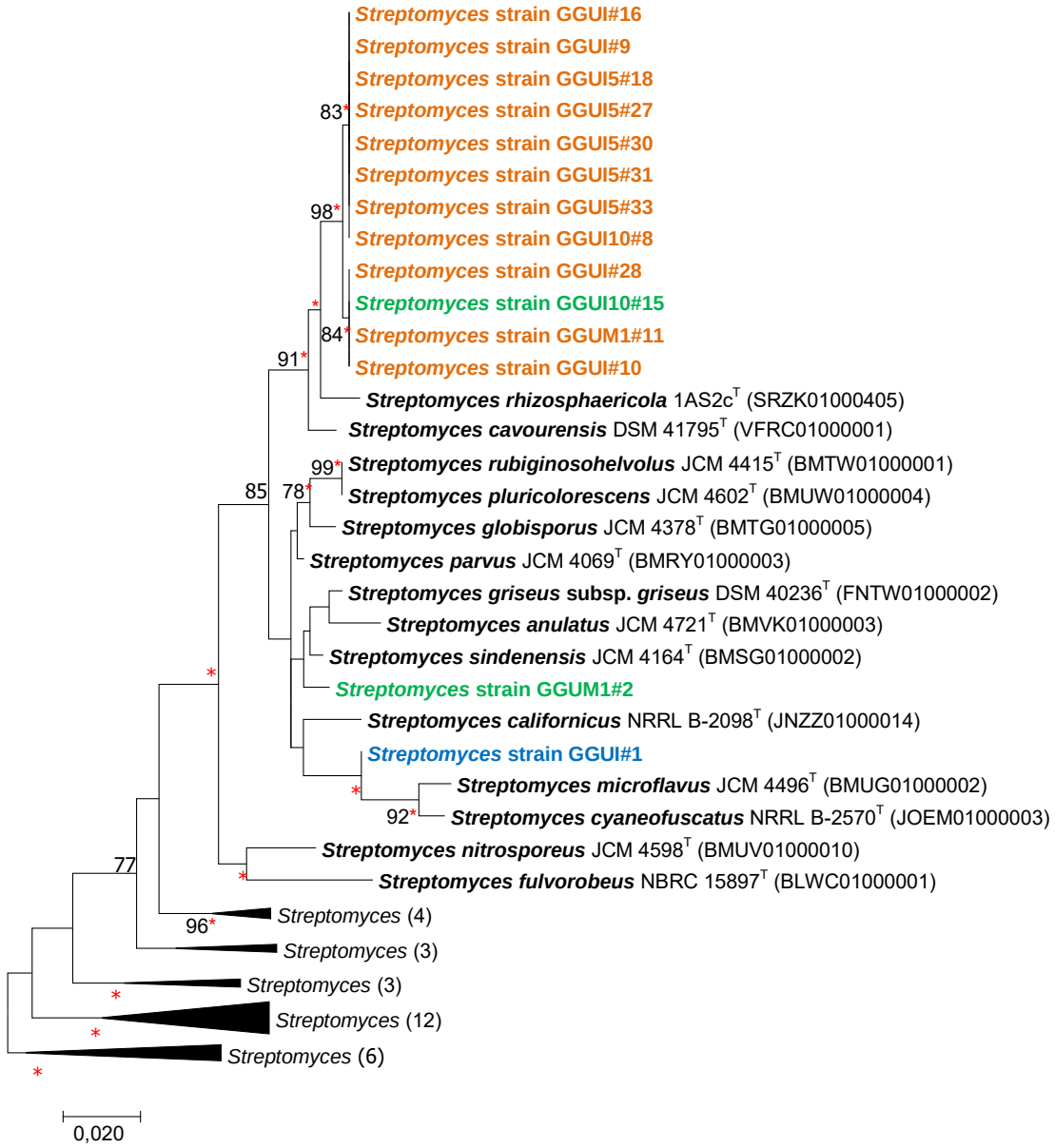

Figure 2.13. Maximum-likelihood phylogenetic tree comparing 14 *Streptomyces* isolates to the type strains of the *Streptomyces* species to which they are most closely related and for which *gyrB* sequences are available. The *Streptomyces* species were chosen based on the top hits obtained from the EzBioCloud and GenBank databases. The tree is based on the alignment of 56 612-nt *gyrB* gene sequences. Numbers at the nodes show the percentage bootstrap support for each node (only values ≥70% are shown). Asterisks (*) indicate nodes that were also obtained in phylogenetic trees constructed with the neighbour joining and maximum parsimony algorithms. The scale bar indicates 20 nucleotide substitutions per 1000 nucleotides.

2.4 REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *Journal of molecular biology*, *215*(3), pp. 403–410.

Anderson, A. S. and Wellington, E. M. H. (2001) 'The taxonomy of *Streptomyces* and related genera', *International Journal of Systematic and Evolutionary Microbiology*, *51*, pp. 797–814.

Bérdy, J. (2012) 'Thoughts and facts about antibiotics: Where we are now and where we are heading', *The Journal of Antibiotics*, *65*, pp. 385–395.

Bredholdt, H., Galatenko, O.A., Engelhardt, K., Fjærvik, E., Terekhova, L.P. and Zotchev, S.B. (2007) 'Rare actinomycete bacteria from the shallow water sediments of the Trondheim fjord, Norway: Isolation, diversity and biological activity', *Environmental Microbiology*, *9*(11), pp. 2756–2764.

Castellani, A. and Chalmers, A.J. (1919) *Manual of tropical medicine, Tindall Baillière Ed.*

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., da Costa, M.S., Rooney, A.P., Yi, H., Xu, X.W., De Meyer, S. and Trujillo, M.E. (2018) 'Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes', *International Journal of Systematic and Evolutionary Microbiology*, *68*(1), pp. 461–466.

Conville, P.S., Brown-Elliott, B.A., Smith, T. and Zelazny, A.M. (2018) 'The Complexities of *Nocardia* Taxonomy and Identification', *Journal of Clinical Microbiology*, 56(1).

Cook, A. E. and Meyers, P. R. (2003) 'Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S rRNA gene restriction fragment patterns', *International Journal of Systematic and Evolutionary Microbiology*, *53*(6), pp. 1907–1915.

Edgar, R. C. (2004) 'MUSCLE: Multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Research*, *32*(5), pp. 1792–1797.

Everest, G. J. and Meyers, P. R. (2009) 'The use of *gyrB* sequence analysis in the phylogeny of the genus *Amycolatopsis*', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *95*(1), pp. 1–11.

Faisal, M., Diamanka, A., Loch, T.P., LaFrentz, B.R., Winters, A.D., García, J.C. and Toguebaye, B.S. (2017) 'Isolation and characterization of *Flavobacterium columnare* strains infecting fishes inhabiting the Laurentian Great Lakes basin', *Journal of Fish Diseases*, *40*(5), pp. 637–648.

Felsenstein, J. (1981) 'Evolutionary trees from DNA sequences: A maximum likelihood approach', *Journal of Molecular Evolution*, *17*(6), pp. 368–376.

Felsenstein, J. (1985) 'Confidence Limits on Phylogenies: An Approach Using the Bootstrap', *Evolution*, *39*(4), p. 783.

Fitch, W. M. (1971) 'Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology', *Systematic Zoology*, *20*(4), p. 406.

Friedman, C.S., Beaman, B.L., Chun, J., Goodfellow, M., Gee, A. and Hedrick, R.P. (1998) *Nocardia crassostreae* sp. nov., the causal agent of nocardiosis in Pacific oysters, *International Journal of Systematic Bacteriology*, *48*(1), pp.237-246.

Goodfellow, M. (2014) 'Selective Isolation of Actinobacteria', *Manual of Industrial Microbiology and Biotechnology*, pp. 13–27.

Goodfellow, M. and Maldonado, L. A. (2015) '*Nocardia*', *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–74.

Hoza, A.S., Mfinanga, S.G., Moser, I. and König, B. (2017) 'Isolation, biochemical and molecular identification of *Nocardia* species among TB suspects in northeastern, Tanzania; A forgotten or neglected threat?', *BMC Infectious Diseases*, *17*(1), pp. 1–9.

Jensen, P. R., Dwight, R. and Fenical, W. (1991) 'Distribution of actinomycetes in near-shore tropical marine sediments', *Applied and Environmental Microbiology*, *57*(4), pp. 1102–1108.

Jensen, P.R., Gontang, E., Mafnas, C., Mincer, T.J. and Fenical, W. (2005) 'Culturable marine actinomycete diversity from tropical Pacific Ocean sediments', *Environmental Microbiology*, *7*(7), pp. 1039–1048.

Jensen, P. R., Maldonado, L. A. and Goodfellow, M. (2015) '*Salinispora*', *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–10.

Kageyama, A., Yazawa, K., Taniguchi, H., Chibana, H., Nishimura, K., Kroppenstedt, R.M. and Mikami, Y . (2005) '*Nocardia concava* sp. nov., isolated from Japanese patients', *International Journal of Systematic and Evolutionary Microbiology*, *55*(5), pp. 2081–2083.

Kämpfer, P. (2015) '*Streptomyces*', *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–414.

Kim, K.O., Shin, K.S., Kim, M.N., Shin, K.S., Labeda, D.P., Han, J.H. and Kim, S.B. (2012a) 'Reassessment of the status of *Streptomyces setonii* and reclassification of *Streptomyces fimicarius* as a later synonym of *Streptomyces setonii* and *Streptomyces albovinaceus* as a later synonym of *Streptomyces globisporus* based on combined 16S rRNA/*gyrB* gene sequence analysis', *International Journal of Systematic and Evolutionary Microbiology*, *62*(12), pp. 2978–2985.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H. and Won, S. (2012b) 'Introducing EzTaxon-e: A prokaryotic 16s rRNA gene sequence database with phylotypes that represent uncultured species', *International Journal of Systematic and Evolutionary Microbiology*, *62*, pp. 716–721.

Kimura, M. (1980) 'A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences.', *J. Mol. Evol.*, *16*, pp. 111–120.

Kirby, B. M., Everest, G. J. and Meyers, P. R. (2010) 'Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: Proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *97*(2), pp. 131–142.

Kreader, C. A. (1996) 'Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein', *Applied and Environmental Microbiology*, *62*(3), pp. 1102–1106.

Kudo, T., Hatai, K. and Seino, A. (1988) '*Nocardia seriolae* sp. nov. causing nocardiosis of cultured fish', *International Journal of Systematic Bacteriology*, *38*(2), pp. 173–178.

Kumar, S., Stecher, G. & Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets Downloaded from', *Mol. Biol. Evol*, *33*(7), pp. 1870–1874.

Law, J.W.F., Tan, K.X., Wong, S.H., Ab Mutalib, N.S. and Lee, L.H. (2018) 'Taxonomic and Characterization Methods of *Streptomyces*: A Review', *Progress In Microbes & Molecular Biology*, *1*(1), pp. 1–13.

Leger, J.S., Begeman, L., Fleetwood, M., Frasca Jr, S., Garner, M.M., Lair, S., Trembley, S., Linn, M.J. and Terio, K.A. (2009) 'Comparative pathology of Nocardiosis in marine mammals', *Veterinary Pathology*, *46*(2), pp.299-308.

Le Roes, M., Goodwin, C. M. and Meyers, P. R. (2008) '*Gordonia lacunae* sp. nov., isolated from an estuary', *Systematic and Applied Microbiology*, *31*, pp. 17–23.

Maldonado, L.A., Fenical, W., Jensen, P.R., Kauffman, C.A., Mincer, T.J., Ward, A.C., Bull, A.T. and Goodfellow, M. (2005) '*Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family *Micromonosporaceae*', *International Journal of Systematic and Evolutionary Microbiology*, *55*(5), pp. 1759–1766.

Meier-Kolthoff, J.P., Göker, M., Spröer, C. and Klenk, H.P. (2013) 'When should a DDH experiment be mandatory in microbial taxonomy?', *Archives of Microbiology*, *195*(6), pp. 413-418.

Mincer, T. J., Fenical, W. and Jensen, P. R. (2005) 'Culture-Dependent and Culture-Independent Diversity within the Obligate Marine Actinomycete Genus *Salinispora*', *Applied and Environmental Microbiology*, *71*(11), p. 7019.

Ogier, J.C., Pages, S., Galan, M., Barret, M. and Gaudriault, S. (2019) 'RpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing', *BMC Microbiology*, *19*(1), pp. 1–16.

Parte, A. C. (2018) 'LPSN - List of prokaryotic names with standing in nomenclature (Bacterio.net), 20 years on', *International Journal of Systematic and Evolutionary Microbiology*, *68*(6), pp. 1825–1829.

Petti, C.A., Bosshard, P.P., Brandt, M.E., Clarridge, J.E., Feldblyum, T.V., Foxall, P., Furtado, M.R., Pace, N. and Procop, G. (2008) 'Interpretive criteria for identification of bacteria and fungi by DNA target sequencing; approved guideline', *Clinical and Laboratory Standards Institute (CLSI) Documents*, *28*, pp.19087-1898

Roth, A., Andrees, S., Kroppenstedt, R.M., Harmsen, D. and Mauch, H. (2003) 'Phylogeny of the genus *Nocardia* based on reassessed 16S rRNA gene sequences reveals underspeciation and division of strains classified as *Nocardia asteroides* into three established species and two unnamed taxons', *Journal of Clinical Microbiology*, 41(2), pp. 851–856.

Saitou, N. and Nei, M. (1987) 'The neighbor-joining method: a new method for reconstructing phylogenetic trees.', *Molecular Biology and Evolution*, *4*(4), pp. 406–425.

Santos, S. R. and Ochman, H. (2004) 'Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins', *Environmental Microbiology*, *6*(7), pp. 754–759.

Sarmiento-Vizcaino, A., Espadas, J., Martín, J., Brana, A.F., Reyes, F., Garcia, L.A. and Blanco, G. (2018) 'Atmospheric precipitations, hailstone and rainwater, as a novel source of *Streptomyces* producing bioactive natural products', *Frontiers in Microbiology*, *9*, p. 773.

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D. and Karsch-Mizrachi, I. (2019) 'GenBank', *Nucleic acids research*, 47, pp.D94-D99

Shirling, E. B. and Gottlieb, D. (1966) 'Methods for characterization of *Streptomyces* species', *International journal of systematic bacteriology*, *16*(3), pp.313-340.

Sivalingam, P., Hong, K., Pote, J. and Prabakar, K (2019) 'Extreme Environment *Streptomyces*: Potential Sources for New Antibacterial and Anticancer Drug Leads?' *International Journal of Microbiology, 2019.*

Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R. and Liesack, W. (1991) 'Designation of streptomycete 16S and 23S rRNA-based target regions for oligonucleotide probes', *Applied and Environmental Microbiology*, *57*(5), pp. 1468–1477.

Takeda, K., Kang, Y., Yazawa, K., Gonoi, T. and Mikami, Y. (2010) 'Phylogenetic studies of *Nocardia* species based on *gyrB* gene analyses', *Journal of Medical Microbiology*, *59*(2), pp. 165–171.

Trevisan V. (1889) 'Genera and species of the batteries [in Italian]', *Milan: Zanaboni and Gabuzzi*

Vyas, T. K. and Dave, B. P. (2011) 'Production of biosurfactant by *Nocardia otitidiscaviarum* and its role in biodegradation of crude oil', *Int. J. Environ. Sci. Tech*, *8*(2), pp. 425–432.

Waksman, Selman A., Henrici, A. T. (1944) 'The nomenclature and classification of map projections', *Empire Survey Review*, *7*(51), pp. 190–200.

Wang, Y., Zhang, Z. & Ruan, J. (1996). 'A Proposal to Transfer *Microbispora bispora* (Lechevalier 1965) to a New Genus, *Thermobispora* gen. nov., as *Thermobispora bispora* comb. nov.' *International Journal of Systematic Bacteriology*. *46*(4), pp. 933-938.

Wellington, E.M.H., Stackebrandt, E., Sanders, D., Wolstrup, J. and Jorgensen, N.O.G. (1992) 'Taxonomic status of *Kitasatosporia*, and proposed unification with *Streptomyces* on the basis of phenotypic and 16S rRNA analysis and emendation of *Streptomyces* Waksman and Henrici 1943, 339(AL)', *International Journal of Systematic Bacteriology*, *42*(1), pp. 156–160.

Wright, L., Katouli, M. and Kurtböke, D. İ. (2021) 'Isolation and characterization of nocardiae associated with foaming coastal marine waters', *Pathogens*, *10*(5), p.579

Zhang, Z., Wang, Y. and Ruan, J. (1997) 'A proposal to revive the genus *Kitasatospora* (Omura, Takahashi, Iwai, and Tanaka 1982)', *International journal of systematic bacteriology*, *47*(4), pp. 1048–1054.

CHAPTER 3:

METAGENOMIC INVESTIGATION OF THE ACTINOBACTERIAL COMPOSITION OF MARINE SEDIMENT SAMPLES FROM GERICKE'S POINT

SUMMARY

Metagenomic data were used to investigate the actinobacterial composition of sediment samples obtained from a rock pool (rocky site), a beach (dry site) and a subtidal zone (ocean site) at Gericke's Point in Sedgefield, South Africa. The 16S-rRNA gene was amplified using the actinobacterium-specific 16S-rRNA gene primer pair, Com2xf and Ac1186r. A total of 1 748 610 partial 16S-rRNA gene sequences were classified using the SILVA 16S-rRNA gene database of which 1 541 544 (~88%) were classified as belonging to actinobacterial strains (the remaining 12% was identified as non-actinobacterial or unclassified strains). Analyses of the actinobacterial diversity revealed that strains belonging to the orders *Candidatus* Microtrichales, *Candidatus* Actinomarinales and *Propionibacteriales* represented the first, second and third highest proportion, respectively, of identified actinobacteria in all three sites. The actinobacterial composition for each site differed with the rocky site displaying the most diversity and the ocean site displaying the least. The percentage compositions of the actinobacterial diversity for each order present in the ocean site sediment samples were as follows: *Ca* Microtrichales, 80%; *Ca* Actinomarinales, 18%; *Propionibacteriales*, 1%; other actinobacterial orders including *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporale*, *Mycobacteriales* and *Pseudonocardiales*, less than 1%. The percentage compositions of the actinobacterial diversity for each order present in the dry site sediment samples were as follows: *Ca* Microtrichales, 78%; *Ca* Actinomarinales, 15%; *Propionibacteriales* , 7%; other actinobacterial orders including *Euzebyales*, *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales*, *Pseudonocardiales*, *Streptomycetales* and *Streptosporangiales*, less than 1%. The percentage compositions of the actinobacterial diversity for each order present in the rocky site sediment samples were as follows: *Ca* Microtrichales, 77%; *Ca* Actinomarinales, 14%; *Propionibacteriales* , 8%; other actinobacterial orders including *Bifidobacteriales*, *Euzebyales*, *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales*, *Pseudonocardiales*, *Streptomycetales* and *Streptosporangiales*, 1%. The majority (99%) of the total actinobacterial strains found at Gericke's Point could not be assigned to a known genus. This represents an abundance of novel actinobacterial diversity that has yet to be fully uncovered.

Table of Contents

3.1 INTRODUCTION

Culture-dependent methods have been used for more than 100 years to study the diversity of bacterial communities. However, the majority of bacteria have yet to be cultured due to limitations such as lack of selectivity of culture media and difficulty dislodging bacteria or spores from environmental samples (Salmonová and Bunešová, 2017). Furthermore, standard culturing conditions favour the growth of only a fraction of the bacterial community (Al-Awadhi *et al.*, 2013). It is particularly hard to culture indigenous marine bacteria (Li and Qin, 2005) and this unculturable diversity represents a very promising source of novel bioactive compounds.

Metagenomics is the analysis of DNA from microbial communities where the DNA is extracted directly from environmental samples. This allows one to study the full range of microbial diversity from environmental samples without prior need for culturing. Advances in next-generation sequencing (NGS) technologies have allowed significant breakthroughs in these types of studies (Oulas *et al.*, 2015). A subset of NGS-based metagenomics involves using amplicons from a taxonomic marker gene to profile microbial communities. Despite its limitations, specifically with regards to species identification, the most common taxonomic marker used in metagenomic studies is the 16S-rRNA gene (Aguiar-Pulido *et al.*, 2016). The other subset of NGS-based metagenomics involves whole-metagenome sequencing which allows, in addition to taxonomic profiling, some insight into the functional profile of a microbial community. Metagenomic analyses have been used to identify novel taxa (Ghai *et al.*, 2013). However, the eventual culturing of strains belonging to these taxa will still be necessary for full classification. Selective isolation techniques informed by insight gained from metagenomic analyses could be used to help achieve this. When used in conjunction with metatranscriptomics and metabolomics, metagenomics could also aid the discovery of novel secondary metabolites.

Metagenomics is a valuable tool in microbial ecology, however there are a few limitations associated with current methodologies available for metagenomic studies. The largest limitation of metagenomic analyses is the reliance of their robustness on the continuous improvement of existing databases and analysis software (Zhang *et al.*, 2021). Other limitations include difficulties regarding amplification bias towards certain microorganisms (particularly in amplicon-based analyses) (Al-Awadhi *et al.*, 2013; Hiraoka *et al.*, 2016). This bias varies depending on the PCR primers used. Additionally, most commonly used

sequencing instruments come with their own set of limitations regarding accuracy, length of reads and, in the case of shotgun sequencing, high cost (Zhang *et al.*, 2021).

The aim of this part of the study was to investigate the actinobacterial composition of sediment samples taken from marine habitats at Gericke's Point, Sedgefield, South Africa. This was done using data obtained from NGS of 16S-rRNA amplicons. The analysis detailed in this chapter served to complement the culture-dependent investigation included in this study (detailed in Chapter 2).

3.2 METHODS AND MATERIALS

3.2.1 16S-rRNA metabarcoding

The protocols described in sections 3.2.1.1 and 3.2.1.2 were performed by members of the BTB research group.

3.2.1.1 Amplicon preparation and sequencing

Metagenomic DNA was extracted from the same sediment samples used in the culture-dependent analysis (Chapter 2) using the DNeasy PowerSoil DNA Isolation Kit (Qiagen) according to the manufacturer's protocol with the following amendment: 0.5g sediment was used as the starting material instead of 0.25g. PCR amplification was performed using the actinobacterium-specific 16S-rRNA gene primer pair Com2xf (5'-AAACTCAAAGGAATTGACGG-3') and Ac1186r (5'-CTTCCTCCGAGTTGACCC-3') designed to target and amplify a ~270bp fragment of the actinobacterial 16S-rRNA gene (Schäfer *et al.*, 2010). PCR reactions (25 µl) contained 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µM of each primer, 10 ng/µl metagenomic DNA and 1x KAPA Taq Readymix. The amplification programme was started by an initial denaturation at 95°C for 3 min, followed by 25 cycles of denaturation at 94 ^{o}C for 30s, annealing for 30s with a temperature gradient (51.6°C – 60.2°C) and extension for 30s at 72 ^{o}C. This was followed by a final extension step of 72°C for 15 minutes. Genomic DNA from *Streptomyces pharetrae* CZA14^{T} was used as a positive control. Amplicons were analysed by electrophoresis on a 1% (w/v) agarose gel containing 10 µg/ml ethidium bromide and visualised under ultra violet (UV) light. Amplicons were purified using the MSB PCRapace Purification Kit (Invitek Molecular) and submitted for next-generation sequencing on the IonTorrent S5 Platform at the Central Analytical Facilities (Stellenbosch University, Stellenbosch, South Africa).

3.2.1.2 Metabarcoding analyses

Metabarcoding analyses was performed remotely using mothur v1.44.0 (Schloss *et al.*, 2009) via the Centre for High Performance Computing (CHPC) platform located at the Council for Scientific and Industrial Research, Rosebank Campus (Cape Town, South Africa). Commands were submitted and queued in batches via *ssh* to the CHPC PBSPro scheduler. All commands are self-contained in the mothur installation and did not require additional

components to be downloaded. The raw reads were quality-filtered (minimum Phred score: 20), ambiguous bases were removed, and the reads were merged into a single file. Chimeras were removed using the VSEARCH algorithm (Rognes *et al*., 2016). The reads were classified using a Bayesian classifier and the SILVA 16S-rRNA gene database (v138, Quast *et al.*, 2013) was used as the reference database. Non-prokaryotic sequences were removed. The sequences were aligned with the reference database and curated to ensure the region of interest overlapped for all samples. Singletons were removed and subsampling was done to normalize the data. The sequences were clustered into Operational Taxonomic Units (OTUs) with a default cutoff of 97% 16S-rRNA gene sequence similarity and an OTU table was generated for downstream analyses.

3.2.2 Analysis of biodiversity

Pie charts were generated (by this author) using Microsoft Excel based on the actinobacterial orders identified in each sample and their relative OTU abundance. For each site, an abundance value for each order was calculated using the sum of the counts present in all six sediment samples and percentage composition was determined relative to the total count of actinobacteria identified to the order level.

3.3 RESULTS AND DISCUSSION

3.3.1 Actinobacterial composition of marine sediment from Gericke's Point

A total of 1 748 610 16S-rRNA sequence reads were obtained from 18 sediment samples (six sub-samples for each of three sampling sites) and classified using the SILVA 16S-rRNA gene database. Of the aforementioned reads, 1 541 544 (~88%) were classified as belonging to the phylum *Actinobacteria*, 185 306 (~11%) were classified as non-actinobacteria and 21 760 (~1%) were classified as bacteria that could not be assigned to a known phylum. It is not surprising that the actinobacterial-specific 16S-rRNA gene primers amplified a small portion of non-actinobacterial 16S-rRNA gene sequences as a similar result was reported by Schäfer *et al.* (2010). Interestingly, although most of the contaminants detailed in Chapter 2, belonged to the classes *Betaproteobacteria* and *Gammaproteobacteria*, only 16S-rRNA gene sequences belonging to bacteria in the class *Alphaproteobacteria* were amplified from the Gericke's Point sediment samples.

Pie charts were generated to represent the actinobacterial diversity found in the ocean (Figure 3.1), dry (Figure 3.2) and rocky (Figure 3.3) site sediment samples. The aforementioned pie charts were based on classifications at the order level as there were too many OTUs to include at the lower taxonomic ranks. Actinobacteria that could not be assigned to a known class or order made up ~41% (644 712 of 1 541 544) of the total actinobacterial strains found in the Gericke's Point sediment samples and were not included in Figures 3.1 – 3.3. The majority (99%; 1 531 320 of 1 541 544) of actinobacterial strains found in the Gericke's Point could not be assigned to a known genus. This represents an abundance of novel actinobacterial diversity that has not yet been cultured or explored. Selective isolation techniques could be used to increase the chances of isolating these strains.

Strains belonging to the orders *Candidatus* Microtrichales, *Candidatus* Actinomarinales and *Propionibacteriales* represented the first, second and third highest proportion respectively of actinobacteria found in the ocean, dry and rocky site sediment samples from Gericke's Point (Figures 3.1, 3.2 and 3.3). Figure 3.1 shows that strains belonging to the orders *Frankiales, Geodermatophilales, Micrococcales, Micromonosporales, Mycobacteriales* and *Pseudonocardiales,* though representing a very small proportion (less than one percent) of the identified actinobacterial diversity, were also found in the sediment samples obtained from the ocean site. Figure 3.2 shows that, in addition to the aforementioned orders present in the

ocean site sediment samples, strains belonging to the orders *Euzebyales*, *Streptomycetales* and *Streptosporangiales* also represented a very small proportion of the actinobacterial diversity found in the sediment samples obtained from the dry site. The rocky site sediment samples contained the most actinobacterial diversity, as they contained strains belonging to all the orders mentioned above plus strains belonging to the order *Bifidobacteriales* (Figure 3.3).

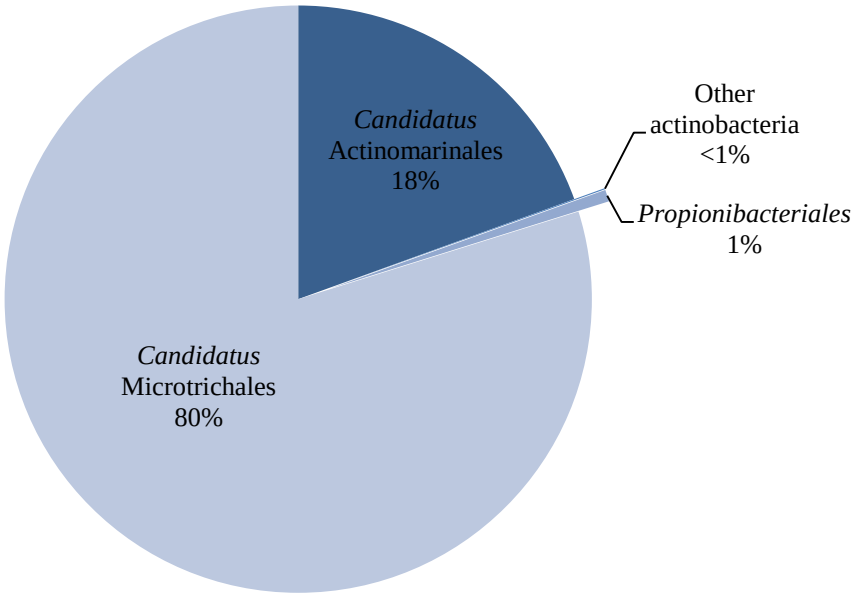

Figure 3.1. Percentage composition of the actinobacterial diversity present in the ocean site sediment, based on identifications to the order level. The results are based on the sum of the counts per order relative to the total count of actinobacteria identified to the order level (284 913) over six replicate sediment samples. The actinobacterial orders that make up 'Other actinobacteria' include *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales* and *Pseudonocardiales*.

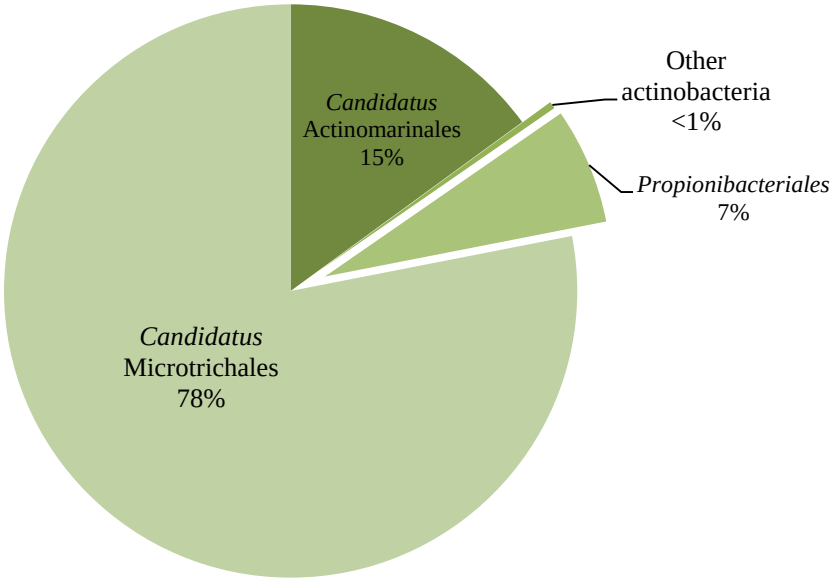

Figure 3.2. Percentage composition of the actinobacterial diversity present in the dry site sediment, based on identifications to the order level. The results are based on the sum of the counts per order relative to the total count of actinobacteria identified to the order level (312 325) over six replicate sediment samples. The actinobacterial orders that make up 'Other actinobacteria' include *Euzebyales*, *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales*, *Pseudonocardiales*, *Streptomycetales* and *Streptosporangiales.*

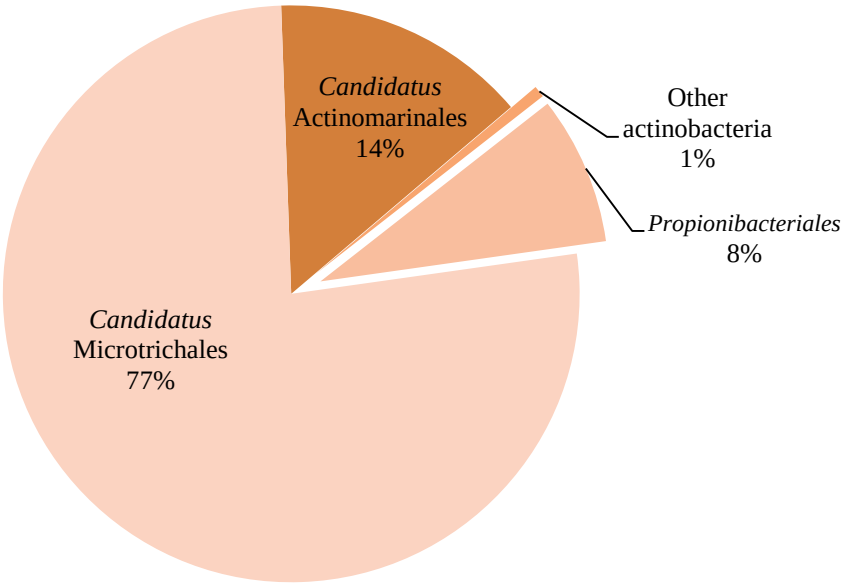

Figure 3.3. Percentage composition of the bacterial diversity present in the rocky site sediment, based on identifications to the order level. The results are based on the sum of the counts per order relative to the total count of actinobacteria identified to the order level (299 594) over six replicate sediment samples. The actinobacterial orders that make up 'Other actinobacteria' include *Bifidobacteriales*, *Mycobacteriales*, *Euzebyales*, *Frankiales*, *Geodermatophilales*, *Micrococcales*, *Micromonosporales*, *Mycobacteriales*, *Pseudonocardiales*, *Streptomycetales* and *Streptosporangiales*

3.3.1.1 Class *Acidimicrobiia*

3.3.1.1.1 Candidatus *Microtrichales*

Figures 3.1, 3.2 and 3.3 show that members of *Ca* Microtrichales represent the highest proportion of actinobacteria found in the sediment samples from Gericke's Point, making up 80%, 78% and 77% of the total actinobacterial diversity at the order level in the ocean, dry and rocky sites, respectively. The SILVA 16S-rRNA gene database classifies members of the families *Iamiaceae, Ilumatobacteraceae* and *Candidatus* Microtrichaceae under *Ca* Microtrichales, though only members of the latter two families were found in the Gericke's Point sediment samples. Since the order *Ca* Microtrichales is not a valid-published name, members of the families *Iamiaceae* and *Ilumatobacteraceae* are currently classified as belonging to the order *Acidimicrobiales.* Furthermore, *Ca* Microtrichaceae contains members of the Sva0996 marine group, which are also currently classified as belonging to the order *Acidimicrobiales* (https://www.uniprot.org/taxonomy/1916143). Strains belonging to the family *Ilumatobacteraceae* made up ~13% (92 861 of 700 992), while strains belonging to *Ca* Microtrichaceae made up ~64% (452 884 of 700 992) of the strains identified as members of *Ca* Microtrichales. The rest of the identified strains could not be assigned to a known family. To date, none of the members of *Ca* Microtrichales have ever been cultured (Huang *et al.*, 2021). Consequently, all of the identified actinobacteria belonging to *Ca* Microtrichales in this study have not been assigned to a known genus. Unlike the type genus of the order *Acidimicrobiales*, *Acidimicrobium*, members of the family *Ilumatobacteraceae* are mesophilic and neutrophilic (Asem *et al.*, 2018). Furthermore, no known acidophilic members of the class *Acidimicrobiia* were found in the Gericke's Point sediment samples. This is not surprising as the ocean, dry and rocky site sediment samples were found to have a pH of 8, 8.8 and 8.7, respectively. All classified members of the type genus of the family *Ilumatobacteraceae, Ilumatobacter*, have been isolated from aquatic habitats (https://lpsn.dsmz.de/genus/ilumatobacter) and the members of the Sva0996 marine group (*Ca* Microtrichaceae) were discovered in the microbiome of a marine sponge (Slaby *et al.*, 2017).

3.3.1.1.2 *Candidatus* Actinomarinales

The order found to represent the second highest proportion of actinobacteria in the Gericke's Point samples was *Ca* Actinomarinales making up 18%, 15% and 14% of the total actinobacterial diversity at the order level in the ocean, dry and rocky sites, respectively

(Figure 3.1; Figure 3.2; Figure 3.3). Similar to *Ca* Microtrichales, *Ca* Actinomarinales is an order, not yet validly published, proposed to belong to the class *Acidimicrobiia*. *Ca* Actinomarinales was proposed by Ghai *et al.* (2013) to designate a deep-branching lineage of obligately marine actinobacteria with extremely small free-living cells and very low GC content (33%). These actinobacteria were discovered using metagenome-assembled genomes and have yet to be cultured. In this study, none of the strains identified as members of *Ca* Actinomarinales could be assigned to a known family or genus. Previously, members of *Ca* Actinomarinales were thought to be found only in marine epipelagic waters, which lead researchers to believe they were photoheterotrophs (Ghai *et al.*, 2013; López-Pérez *et al.*, 2020). However, members of *Ca* Actinomarinales have since been found in coastal sediment, implying that not all members of this order are phototrophic (Huang *et al.*, 2021).

3.3.1.2 Class *Actinobacteria*

3.3.1.2.1 Order *Bifidobacteriales*

Two strains belonging to the order *Bifidobacteriales* were found in one sediment sample from the rocky site. These strains belonged to the genus *Bifidobacterium*. Members of the order *Bifidobacteriales* are anaerobic and most commonly found in the gastrointestinal tract of humans and other animals (Zhang *et al.*, 2016). *Bifidobacterium* strains have also been assessed as indicators of human faecal pollution in freshwater and marine environments (Resnick and Levin, 1981; Morrison *et al.*, 2008)

3.3.1.2.2 Order *Frankiales*

Eleven (11), 36 and 652 strains belonging to the order *Frankiales* were found in the ocean, dry and rocky site sediment samples, respectively. The majority (~98%; 689 of 699) of these strains, including all *Frankiales* strains found in the dry site sediment samples, could not be assigned to a known family or genus. The remaining ten *Frankiales* strains belonged to the genus *Frankia* of the family *Frankiaceae*. Strains belonging to this genus are nitrogen-fixing actinobacteria that are found as free-living microbes in soil and in symbiotic associations with actinorhizal plants (Rehan *et al.*, 2016). *Frankia* strains have also been shown to tolerate a narrow range of temperatures, pH levels and salinity (Wijnholds and Young, 2000; Srivastava *et al.*, 2017). To date, *Frankiales* strains have only been found in terrestrial and freshwater habitats (Normand and Fernandez, 2020; Betiku *et al.*, 2021). Therefore, it is

surprising and noteworthy that members of this order were found in the marine habitats at Gericke's Point.

3.3.1.2.3 Order *Geodermatophilales*

Forty-three (43), 78 and 824 strains belonging to the order *Geodermatophilales* were found in the ocean, dry and rocky site sediment samples, respectively. All of these strains belonged to the family *Geodermatophilaceae*. Almost half (426 of 945) of the strains belonging to *Geodermatophilaceae* could not be assigned to a known genus. The rest of the strains belonged to the genera *Blastococcus* (38 of 945) and *Geodermatophilus* (481 of 945). Members of the family *Geodermatophilaceae* have been isolated mainly from soils, seawater, and stone surfaces (Normand *et al.*, 2014).

3.3.1.2.4 Order *Micrococcales*

Eighty-five (85), 128 and 180 strains belonging to the order *Micrococcales* were found in the ocean, dry and rocky site sediment samples, respectively. Members of the order *Micrococcales* are non-spore forming bacteria that have been found to inhabit mammalian skin, various soil samples and marine environments (Dastager *et al.*, 2014; Salam *et al.*, 2020). More than half (257 of 393) of the identified *Micrococcales* strains could not be assigned to a known genus. The majority of these strains (209 of 257) could not be assigned to a known family either. The rest of the strains identified as *Micrococcales* were evenly distributed across seven of the 17 known families currently belonging to the order, namely *Cellulomonadaceae*, *Dermabacteraceae*, *Dermacoccaceae*, *Intrasporangiaceae*, *Microbacteriaceae*, *Micrococcaceae* and *Promicromonosporaceae* (https://lpsn.dsmz.de/order/micrococcales). All strains belonging to *Intrasporangiaceae* could not be assigned to a known genus in this family. The following genera were found in all Gericke's Point sediment samples: *Microbacterium* (family *Microbacteriaceae*), *Isoptericola* (family *Promicromonosporaceae*) and *Arthrobacter* and *Micrococcus* from the family *Micrococcaceae*. Members of the genus *Kocuria* (family *Micrococcaceae*) were found only in the ocean and dry site sediment samples. The genera *Glutamicibacter* and *Nesterenkonia* (family *Micrococcaceae*) and *Barrientosiimonas* (family *Dermacoccaceae*) were exclusively found in the ocean site sediment and the genera *Cellulomonas* (family *Cellulomonadaceae*), *Brachybacterium* (family *Dermabacteraceae*) and *Agromyces* (family *Microbacteriaceae)* were exclusively found in the rocky site sediment. Two strains of the genus *Candidatus* Planktoluna (family *Microbacteriaceae*) were found in sediment samples from the dry site.

3.3.1.2.5 Order *Micromonosporales*

Fourteen (14), 141 and 69 strains belonging to the order *Micromonosporales* were found in the ocean, dry and rocky site sediment samples, respectively. This order currently contains only one family, *Micromonosporaceae* (https://lpsn.dsmz.de/order/micromonosporales). *Micromonosporaceae* strains have been isolated from diverse habitats including soil, fresh and marine water, rhizosphere soil, and plant tissues (Trujillo *et al.* , 2014). The family *Micromonosporaceae* is comprised of 30 genera, including the obligate marine genus *Salinispora* (https://lpsn.dsmz.de/family/micromonosporaceae). Seventy five percent (75%; 168 of 224) of the *Micromonosporales* strains identified in the Gericke's Point sediment samples could not be assigned to a known genus. Thirty-nine (39) strains belonging to the genus *Micromonospora* was found in sediment samples from all three sites. One and two strains belonging to the genus *Actinoplanes*, and two and 12 strains belonging to the genus *Salinispora*, were found in the ocean and dry site sediment samples, respectively. It is interesting that more strains of the obligately marine genus *Salinispora* were found in the dry sediment as opposed to sediment from the subtidal zone.

3.3.1.2.6 Order *Mycobacteriales*

Fourteen (14), 71 and 104 strains belonging to the order *Mycobacteriales* were found in the ocean, dry and rocky site sediment samples, respectively. Members of this order are found mainly in soil ecosystems, but their presence has also been reported in freshwater and marine environments (Goodfellow and Jones, 2015; Kurilkina *et al.*, 2016; Wright *et al.*, 2021). Of the eight known families currently belonging to the order *Mycobacteriales* (https://lpsn.dsmz.de/order/mycobacteriales), only members of the families *Corynebacteriaceae* (~11%; 22 of 189*)*, *Mycobacteriaceae* (~60%; 115 of 189*)*, *Gordoniaceae* (~1%; 3 of 189) and *Nocardiaceae* (~25%; 48 of 189*)* were found. One identified *Mycobacteriales* strain could not be assigned to a known family. Genera found include *Corynebacterium* (family *Corynebacteriaceae*), *Mycobacterium* (family *Mycobacteriaceae*), *Gordonia* (family *Gordoniaceae*) and *Nocardia* (23 of 48) and *Rhodococcus* (10 of 48) belonging to the family *Nocardiaceae.* Fifteen (15) strains belonging to the family *Nocardiaceae* could not be assigned to a known genus. A marine *Nocardia* strain was isolated in this study (Chapter 2).

3.3.1.2.7 Order *Propionibacteriales*

The order *Propionibacteriales* represented 1%, 7% and 8% of the total actinobacterial diversity (identified to the order level) found in the ocean, dry and rocky site sediment samples, respectively (Figures 3.1, 3.2 and 3.3). This represents the third highest proportion of actinobacteria found in Gericke's Point. Members of the order *Propionibacteriales* are less frequently isolated in culture-dependent studies compared to other genera, such as *Streptomyces*, but have been reported to occur in various environments such as soil, marine invertebrates, plants and human samples (Stackebrandt, 2014; Tóth and Borsodi, 2014; Shi *et al.*, 2022). The order *Propionibacteriales* currently contains the families *Actinopolymorphaceae*, *Kribbellaceae*, *Nocardioidaceae* and *Propionibacteriaceae* (https://lpsn.dsmz.de/order/propionibacteriales), though only members of *Nocardioidaceae* and *Propionibacteriaceae* were found in the Gericke's Point sediment samples. The family *Nocardioidaceae* currently contains four genera (https://lpsn.dsmz.de/family/nocardioidaceae), but only members of the genera *Marmoricola* and *Nocardioides* were found in the Gericke's Point sediment samples. The family *Propionibacteriaceae* currently contains 24 genera (https://lpsn.dsmz.de/family/propionibacteriaceae), but only members of the genera *Cutibacterium*, *Microlunatus* and *Arachnia* were found in the Gericke's Point sediment samples. *Nocardioidaceae* strains made up ~70% (33 223 of 47291) of the strains identified as members of the order *Propionibacteriales.* Strains belonging to the family *Propionibacteriaceae* represented a very small portion (~0.08%; 38 of 47 291) of the strains identified as members of the order *Propionibacteriales.* The remaining strains identified as *Propionibacteriales* could not be assigned to a known family. Furthermore, ~18% (6018 of 33 223) of the strains identified as members of the family *Nocardioidaceae* and ~5% (2 of 38) of the strains identified as members of the family *Propionibacteriaceae* could not be assigned to a known genus.

3.3.1.2.8 Order *Pseudonocardiales*

The order *Pseudonocardiales* is comprised of a single family, namely *Pseudonocardiaceae* (https://lpsn.dsmz.de/order/pseudonocardiales). Members of the family *Pseudonocardiaceae* exhibit a wide range of different physiologies and are known antibiotic producers (Platas *et al.*, 1998). *Pseudonocardiaceae* strains were found in all Gericke's Point samples, but the majority (~93%, 784 of 837) of these strains were found in the dry site sediment samples.

Sixteen (16) and 36 *Pseudonocardiaceae* strains were found in the ocean and rocky site samples, respectively. The majority (~96%; 808 of 837) of strains identified as *Pseudonocardiaceae* could not be assigned to a known genus. Two of the 39 genera currently belonging to the family *Pseudonocardiaceae* (https://lpsn.dsmz.de/family/pseudonocardiaceae), *Amycolatopsis* (6 of 837) and *Pseudonocardia (*23 of 837) were found in the Gericke's Point samples with members of *Amycolatopsis* found exclusively in the dry site sediment.

3.3.1.2.9 Order *Streptomycetales*

One and ten strains belonging to the order *Streptomycetales* were found in the dry and rocky site sediment samples, respectively. All of these strains belonged to the genus *Streptomyces*, one of six genera in the sole *Streptomycetales* family, *Streptomycetaceae* (https://lpsn.dsmz.de/family/streptomycetaceae). The other genera in the family are *Allostreptomyces*, *Embleya*, *Kitasatospora*, *Streptacidiphilus* and *Yinghuangia*. The majority of strains cultured in this study were identified as *Streptomyces*. It is interesting that no strains belonging to the genus *Streptomyces* were found in the metagenomic analysis of the ocean site sediment samples as both strains isolated from this sample site (GGUI#1 and GGUI10#6; detailed in Chapter 2) were identified as *Streptomyces*.

3.3.1.2.10 Order *Streptosporangiales*

The order *Streptosporangiales* currently contains the families *Nocardiopsaceae*, *Streptosporangiaceae*, *Thermomonosporaceae* and *Treboniaceae* (https://lpsn.dsmz.de/order/streptosporangiales). Four and 31 strains belonging to the genus *Nocardiopsis* (family *Nocardiopsaceae*) were found in the dry and rocky site sediment samples, respectively. It has been shown that species of the family *Nocardiopsaceae* can be isolated from the marine environment with the help of selective isolation techniques (Ng and Tan, 2018). Nine strains belonging to the genus *Actinomadura* (family *Thermomonosporaceae*) were found exclusively in the rocky site samples. An additional nine strains belonging to the family *Thermomonosporaceae* were found in the dry and rocky site sediment, but these strains could not be assigned to a known genus.

3.3.1.3 Class *Nitriliruptoria*

3.3.1.3.1 Order *Euzebyales*

One and two strains belonging to the order *Euzebyales* were found in a sediment sample from the dry site and a sediment sample from the rocky site, respectively. These strains belonged to the genus *Euzebya*. All described members of this genus have been isolated from the marine environment (Kurahashi *et al.*, 2010; Yin *et al.*, 2018; Jian *et al.*, 2021).

3.4. REFERENCES

Aguiar-Pulido, V., Huang, W., Suarez-Ulloa, V., Cickovski, T., Mathee, K. and Narasimhan, G. (2016) 'Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis', *Evolutionary Bioinformatics*, *12*, pp. 5–16.

Al-Awadhi, H., Dashti, N., Khanafer, M., Al-Mailem, D., Ali, N. and Radwan, S. (2013) 'Bias problems in culture-independent analysis of environmental bacterial communities: A representative study on hydrocarbonoclastic bacteria', *SpringerPlus*, *2*(1), pp. 1-11.

Asem, M.D., Shi, L., Jiao, J.Y., Wang, D., Han, M.X., Dong, L., Liu, F., Salam, N. and Li, W.J. (2018) '*Desertimonas flava* gen . nov ., sp . nov . isolated from a desert soil , and proposal of *Ilumatobacteraceae* fam . nov .' *International Journal of Systematic and Evolutionary Microbiology*, *68*(11), pp. 3593–3599.

Betiku, O.C., Sarjeant, K.C., Ngatia, L.W., Aghimien, M.O., Odewumi, C.O. and Latinwo, L.M. (2021) 'Evaluation of microbial diversity of three recreational water bodies using 16S rRNA metagenomic approach', *Science of the Total Environment*, *771*, p. 144773.

Dastager, S.G., Krishnamurthi, S., Rameshkumar, N. and Dharne, M. (2014) 'The family *Micrococcaceae*', *The Prokaryotes: Actinobacteria*, pp. 455–498.

Ghai, R., Mizuno, C.M., Picazo, A., Camacho, A. and Rodriguez-Valera, F. (2013) 'Metagenomics uncovers a new group of low GC and ultra-small marine Actinobacteria', *Scientific Reports*, *3*(1), pp. 1-8.

Goodfellow, M. and Jones, A. L. (2015) ' *Corynebacteriales* ord. nov. ', *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–14.

Hiraoka, S., Yang, C. C. and Iwasaki, W. (2016) 'Metagenomics and bioinformatics in microbial ecology: Current status and beyond', *Microbes and Environments*, *31*(3), pp. 204–212.

Huang, Z., Mo, S., Yan, L., Wei, X., Huang, Y., Zhang, L., Zhang, S., Liu, J., Xiao, Q., Lin, H. and Guo, Y. (2021) 'A Simple Culture Method Enhances the Recovery of Culturable Actinobacteria From Coastal Sediments', *Frontiers in Microbiology*, p.1451.

Jian, S.L., Xu, L., Meng, F.X., Sun, C. and Xu, X.W. (2021) '*Euzebya pacifica* sp. nov., a novel member of the class *Nitriliruptoria*', *International Journal of Systematic and Evolutionary Microbiology*, *71*(7), p. 004864.

Kurahashi, M., Fukunaga, Y., Sakiyama, Y., Harayama, S. and Yokota, A. (2010) '*Euzebya tangerina* gen. nov., sp. nov., a deeply branching marine actinobacterium isolated from the sea cucumber Holothuria edulis, and proposal of *Euzebyaceae* fam. nov., *Euzebyales* ord. nov. and *Nitriliruptoridae* subclassis nov.', *International Journal of Systematic and Evolutionary Microbiology*, *60*(10), pp. 2314–2319.

Kurilkina, M.I., Zakharova, Y.R., Galachyants, Y.P., Petrova, D.P., Bukin, Y.S., Domysheva, V.M., Blinov, V.V. and Likhoshway, Y.V. (2016) 'Bacterial community composition in the water column of the deepest freshwater Lake Baikal as determined by next-generation sequencing', *FEMS Microbiology Ecology*, *92*, p. 94.

Li, X. and Qin, L. (2005) 'Metagenomics-based drug discovery and marine microbial diversity', *Trends in Biotechnology*, *23*(11), pp. 539–543.

López-Pérez, M., Haro-Moreno, J.M., Iranzo, J. and Rodriguez-Valera, F. (2020) 'Genomes of the "*Candidatus* Actinomarinales" Order: Highly Streamlined Marine Epipelagic Actinobacteria.', *Msystems*, *5*(6), pp.e01041-20.

Morrison, C. R., Bachoon, D. S. and Gates, K. W. (2008) 'Quantification of enterococci and bifidobacteria in Georgia estuaries using conventional and molecular methods', *Water Research*, *42*(14), pp. 4001–4009.

Ng, Z. Y. and Tan, G. Y. A. (2018) 'Selective isolation and characterisation of novel members of the family Nocardiopsaceae and other actinobacteria from a marine sediment of Tioman Island', *Antonie van Leeuwenhoek*, *111*(5), pp. 727–742.

Normand, P., Daffonchio, D. and Gtari, M. (2014) 'The family *Geodermatophilaceae*', *The Prokaryotes: Actinobacteria*, pp. 361–379.

Normand, P. and Fernandez, M. P. (2020) '*Frankiales*', *Bergey's Manual of Systematics of Archaea and Bacteria*, pp. 1–3.

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G.A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C. and Iliopoulos, L. (2015) 'Metagenomics: Tools and insights for analyzing next-generation sequencing data derived from biodiversity studies', *Bioinformatics and Biology Insights*, *9*, pp. 75–88.

Platas, G., Morón, R., González, I., Collado, J., Genilloud, O., Peláez, F. and Diez, M.T. (1998) 'Production of antibacterial activities by members of the family Pseudonocardiaceae: Influence of nutrients', *World Journal of Microbiology and Biotechnology*, *14*(4), pp.521-527.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) 'The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools', *Nucleic Acids Research*, *41*, pp. D590-D596.

Rehan, M., Swanson, E. and Tisa, L. S. (2016) '*Frankia* as a Biodegrading Agent', *Actinobacteria - Basics and Biotechnological Applications*.

Resnick, I. G. and Levin, M. A. (1981) 'Assessment of bifidobacteria as indicators of human fecal pollution', *Applied and Environmental Microbiology*, *42*(3), pp. 433–438.

Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. (2016) 'VSEARCH: a versatile open source tool for metagenomics', *PeerJ*, *4*, p.e2584.

Salam, N., Jiao, J.Y., Zhang, X.T. and Li, W.J. (2020) 'Update on the classification of higher ranks in the phylum Actinobacteria', *International Journal of Systematic and Evolutionary Microbiology*, *70*(2), pp. 1331–1355.

Salmonová, H. and Bunešová, V. (2017) 'Methods of Studying Diversity of Bacterial Comunities: A Review', *Scientia Agriculturae Bohemica*, *48*(3), pp. 154–165.

Schäfer, J., Jäckel, U. and Kämpfer, P. (2010) 'Development of a new PCR primer system for selective amplification of Actinobacteria', *FEMS Microbiology Letters*, *311*(2), pp. 103–112.

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. and Sahl, J.W. (2009) 'Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities', *Applied and Environmental Microbiology*, *75*(23), pp. 7537–7541.

Shi, S.B., Cui, L.Q., Zeng, Q., Long, L.J. and Tian, X.P. (2022) '*Nocardioides coralli* sp. nov., an actinobacterium isolated from stony coral in the South China Sea', *International Journal of Systematic and Evolutionary Microbiology*, *72*(5), p. 005342.

Slaby, B.M., Hackl, T., Horn, H., Bayer, K. and Hentschel, U. (2017) 'Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization', *Nature Publishing Group*, *11*, pp. 2465–2478.

Srivastava, A., Singh, A., Singh, S.S. and Mishra, A.K. (2017) 'Salt stress-induced changes in antioxidative defense system and proteome profiles of salt-tolerant and sensitive *Frankia* strains', *Journal of environmental science and health , 52*(5), pp. 420–428.

Stackebrandt, E. (2014) 'The Family *Propionibacteriaceae*: Genera other than *Propionibacterium*', *The Prokaryotes: Actinobacteria*, pp. 725–741.

Tóth, E. M. and Borsodi, A. K. (2014) 'The Family *Nocardioidaceae*', *The Prokaryotes: Actinobacteria*, pp. 651–694.

Trujillo, M. E., Hong, K. and Genilloud, O. (2014) 'The Family *Micromonosporaceae*', *The Prokaryotes: Actinobacteria*, pp. 499–569.

Wijnholds, A. E. and Young, D. R. (2000) 'Interdependence of *Myrica cerifera* seedlings and the nodule forming actinomycete, *Frankia*, in a coastal environment', *Journal of Coastal Research*, *16*(1), pp. 139–144.

Wright, L., Katouli, M. and Kurtböke, D.İ. (2021) 'Isolation and Characterization of Nocardiae Associated with Foaming Coastal Marine Waters', *Pathogens*, *10*(5), p.579.

Yin, Q., Zhang, L., Song, Z.M., Wu, Y., Hu, Z.L., Zhang, X.H., Zhang, Y., Yu, M. and Xu, Y. (2018) '*Euzebya rosea* sp. nov., a rare actinobacterium isolated from the east China sea and analysis of two genome sequences in the genus *Euzebya*', *International Journal of Systematic and Evolutionary Microbiology*, *68*(9), pp. 2900–2905.

Zhang, G., Gao, B., Adeolu, M., Khadka, B. and Gupta, R.S.(2016) 'Phylogenomic Analyses and Comparative Studies on Genomes of the Bifidobacteriales: Identification of Molecular Signatures Specific for the Order Bifidobacteriales and Its Different Subclades', *Frontiers in Microbiology*, *7*, p. 978.

Zhang, L., Chen, F., Zeng, Z., Xu, M., Sun, F., Yang, L., Bi, X., Lin, Y., Gao, Y., Hao, H. and Yi, W. (2021) 'Advances in Metagenomics and Its Application in Environmental Microorganisms', *Frontiers in Microbiology*, *12*, p. 766364.

CHAPTER 4:

SCREENING OF ACTINOBACTERIAL ISOLATES FOR ANTIBACTERIAL ACTIVITY

SUMMARY

Twenty six (26) *Streptomyces* strains, one *Nocardia* strain and two presumptive actinobacterial strains isolated from marine sediment were screened for antimycobacterial activity against the non-pathogenic bacterium *Mycobacterium aurum* strain A+ using a standard agar overlay method. The antibiotic spectra of the isolates were then investigated against *Escherichia coli* strain ATCC 25922 (Gram-negative) and *Staphylococcus aureus* strain ATCC 25923 (Gram-positive). Twenty one (21) out of 29 isolates (20 *Streptomyces* strains and one unidentified strain) demonstrated strong to very strong antimycobacterial activity. Only two isolates (*Streptomyces* strain GGUI#10 and *Nocardia* strain GGUI10#20) were active against *E. coli* ATCC 25922 and only six strains (*Streptomyces* strains GGUI10#6, GGUI#9, GGUM1#11, GGUI5#25 and GGUM1#29 and presumptive actinobacterial strain GGUI5#21) displayed notable activity (moderate to very strong) against *S. aureus* ATCC 25923. Notably, strain GGUI10#6 was the only isolate to display very strong activity against both *M. aurum* strain A+ and *S. aureus* ATCC 25923. The compounds that displayed strong antimycobacterial activity were extracted and analysed using HPLC-MS. In order to assess the novelty of the extracted compounds, the resulting mass spectra were compared to the mass spectra of known compounds within the Global Natural Products Social Molecular Networking (GNPS) database. Eighteen (18) out of 21 strains produced compounds with no matches in the GNPS database (*Streptomyces* strains GGUI#1, GGUM1#4, GGUI10#6, GGUI#7, GGUI10#8, GGUI#9, GGUI#10, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI#23, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33 and presumptive actinobacterial strain GGUI5#21) while, one isolate (*Streptomyces* strain GGUI5#33) produced a compound that could be an analogue of abyssomicin L, a member of a recently discovered rare class of antibiotics. Overall, the results presented in this chapter indicate that actinobacterial strains isolated from the marine environment could be a good source of potentially novel compounds with strong antimycobacterial activity.

Table of Contents

4.1 INTRODUCTION

According to the WHO, antibiotic resistance is one of the biggest threats to global health, food security and development today. Consequently, novel antibiotics with novel modes of action are urgently needed (Hutchings *et al.*, 2020). This has a caused a shift in research to the enhancement and activation of biosynthetic genes to increase the chances of discovering novel compounds (Ochi, 2016). Actinobacteria, particularly actinomycetes, are of particular interest in this field as they are prolific antibiotic producers. Various factors have been shown to influence biosynthetic gene expression in actinomycetes, including morphological differentiation and environmental stresses such as nutrient deprivation, heat shock and extreme pH levels (Yepes-García *et al.*, 2020; Zong *et al.*, 2021). Furthermore, the occurrence of competitors in the environment has also been shown to induce the biosynthesis of certain antibiotics, leading researchers to believe that co-culturing methods could be used to enhance antibiotic production (Westhoff *et al.*, 2021). In the same vein, researchers have also discovered that other secondary metabolites with regulatory functions (including known antibiotics) can be utilised to activate antibiotic biosynthetic pathways (Niu *et al.*, 2016; Okada and Seyedsayamdost, 2017). Genome mining has allowed researchers to gain more insight into biosynthetic genes and the pathways in which they are involved (Niu and Li, 2019). That being said, there is still a great deal unknown about the complex regulatory networks associated with SM production and more insight is required to develop new strategies for novel antibiotic discovery. In addition to the manipulation of biosynthetic gene regulation, researchers are also focused on exploring understudied habitats as sources of novel antibiotics.

One of the most concerning cases of acquired antibiotic resistance is that of *M. tuberculosis* (*Mtb*), the bacterial pathogen which causes TB. Worldwide, TB is the 13th leading cause of death and the second leading infectious killer after SARS-CoV-2 (the causal agent of COVID-19) and, in 2020, only about one in three people with drug-resistant TB accessed treatment (https://www.who.int/news-room/fact-sheets/detail/tuberculosis). Antibiotic resistance in TB can be separated into four categories. Single drug-resistant TB (SDR-TB) is defined as resistance of *Mtb* to only one first-line drug. First-line drugs include isoniazid, **r**ifampin, moxifloxacin, ethambutol and pyrazinamide. Multi-drug-resistant TB (MDR-TB) is defined as resistance of *Mtb* to both rifampin and isoniazid (i.e. the core first-line drugs). Extensively drug-resistant TB (XDR-TB) is defined as MDR-TB coupled with the resistance of *Mtb* to any fluoroquinolone and at least one injectable second-line drug (amikacin,

kanamycin, or capreomycin). Totally drug-resistant (TDR-TB) is defined as resistance of *Mtb* to all available first-line and second-line drugs (https://www.cdc.gov/tb/topic/drtb/default.htm; Singh and Chibale, 2021). Unlike other bacterial pathogens, resistance in *Mtb* is not acquired via HGT, but by spontaneous mutations (i.e. single nucleotide polymorphisms, insertions and deletions) in genes that code for drug targets (Dookie *et al.*, 2018). Therefore, it is possible to predict resistance in *Mtb* clinical isolates to first- and second-line drugs using NGS-based tests in order to develop effective treatment regimens (Migliori *et al.*, 2020). Current regimens are usually composed of four drugs likely to be effective of which at least two are essential and two are companion drugs (Rendon *et al.*, 2016). TDR-TB it is practically untreatable. Non-adherence of patients to treatment regimens is one of the main driving forces of drug resistance in TB. Thus, although the only way to truly fight drug resistance is the introduction of new effective drugs, using available drugs in the correct manner could still prevent the spread of drug-resistant TB (Migliori *et al.*, 2020). South Africa had an estimated incidence of 328 000 (554 per 100 000 population) cases of active TB in 2020 (World Health Organization, 2021). This resulted in 61 000 TB related deaths. Approximately 7517 of the aforementioned cases were laboratory-confirmed instances of MDR-TB and XDR-TB.

This chapter investigates antibiotic production by actinobacterial isolates from sediment from three different habitats (subtidal zone, rock pool and beach/dune) in the marine environment. The isolates were screened against *Mycobacterium aurum* strain A+, a non-pathogenic species with a similar antibiotic susceptibility profile to that of *M. tuberculosis* (Chung *et al.*, 1995). Therefore, screening for activity against *M. aurum* strain A+ could give us more insight into the potential of the actinobacterial isolates to produce compounds that could also be effective against *Mtb*. The ability of each actinobacterial isolate to produce antibacterial compounds was further investigated by screening for activity against another Gram- positive bacterium (*Staphylococcus aureus* strain ATCC 25923) and a Gram-negative bacterium (*Escherichia coli* strain ATCC 25922) with clinical relevance of their own. Additionally, the novelty of strong antimycobacterial compounds was determined by Dr Daniel Watson (Division of Clinical Pharmacology, Faculty of Health Sciences, University of Cape Town) using HPLC-analysis. This was done to prioritize potentially novel compounds for further study.

4.2 METHODS AND MATERIALS

4.2.1 Screening for antibacterial activity

Presumptive actinobacterial isolates were screened against *M. aurum* A+, *E. coli* ATCC 25922 and *S. aureus* ATCC 25923 using a standard agar-overlay technique (Curtis, 2015). Isolates were stab inoculated onto ISP2 (without calcium carbonate), ISP5, M1 (section 2.2.1.3), Difco Middlebrook 7H9 containing 100mM glucose (Becton Dickinson, USA), and R2A medium (https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium830.pdf) agar plates supplemented with 3.8% Red Sea Salt using sterile toothpicks. All agar plates were incubated at 30°C for 10 days. The day before overlays were performed, *M. aurum* A+ was inoculated into 5ml 2xYT medium [16g tryptone, 10g yeast extract, 5g NaCl, distilled water to 1L, pH 7.0] and *E. coli* ATCC 25922 and *S. aureus* ATCC 25923 were inoculated into 5ml Luria-Bertani medium (LB) (Sambrook *et al*., 1989) before being incubated overnight at 37°C with shaking. The bacterial cultures were examined by streaking for single colonies on appropriate agar plates and Gram-staining in order to ensure no contamination was present. In order to ensure the concentration of cells was standardized across all overlay experiments, the optical density at 600nm (OD600nm) of each culture was measured using a Beckman DU®-64 spectrophotometer. The optical density reading was then used to determine the volume of culture to add to each plate using the following empirical formulae: OD600nm × volume (µl) = 160 for *M. aurum* A+ and OD600nm ×volume (µl) = 4 for *E. coli* ATCC 25922 and *S. aureus* ATCC 25923. The determined volume of culture was inoculated into test tubes containing 6ml 2YT/LB sloppy agar (0.7% agar) and poured over the stab inoculated plates. The overlaid agar was allowed to set and the plates were incubated at 37°C overnight for screening against *E. coli* ATCC 25922 and *S. aureus* ATCC 25923 and for 48 hours for screening against *M. aurum* A+. Antibacterial activity was observed as zones of growth inhibition of test bacteria around the actinobacterial colonies. The area of the zone of inhibition (ZOI) was calculated by subtracting the area of the colony (mm^2) from the area of the zone of growth inhibition (mm^2). ZOIs less than 1000 mm^2 indicated weak activity, ZOIs between 1000 mm^2-2000mm^2 indicated moderate activity, ZOIs between 2000mm^2-3000mm^2 indicated strong activity and ZOIs of more than 3000mm^2 indicated very strong activity.

4.2.2 Identification of active antimycobacterial compounds

4.2.2.1 Preparation of crude extract

Actinobacterial isolates displaying strong and very strong antimycobacterial activity were stabbed on fresh plates (one strain per plate) to confirm the size of the ZOIs. The ZOIs were excised as 1x1 cm pieces of agar from the overlaid plates and stored at -20°C before extraction. The extraction procedure detailed below was performed by Dr Daniel Watson. Agar pieces were thawed and added to a mixture of 5 mL methanol and 35 mL ethyl acetate (EtAc) and left to shake on a multipurpose shaker at 120 rpm for at least one hour at room temperature. The resultant liquid samples were washed with 40 mL Millipore water in a separating funnel to remove salts and polar constituents that could affect the HPLC-MS analysis. The EtAc fraction was collected and dried in a fume hood. The dried sample was concentrated by re-dissolving the fraction in 1.5 ml of EtAc and transferred to a previously weighed microcentrifuge tube. The samples were dried down again using a sample concentrator and the weight of the sample was determined. The final crude extract was made up to a concentration of 10 mg/mL using ethyl acetate.

4.2.2.2 HPLC – MS analysis and dereplication

The HPLC-MS and dereplication analyses were also performed by Dr Daniel Watson. These analyses were based on the spectroscopic analysis and Global Natural Products Social Molecular Networking protocols set out in Watson *et al.* (2021) with a few modifications. For the spectroscopic analysis, the HPLC method utilised a C18 HPLC column (Agilent Poroshell 120). The aqueous mobile phase used was 1 mM ammonium formate in water and the organic mobile phase was MeOH/0.1% formic acid. A gradient method was used that ran from 2% to 98% organic phase for 25 minutes with a flow rate of 600 µl/min. The gradient was then held at 98% organic phase for a further 2 minutes before returning to 2% organic phase to equilibrate the column for the next run. Spectral data were obtained using information dependent acquisition (IDA) at a mass range of 50–1500 Da. For the Global Natural Products Social Molecular Networking protocol, a network was created where edges were filtered to have a cosine score above 0.7 and more than 5 matched peaks. All matches kept between network spectra and library spectra were also required to have a cosine score above 0.70 and at least 5 matched peaks.

4.3 RESULTS AND DISCUSSION

4.3.1 Screening for antibacterial activity

4.3.1.1 Antimycobacterial activity against *M. aurum* strain A+

Twenty-seven confirmed and two presumptive actinobacterial strains (described in Chapter 2) were screened for antimycobacterial activity against *M. aurum* strain A+. Seven strains displayed weak or no activity (*Streptomyces* strains GGUI#3, GGUM1#5, GGUI5#24, GGUI5#25, GGUM1#26, *Nocardia* strain GGUI10#20 and presumptive actinobacterial strain GGUI5#32), one strain displayed moderate activity (*Streptomyces* strain GGUM1#2), four strains (*Streptomyces* strains GGUM1#2, GGUI#16, GGUI#28 and GGUM1#29) displayed strong activity and seventeen strains (*Streptomyces* strains GGUI#1, GGUM1#4, GGUI10#6, GGUI#7, GGUI10#8, GGUI#9, GGUI#10, GGUM1#11, GGUI#12, GGUI10#15, GGUI5#18, GGUI#23, GGUI5#27, GGUI5#30, GGUI5#31, GGUI5#33 and presumptive actinobacterial strain GGUI5#21), displayed very strong activity against *M. aurum* strain A+ (Table 4.1). The results show that strong antimycobacterial activity was observed for strains isolated from all three sub-sites. The novelty of the compounds produced by strains displaying strong and very strong antimycobacterial activity was assessed (see section 4.3.2).

The nutrient composition of the environment is known to influence antibiotic biosynthesis (Martin and Demain, 1980). This is suggested as an explanation for why all the aforementioned strains displayed strong activity on at least one medium, but not on all five media. The media that induced the strongest antimycobacterial activity were M1 and R2A. This is not surprising as nutrients in R2A are present in low concentrations and, as mentioned in section 4.1, antibiotic biosynthesis has been shown to be regulated in response to environmental stresses such as nutrient deprivation. The media which induced the weakest antimycobacterial activity were ISP2 and 7H9. This could possibly be a result of the presence of glucose in ISP2 and 7H9, as glucose is known to inhibit the biosynthesis of several antibiotics (Martin and Demain, 1980; Grasso *et al.*, 2016). Despite this, strong and very strong activity was still observed for strains GGUI5#30 and GGUM1#4 on ISP2. Interestingly, strain GGUI5#30 was also the only strain to display moderate to very strong activity against *M. aurum* strain A+ on all five test media. ISP5 is a defined medium and it was observed that growth of some actinobacterial strains was weaker on this medium compared to complex media such as ISP2 and M1. Martin and Demain (1980) suggested that

cultivation on defined media supporting only slow growth might favour antibiotic production, therefore, it is surprising that the activity induced on ISP5 was moderate for the most part. However, it is possible that the nutrients in this specific defined medium (ISP5) were not able to induce strong antibiotic synthesis in the isolates under investigation.

4.3.1.2 Antibacterial activity against *E. coli* and *S. aureus*

Antibiotic activity was also tested against *S. aureus* strain ATCC 25923 (Gram-positive bacterium) and *E. coli* strain ATCC 25922 (Gram-negative bacterium) to determine the antibiotic spectrum of each isolate tested in section 4.3.1.1. Only two isolates (*Streptomyces* strain GGUI#10 and *Nocardia* strain GGUI10#20) displayed activity against *E. coli* strain ATCC 25922 (Table 4.1). Strain GGUI#10 displayed moderate activity (1166 mm^2) and strain GGUI10#20 displayed weak to moderate activity (980 mm^2). It is interesting that *Nocardia* strain GGUI10#20 displayed higher activity against *E. coli* ATCC 25922 than *M. aurum* A+ or *S. aureus* ATCC 25923, as, unlike Gram-positive bacteria, Gram-negative bacteria have an outer membrane which limits the access of most antibiotics to their intracellular targets (Breijyeh *et al.*, 2020). Table 4.1 shows only six isolates displayed notable (i.e. moderate to very strong) antibiotic activity against *S. aureus* strain ATCC 25923. *Streptomyces* strains GGUI#9, GGUM1#11, GGUI5#25 and GGUM1#29 displayed moderate activity and presumptive actinobacterial strain GGUI5#21 displayed strong activity against *S. aureus* ATCC 25923 on at least one of the five test media. Notably, *Streptomyces* strain GGUI10#6 displayed very strong activity against *S. aureus* strain ATCC 25923 on all five test media.

Five strains (*Streptomyces* strains GGUI#7 and GGUI10#19, GGUI5#24 and GGUM1#26 and presumptive actinobacterial strain GGUI5#32) showed activity against *M. aurum* strain A+ only (weak to very strong activity), suggesting that the antibiotics they are producing are specific to *Mycobacterium*. *Streptomyces* strains GGUI#3, GGUM1#5, GGUI5#24, GGUM1#26, *Nocardia* strain GGUI10#20 and strain GGUI5#32 displayed weak or no activity against all test bacteria. However, this does not mean these strains are incapable of producing strongly active antibacterial compounds. It is possible that the compounds produced by these strains were not active against the test bacteria used in this study, or that production of active compounds by these strains was not induced on the test media used.

Table 4.1. The antibiotic activity of all actinobacterial isolates against test bacteria *M. aurum* strain A+, *S. aureus* strain ATCC 25923 and *E. coli* strain ATCC 25922. Each isolate was stab inoculated on ISP2, ISP5, M1, Middlebrook 7H9 (MB) and R2A media. The zone of inhibition (ZOI) was calculated by subtracting the area of the colony from the area of the zone of growth inhibition. Arbitrary assignment of the strength of antibiotic activity: - = no activity, ZOI < 1000 mm^2 = weak activity, 1000mm^2 < ZOI ≤ 2000mm^2 = moderate activity (highlighted in yellow), 2000mm^2 < ZOI ≤ 3000mm^2 = strong activity (highlighted in purple) and ZOI > 3000mm^2 = very strong activity (highlighted in red). Antimycobacterial compound identifications are shown next to the agar media from which the compound was extracted. 'No match' indicates the spectrum obtained for the compound did not match any known compounds in the GNPS database.

ISOLATE	MEDIA	*M. aurum* A+	*E. coli* ATCC 25922	*S. aureus* ATCC 25923	COMPOUND
OCEAN SITE					
***Streptomyces* strain GGUI#1**	**ISP2**	657	-	-	
	ISP5	679	-	-	
	M1	3196	-	433	No match
	MB	-	-	-	
	R2A	-	-	463	
***Streptomyces* strain GGUI10#6**	**ISP2**	748	-	4068	
	ISP5	694	-	4894	
	M1	1122	-	3213	
	MB	1074	-	3628	
	R2A	5014	-	3841	No Match
DRY SITE					
***Streptomyces* strain GGUM1#2**	**ISP2**	-	-	170	
	ISP5	1208	-	-	
	M1	2651	-	440	Rabelomycin
	MB	-	-	359	
	R2A	1080	-	283	
***Streptomyces* GGUI#3**	**ISP2**	160	-	337	
	ISP5	-	-	657	
	M1	-	-	-	
	MB	286	-	-	
	R2A	-	-	-	
***Streptomyces* strain GGUM1#4**	**ISP2**	3817	-	163	Desferrioxamine E
	ISP5	1206	-	-	
	M1	5717	-	-	No match
	MB	-	-	-	
	R2A	1837	-	262	
***Streptomyces* strain GGUI10#15**	**ISP2**	-	-	286	
	ISP5	1021	-	-	
	M1	3267	-	359	Monactin
	MB	-	-	402	
	R2A	2211	-	176	No match
***Streptomyces* strain GGUM1#29**	**ISP2**	1897	-	378	
	ISP5	2060	-	-	Antimycin
	M1	1435	-	189	
	MB	1103	-	-	
	R2A	399	-	1179	

Table 4.1. *(continued)*

ISOLATE	MEDIA	*M. aurum* A+	*E. coli* ATCC 25922	*S. aureus* ATCC 25923	COMPOUND
ROCKY SITE					
***Streptomyces* strain GGUM1#5**	**ISP2**	691	-	-	
	ISP5	236	-	236	
	M1	-	-	-	
	MB	-	-	-	
	R2A	-	-	-	
***Streptomyces* strain GGUI#7**	**ISP2**	-	-	-	
	ISP5	-	-	-	
	M1	-	-	-	
	MB	302	-	-	
	R2A	4287	-	-	No Match
***Streptomyces* strain GGUI10#8**	**ISP2**	-	-	320	
	ISP5	1680	-	-	
	M1	3003	-	-	No match
	MB	-	-	717	
	R2A	1184	-	-	
***Streptomyces* strain GGUI#9**	**ISP2**	357	-	-	
	ISP5	1124	-	-	
	M1	3349	-	-	No Match
	MB	-	-	1179	
	R2A	1207	-	-	
***Streptomyces* strain GGUI#10**	**ISP2**	672	-	163	
	ISP5	1457	-	191	
	M1	3844	-	204	No match
	MB	-	1166	705	
	R2A	1529	-	403	
***Streptomyces* strain GGUM1#11**	**ISP2**	411	-	402	
	ISP5	1131	-	219	
	M1	3377	-	-	Monactin
	MB	-	-	1145	
	R2A	1850	-	126	
***Streptomyces* strain GGUI#12**	**ISP2**	339	-	-	
	ISP5	1178	-	-	
	M1	3656	-	-	Monactin/Nonactin
	MB	-	-	453	
	R2A	2313	-	-	No match
***Streptomyces* strain GGUI#16**	**ISP2**	285	-	-	
	ISP5	754	-	-	
	M1	2510	-	-	No match
	MB	-	-	503	
	R2A	2045	-	-	Monactin
***Streptomyces* strain GGUI5#18**	**ISP2**	1021	-	236	
	ISP5	884	-	201	
	M1	1922	-	-	
	MB	-	-	817	
	R2A	3817	-	-	No Match

Table 4.1. *(continued)*

ISOLATE	MEDIA	*M. aurum* A+	*E. coli* ATCC 25922	*S. aureus* ATCC 25923	COMPOUND
Streptomyces	**ISP2**	308	-	-	
strain	**ISP5**	1070	-	-	
GGUI10#19	**M1**	-	-	-	
	MB	-	-	-	
	R2A	-	-	-	
***Nocardia* strain**	**ISP2**	337	-	-	
GGUI10#20	**ISP5**	-	-	-	
	M1	-	980	-	
	MB	-	-	-	
	R2A	216	-	-	
Presumptive	**ISP2**	-	-	776	
actinobacterial	**ISP5**	1935	-	2814	
strain	**M1**	2262	-	1254	No match
GGUI5#21	**MB**	-	-	-	
	R2A	3820	-	295	Possible Mitomycin C analogue
Streptomyces	**ISP2**	214	-	579	
strain	**ISP5**	3885	-	-	Streptomycin
GGUI#23	**M1**	4713	-	-	No match
	MB	361	-	939	
	R2A	-	-	-	
Streptomyces	**ISP2**	679	-	-	
strain	**ISP5**	-	-	-	
GGUI5#24	**M1**	-	-	-	
	MB	-	-	-	
	R2A	521	-	-	
Streptomyces	**ISP2**	-	-	1272	
strain	**ISP5**	-	-	-	
GGUI5#25	**M1**	-	-	493	
	MB	-	-	699	
	R2A	181	-	463	
Streptomyces	**ISP2**	-	-	-	
strain	**ISP5**	283	-	-	
GGUM1#26	**M1**	-	-	-	
	MB	-	-	-	
	R2A	330	-	-	
Streptomyces	**ISP2**	463	-	219	
strain	**ISP5**	3735	-	-	Streptomycin
GGUI5#27	**M1**	2626	-	-	No match
	MB	-	-	955	
	R2A	2956	-	-	No match
Streptomyces	**ISP2**	-	-	327	
strain	**ISP5**	537	-	139	
GGUI#28	**M1**	1124	-	320	
	MB	-	-	479	
	R2A	2509	-	-	No Match

Table 4.1. *(continued)*

ISOLATE	MEDIA	*M. aurum* A+	*E. coli* ATCC 25922	*S. aureus* ATCC 25923	COMPOUND
Streptomyces	**ISP2**	2158	-	-	No match
strain	**ISP5**	1168	-	179	
GGUI5#30	**M1**	2989	-	-	No match
	MB	1056	-	968	
	R2A	4123	-	-	No Match
Streptomyces	**ISP2**	1209	-	-	
strain	**ISP5**	1021	-	201	
GGUI5#31	**M1**	3694	-	203	No Match
	MB	-	-	754	
	R2A	506	-	-	
Presumptive	**ISP2**	-	-	-	
actinobacterial	**ISP5**	317	-	-	
strain	**M1**	-	-	-	
GGUI5#32	**MB**	-	-	-	
	R2A	302	-	-	
Streptomyces	**ISP2**	1125	-	182	
strain	**ISP5**	1850	-	-	
GGUI5#33	**M1**	4711	-	-	Possible Abyssomycin L analogue
	MB	-	-	633	
	R2A	3455	-	-	No Match

4.3.2 Dereplication

Eighteen (18) of the isolates described in this study (Strains GGUI#1, GGUM1#4, GGUI10#6, GGUI#7, GGUI10#8, GGUI#9, GGUI#10, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#21, GGUI#23, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33) produced at least one compound with strong antimycobacterial activity that did not match any known compounds in the GNPS database (Table 4.1). These compounds with strong antimycobacterial activity may represent novel antibiotics and, therefore, are worth investigating further. The strains mentioned above include isolates from all three sub-sites. This illustrates the great potential of marine habitats as a source of novel antimycobacterial compound producing actinobacterial strains. These data need to be confirmed by re-testing against *M. aurum* strain A+ using agar overlays. Thereafter, notable antitubercular activity can be determined by cultivating the strains responsible for the production of the unknown compounds in liquid versions of the media in which the compounds were produced and then testing for activity against clinically relevant strains of *Mtb*. Compounds which show strong activity against *Mtb* can be prioritized for purification and structural elucidation. Considering some of the unknown compounds were produced on the same media by strains identified to be potential clones in Chapter 2 (e.g. GGUI5#18, GGUI5#27, GGUI5#30 and GGUI5#33 all produced unknown compounds on R2A), it is possible these strains are producing the same active antimycobacterial compound. This can be investigated further by comparing the mass spectra of these unknown compounds and determining if their fragmentation patterns are the same. It is important to note that inferences made about the novelty of the compounds produced in this study are based only on the information available in the GNPS database (Wang *et al.*, 2016) and other natural product databases could be used to support these results.

Eleven of the isolates described in this study were shown to produce compounds that matched known compounds in the GNPS database (Table 4.2). These include *Streptomyces* strains GGUM1#2, GGUM1#4, GGUM1#11, GGUI#12, GGUI10#15, GGUI#16, GGUI#23, GGUI5#27, GGUM1#29, GGUI5#33 and presumptive actinobacterial strain GGUI5#21.The similarity between the query compounds and the known antibiotics was evaluated using the cosine score and mass difference parameters. The cosine score is the measure of similarity between the fragmentation patterns of a query compound and a known compound in the GNPS library. A cosine score of 1 represents identical spectra while a cosine score of 0 denotes no similarity at all (https://ccms-ucsd.github.io/GNPSDocumentation/lexicon/). Only

matches with a cosine score of 0.70 and above are shown in this study. Mass difference is a quantification of the shift in spectral peaks when comparing the mass spectra of the query and known compounds. The greater the mass difference, the less closely the spectra are aligned. Some mass differences may be attributable to an adduct formed during ionization before analysis by mass spectrometry.

Table 4.2 GNPS database matches for compounds produced by actinobacterial isolates on various growth media. Cosine Score and Mass Difference indicate the similarity between the mass spectra of the query compound and known compound.

Isolate	Medium	Compound	Cosine Score	Mass Difference (Da)
***Streptomyces* strain GGUM1#2**	M1	Rabelomycin	0.74	0.013
***Streptomyces* strain GGUM1#4**	ISP2	Desferrioxamine E	0.85	0.001
***Streptomyces* strain GGUM1#11**	M1	Monactin	0.80	0.002
***Streptomyces* strain GGUI#12**	M1	Monactin	0.83	0.24
		Nonactin	0.88	17.02
***Streptomyces* strain GGUI10#15**	M1	Monactin	0.73	0.002
***Streptomyces* strain GGUI#16**	R2A	Monactin	0.71	0.01
Presumptive actinobacterial strain GGUI5#21	R2A	Possible Mitomycin C analogue	0.76	45,09
***Streptomyces* strain GGUI#23**	ISP5	Streptomycin	0.82	11.5
***Streptomyces* strain GGUI5#27**	ISP5	Streptomycin	0.81	11.5
***Streptomyces* strain GGUM1#29**	ISP5	Antimycin A3	0.91	42
		Antimycin A4	0.93	28
***Streptomyces* strain GGUI5#33**	M1	Possible Abyssomicin L analogue	0.87	7.04

Table 4.2 shows that the compound produced by *Streptomyces* strain GGUM1#2 on M1 medium matched the antibiotic rabelomycin with a moderate cosine score of 0.74 and a very low mass difference of 0.013 Da. Rabelomycin is an angucycline antibiotic isolated from

Streptomyces olivaceus (Liu *et al.*, 1970) that has been reported to have antibacterial activity against a range of Gram-positive bacteria including *Mycobacterium smegmatis* and *Staphylococcus aureus* (Bao *et al.*, 2018). Thus, since the mass difference is also too low to be considered significant, it is very likely that rabelomycin is the active antimycobacterial compound produced by strain GGUM1#2 on M1 medium.

The compound produced by *Streptomyces* strain GGUM1#4 on ISP2 medium was identified to be desferrioxamine E with a high cosine score of 0.85 and a negligible mass difference of 0.001 Da (Table 4.2). Desferrioxamine E, also known as nocardamine, is a siderophore isolated from a *Nocardia* species reported to display specific activity against mycobacteria (Stoll *et al.*, 1951; Ishida *et al.*, 2011; Thompson *et al.*, 2012) This compound is produced by several species of actinobacteria and has been shown to stimulate secondary metabolite formation and morphological differentiation in various actinomycete strains (Yamanaka *et al.*, 2005). The desferrioxamine biosynthetic pathway belongs to a family of pathways that are biochemically distinct from the NRPS and PK pathways (Barona-Gómez *et al.*, 2004). Siderophores have been shown to have higher production when a monosaccharide such as glucose is used as a carbon source (Franco-Correa and Chavarro-Anzola, 2016). This might explain why strain GGUM1#4 was the only isolate to exhibit very strong antimycobacterial activity on ISP2 medium. Based on the data, it very likely that, desferrioxamine E is the active antimycobacterial compound produced by strain GGUM1#4 on ISP2 medium.

Streptomyces strains GGUM1#11, GGUI#12, GGUI10#15 and GGUI#16 were shown to produce compounds on M1 and R2A media matching the antibiotic monactin with moderate to high cosine scores (0.71-0.83) and negligible mass differences (Table 4.2). Furthermore, strain GGUI#12 was also shown to produce the homologue, nonactin. This match was supported by a high cosine score of 0.88 and a mass difference of 17.02 Da attributable to an ammonium adduct. Nonactin and monactin are members of a family of cyclic ionophores known as macrotetrolide antibiotics (Corbaz *et al.*, 1955; Gerlach and Prelog, 1963). Monactin is a homologue of nonactin containing an additional methyl group. These compounds are produced by a range of *Streptomyces* species and are potent uncouplers of oxidative phosphorylation (Graven *et al.*, 1965). Nonactins were originally thought to have little to no antimicrobial activity (hence the name), however, they have been shown to be active *in vitro* against Gram-positive bacteria such as *Mycobacterium bovis* and *S. aureus* (Meyers *et al.*, 1965). Nonactin was noted as being the least active of the homologues. It has

also been noted that the antibacterial effects of nonactin depend upon its ability to support the passive diffusion of cations, such as K^+, Na^+ and NH_4^+ across cell membranes (Kusche *et al.*, 2010). *Streptomyces* strains GGUM1#11, GGUI#12, GGUI10#15 and GGUI#16 were shown to be genetically similar in Chapter 2. Thus, it is possible these strains are producing the same antimycobacterial compound on M1 and R2A (strain GGUI#16). Based on the evidence presented above and in Table 4.2, it also likely that this compound is in fact monactin (produced alongside nonactin by strain GGUI#12).

Presumptive actinobacterial strain GGUI5#21 produced a compound on R2A medium that could be an analogue of the antibiotic mitomycin C with a moderate cosine score of 0.76, but a high mass difference of 45.09 (Table 4.2). Mitomycin C is a DNA crosslinking agent isolated from *Streptomyces caespitosus* (Szybalski and Iyer, 1964). This compound is known for its antitumor activity and is used as a chemotherapy drug in the treatment of bladder cancer (Zargar *et al.*, 2014). In terms of antibacterial activity, mitomycin C has also been reported to show some *in vitro* activity against growing *M. bovis* (Peh *et al.*, 2001). Therefore, it is possible that an analogue of mitomycin C could be the active antimycobacterial compound produced by strain GGUI5#21 on R2A medium.

Streptomyces strains GGUI#23 and GGUI5#27 both produced compounds on ISP5 medium matching the antibiotic streptomycin with high cosine scores of 0.82 and 0.81, respectively (Table 4.2). The corresponding mass differences for these matches were both 11.5 Da, indicating that the structure of the compounds produced by strains GGUI#23 and GGUI5#27 could be identical. Due to the mass difference, these compounds could also be analogues of streptomycin. After the isolation of streptomycin from *Streptomyces griseus* in 1943 (Schatz *et al.*, 1944), it became the first clinically relevant antibiotic to treat TB. It was used as a monotherapy for TB for 3-4 years before the multitherapy regimen was introduced to combat streptomycin-resistant strains of *Mtb* (Rocha *et al.*, 2021). Streptomycin currently has limited clinical application due to the high incidence of resistance in *Mtb* strains. It is highly likely the active antimycobacterial compound produced by strains GGUI#23 and GGUI5#27 is streptomycin or a streptomycin analogue.

Two compounds were co-produced by *Streptomyces* strain GGUM1#29 on ISP5 medium. These compounds matched antimycin A3 and antimycin A4 (homologues of the antibiotic antimycin A) with very high cosine scores of 0.91 (mass difference of 42 Da) and 0.93 (mass difference of 28 Da), respectively (Table 4.2). The difference in mass between successive

members of the homologous series of antimycin A is due to the loss of a methylene bridge (~14 Da) (https://www.ebi.ac.uk/chebi/init.do). Therefore, due to the mass differences observed, it is likely the compounds produced by strain GGUM1#29 are not antimycin A3 and A4, but rather other homologues within the antimycin series. Antimycin A is generally considered to have little to no antibacterial activity. Furthermore, the little antibacterial activity it does exhibit only occurs when the antibiotic is present in high concentration (Marquis, 1965). For this reason, it is unlikely the active antimycobacterial compound produced by strain GGUM1#29 is a homologue of antimycin A. It is possible that the active compound and antimycin A homologues were being co-produced by strain GGUM1#29 and the active compound went undetected as it did not match any known compound in the GNPS database.

The compound produced by *Streptomyces* strain GGUI5#33 on M1 medium was found to be a potential analogue of the antibiotic abyssomicin L with a high cosine score of 0.87 and a mass difference of 7.04 Da (Table 4.2). Abyssomicin L was isolated from a species of the actinomycete genus *Verrucosispora* (isolated from deep-sea sediment) and forms part of a rare class of polyketides (Wang *et al.*, 2013). All members of the genus *Verrucosispora* have since been reassigned to the genus *Micromonospora*; https://lpsn.dsmz.de/genus/verrucosispora). Abyssomicins inhibit the production of an intermediate of the folic acid pathway thereby decreasing the availability of folates for nucleic acid synthesis (Fiedler, 2021). Abyssomicin L has been reported to have antimycobacterial activity against *M. bovis* (Wang *et al.*, 2013). Therefore, it is possible that an analogue of this compound is active against *M. aurum* A+. Due to the novelty of the abyssomicin family, the chemical structure of a potential analogue of abyssomicin L might be worth studying further.

Strains GGUM1#4, GGUI#12, GGUI10#15, GGUI#16, GGUI5#21, GGUI#23, GGUI5#27 and GGUI5#33 were observed to produce compounds with matches in the GNPS database on certain media and compounds with no matches on other media (Table 4.1). This demonstrates the influence of nutrient composition on secondary metabolite production.

4.4 REFERENCES

Bao, J., He, F., Li, Y., Fang, L., Wang, K., Song, J., Zhou, J., Li, Q. and Zhang, H. (2018) 'Cytotoxic antibiotic angucyclines and actinomycins from the *Streptomyces* sp. XZHG99T', *The Journal of Antibiotics*, *71*(12), pp. 1018–1024.

Barona-Gómez, F., Wong, U., Giannakopulos, A. E., Derrick, P. J., & Challis, G. L. (2004) 'Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145', *Journal of the American Chemical Society*, *126*(50), pp.16282–16283.

Breijyeh, Z., Jubeh, B. and Karaman, R. (2020) 'Resistance of Gram-Negative Bacteria to Current Antibacterial Agents and Approaches to Resolve It', *Molecules*, *25*(6), p.1340

Chung, G. A. C., Aktar, Z., Jackson, S. & Duncan, K. (1995) 'High-throughput screen for detecting antimycobacterial agents', *Antimicrobial agents and chemotherapy*, *39*(10), pp.2235-2238.

Corbaz, R., Ettlinger, L., Gäumann, E., Keller-Schierlein, W., Kradolfer, F., Neipp, L., Prelog, V. and Zähner, H. (1955) 'Stoffwechselprodukte von Actinomyceten. 3. Mitteilung. Nonactin', *Helvetica Chimica Acta*, *38*(6), pp. 1445–1448.

Curtis, S. M. (2015) 'Enhanced phylogenetic analysis and targeted search for the genus *Kribbella'*, Ph.D. thesis. Department of Molecular and Cell Biology. University of Cape Town. Supervisor: Paul Meyers

Dookie, N., Rambaran, S., Padayatchi, N., Mahomed, S. and Naidoo, K. (2018) 'Evolution of drug resistance in *Mycobacterium tuberculosis*: A review on the molecular determinants of resistance and implications for personalized care', *Journal of Antimicrobial Chemotherapy*, *73*(5), pp. 1138–1151.

Fiedler, H. (2021) 'Abyssomicins — A 20-Year Retrospective View'. *Marine Drugs*, *19*(6), p.299.

Franco-Correa, M. and Chavarro-Anzola, V.(2016) 'Actinobacteria as plant growth promoting rhizobacteria' *Actinobacteria-basis and biotechnological application*, pp.249-270.

Gerlach, H. and Prelog, V. (1963) 'Über die Konfiguration der Nonactinsäure', *Justus Liebigs Annalen der Chemie*, 669(1), pp. 121–135.

Grasso, L. Lo, Martino, D. C. and Alduina, R. (2016) 'Production of Antibacterial Compounds from Actinomycetes', *Actinobacteria - Basics and Biotechnological Applications*, *214*(11), pp.272-282

Graven, S. N., Lardy, H. A. and Rutter, A. (1965) 'Antibiotics As Tools for Metabolic Studies. VI. Damped Oscillatory Swelling of Mitochondria Induced by Nonactin, Monactin, Dinactin, and Trinactin', *5*(5), pp. 1735–1742.

Hutchings, M. I., Truman, A. W. and Wilkinson, B. (2020) 'Antibiotics : past , present and future', *Current Opinion in Microbiology*, *51*, pp. 72–80.

Ishida, S., Arai, M., Niikawa, H. and Kobayashi, M. (2011) 'Inhibitory effect of cyclic trihydroxamate siderophore, desferrioxamine E, on the biofilm formation of *Mycobacterium* species.', *Biological & Pharmaceutical Bulletin*, *34*(6), pp. 917–920.

Kusche, B.R., Smith, A.E., McGuirl, M.A. and Priestley, N.D. (2010) 'The alternating pattern of stereochemistry in the nonactin macrocycle is required for antibacterial activity and efficient ion binding', *Journal of the American Chemical Society, 131*(47), pp. 17155–17165.

Liu, W.C., Parker, W.L., Slusarchyk, D.S., Greenwood, G.L., Graham, S.F. and Meyers, E. (1970) 'Isolation, characterization, and structure of rabelomycin, a new antibiotic', *The Journal of Antibiotics*, *23*(9), pp.437-441.

Marquis, R. E. (1965) 'Nature of the Bactericidal Action of Antimycin A for *Bacillus megaterium*', *Journal of Bacteriology*, *89*(6), pp.1453-1459.

Martin, J. F. and Demain, A. L. (1980) 'Control of antibiotic biosynthesis', *Microbiological Reviews*, *44*(2), pp. 230–251.

Meyers, E., Pansy, F.E., Perlman, D., Smith, D.A. and Weisenborn, F.L. (1965) 'The in vitro activity of nonactin and its homologs: monactin, dinactin and trinactin', *The Journal of Antibiotics*, *18*(3), pp.128-129

Migliori, G.B., Tiberi, S., Zumla, A., Petersen, E., Chakaya, J.M., Wejse, C., Torrico, M.M., Duarte, R., Alffenaar, J.W., Schaaf, H.S. and Marais, B.J. (2020) 'MDR/XDR-TB management of patients and contacts: Challenges facing the new decade. The 2020 clinical update by the Global Tuberculosis Network', *International Journal of Infectious Diseases*, *92*, pp. S15–S25.

Niu, G., Chater, K.F., Tian, Y., Zhang, J. and Tan, H. (2016) 'Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp', *FEMS microbiology reviews*, *40*(4), pp. 554–573.

Niu, G. and Li, W. (2019) 'Next-Generation Drug Discovery to Combat Antimicrobial Resistance', *Trends in Biochemical Sciences*, *44*, pp. 961–972.

Ochi, K. (2016) 'Insights into microbial cryptic gene activation and strain improvement: principle, application and technical aspects', *The Journal of Antibiotics*, *70*, pp. 25–40.

Okada, B. K. and Seyedsayamdost, M. R. (2017) 'Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules', *FEMS Microbiology Reviews*, *41*(1), pp. 19–33.

Peh, H.L., Toh, A., Murugasu-Oei, B. and Dick, T. (2001) 'In vitro activities of mitomycin C against growing and hypoxic dormant tubercle bacilli', *Antimicrobial Agents and Chemotherapy*, *45*(8), pp. 2403–2404.

Rendon, A., Tiberi, S., Scardigli, A., D'Ambrosio, L., Centis, R., Caminero, J.A. and Migliori, G.B. (2016) 'Classification of drugs to treat multidrug-resistant tuberculosis (MDR-TB): Evidence and perspectives', *Journal of Thoracic Disease*, *8*(10), pp. 2666–2671.

Rocha, D.M., Viveiros, M., Saraiva, M. and Osório, N.S. (2021) 'The Neglected Contribution of Streptomycin to the Tuberculosis Drug Resistance Problem', *Genes*, *12*(12), p.2003

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning, a laboratory manual (2nd Edition) Cold Spring Harbor Laboratory Press.

Schatz, A., Bugle, E. and Waksman, S.A.(1944) 'Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria', *Proceedings of the Society for Experimental Biology and Medicine*, *55*(1), pp.66-69.

Singh, V. and Chibale, K. (2021) 'Strategies to Combat Multi-Drug Resistance in Tuberculosis', *Accounts of Chemical Research*, *54*(10), pp. 2361–2376.

Stoll, A., Renz, J. and Brack, A. (1951) 'Beiträge zur Konstitutionsaufklärung des Nocadamins. 10. Mitteilung über antibakterielle Stoffe', *Helvetica Chimica Acta*, *34*(3), pp. 862–873.

Szybalski, W., and Iyer, V. N. (1964) 'Crosslinking of DNA by enzymatically or chemically activated mitomycins and porfiromycins, bifunctionally "alkylating" antibiotics', *Federation proceedings*, *23*, 946–957

Thompson, M.G., Corey, B.W., Si, Y., Craft, D.W. and Zurawski, D.V. (2012) 'Antibacterial Activities of Iron Chelators against Common Nosocomial Pathogens', *Antimicrobial agents and chemothera*py, *56*(10), pp.5419-5421.

Wang, Q., Song, F., Xiao, X., Huang, P., Li, L., Monte, A., Abdel-Mageed, W.M., Wang, J., Guo, H., He, W. and Xie, F .(2013) 'Abyssomicins from the South China Sea Deep-Sea Sediment *Verrucosispora* sp. natural thioether Michael addition adducts as antitubercular prodrugs', *Angewandte Chemie*, *125*(4), pp.1269-1272.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T. and Porto, C.(2016) 'Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking'. *Nature biotechnology*, *34*(8), pp.828-837

Watson D.J., Meyers P.R., Acquah K.S., Dziwornu G.A., Barnett C.B., Wiesner L. (2021) 'Discovery of novel cyclic ethers with synergistic antiplasmodial activity in combination with valinomycin', *Molecules*, *26*(24), pp. 1–19.

Westhoff, S., Kloosterman, A.M., van Hoesel, S.F., van Wezel, G.P. and Rozen, D.E. (2021) 'Competition sensing changes antibiotic production in *Streptomyces*', *mBio*, *12*(1), pp. 1–13.

World Health Organization (2021) 'Global tuberculosis report 2021', *World Health Organization, 2021*, Licence: CC BY-NC-SA 3.0 IGO, Geneva.

Yamanaka, K., Oikawa, H., Ogawa, H.O., Hosono, K., Shinmachi, F., Takano, H., Sakuda, S., Beppu, T. and Ueda, K. (2005) 'Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*', *Microbiology*, *151*(9), pp. 2899–2905.

Yepes-García, J., Caicedo-Montoya, C., Pinilla, L., Toro, L.F. and Ríos-Estepa, R. (2020) 'Morphological differentiation of *Streptomyces clavuligerus* exposed to diverse environmental conditions and its relationship with clavulanic acid biosynthesis', *Processes*, *8*(9), p.1038.

Zargar, H., Aning, J., Ischia, J., So, A. and Black, P. (2014). 'Optimizing intravesical mitomycin C therapy in non-muscle-invasive bladder cancer', *Nature Reviews Urology*, *11*(4), pp.220-230

Zong, G., Fu, J., Zhang, P., Zhang, W., Xu, Y., Cao, G. and Zhang, R. (2021) 'Use of elicitors to enhance or activate the antibiotic production in *Streptomyces*', *Critical Reviews in Biotechnology*, pp. 1–24.

CHAPTER 5:

GENERAL DISCUSSION AND CONCLUSION

Due to antibiotic resistance, there is an urgent need for novel antibacterial compounds. The most prolific antibiotic producers are members of the phylum *Actinobacteria* (Jose *et al.*, 2021). Consequently, exploring understudied habitats for the isolation of novel actinobacteria (with the ability to produce novel bioactive compounds) has become an active and important field of research. South Africa ranks as the third most biodiverse country in the world (https://www.biofin.org/south-africa), yet very little is known about the biodiversity in South African marine environments. This study made use of traditional culture-based methods, in combination with data obtained from metagenomic analysis, to investigate the actinobacterial biodiversity in marine sediment from Gericke's Point in Sedgefield, South Africa. The marine actinobacteria isolated in this study were screened for their ability to produce antibacterial compounds. There was particular emphasis on antimycobacterial compounds due to increasing antibiotic resistance in *Mtb* strains. South Africa is among WHO's list of 20 high-burden multidrug-resistant tuberculosis countries (World Health Organization, 2021).

Thirty-three (33) presumptive actinobacterial strains were isolated in this study by means of culture-based techniques, 27 of which were confirmed as actinobacterial strains. Twenty six (26) isolates were identified as *Streptomyces* strains and one isolate was identified as belonging to the genus *Nocardia*. The chances of isolating more actinobacteria, specifically strains of non-*Streptomyces* genera, would have been greater had it not been for the fungal overgrowth on the isolation plates due to the extended hard lockdown in South Africa in 2020. In the future, sediment samples could also be pre-treated to enhance the probability of isolating novel actinobacterial strains (Goodfellow and Fiedler, 2010). However, it is important to note that the metagenomic analysis revealed that the majority of actinobacteria present in the Gericke's Point sediment samples have yet to be cultured. Thus, these actinobacteria may require very specific culture conditions to be isolated.

All but one of the isolates were identified as *Streptomyces* strains indicating a skewed isolation pattern towards this genus. Furthermore, there was also a lack of diversity seen among the *Streptomyces* strains isolated in this study, as three separate groups were identified as potential clones with identical partial 16S-rRNA gene sequences. The groups were as follows: 1) strains GGUI10#19, GGUI5#24 and GGUM1#26 (PC1), 2) strains GGUI10#8, GGUI#9, GGUI#10, GUUM1#11, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33 (PC2) and 3) strains GGUI#1 and GGUM1#2 (PC3). It was noted that the lack of variation within the partial 16S-rRNA gene

sequences obtained for *Streptomyces* strains in this study could be due to the exclusion of the gamma region (nucleotide positions 158-203) of the 16S-rRNA gene. This region is important for species discrimination within the genus *Streptomyces*. Analysis of the partial *gyrB* gene sequences of potential clones revealed that the strains in PC1 and PC3 are closely related, but are not clones. Strains GGUI#10, GUUM1#11, GGUI10#15, GGUI#28 (PC2) had identical partial *gyrB* sequences, but significantly different phenotypic characteristics. Therefore, though it is unlikely these strains are clones, only obtaining their full genome sequences would allow one to definitively establish whether these strains belong to the same genomic species. The same can be said for strains GGUI#9 and GGUI#16 (PC2), as they had identical partial *gyrB* gene sequences to strains GGUI10#8, GGUI5#18, GGUI5#27, GGUI5#30, GGUI5#31 and GGUI5#33 (PC2) but were significantly different, phenotypically, to these strains and to each other.

Streptomyces strain GGUI10#6 had the lowest 16S-rRNA gene sequence similarity to its closest relative, *S. synnematoformans* S155^{T}, amongst all of the identified actinobacteria (98.91% 16S-rRNA gene sequence similarity over a length of 922 nucleotides). It was also noted that the phenotypic characteristics of strain GGUI10#6 observed in this study (dark brown/black substrate mycelium and white aerial mycelium) were slightly different to that observed for the type strain of *S. synnematoformans* (blackish red substrate mycelium and greyish red aerial mycelium; Hozzein and Goodfellow, 2007). Therefore, strain GGUI10#6 would be worth investigating further to determine whether it represents a new species. Currently, there is no genome sequence available for *S. synnematoformans* S155^{T}. This means neither ANI nor dDDH can be performed to compare these strains until both genomes are sequenced. However, there are partial sequences available for the *atpD*, *recA*, *rpoB*, and *trpB* housekeeping genes for *S. synnematoformans* DSM 41902^{T}. Therefore, a multilocus sequence analysis could be done to provide further evidence of the novelty of strain GGUI10#6 and to establish whether it is worth sequencing the genomes of this strain and the type strain of *S. synnematoformans* closest relative for further analysis.

Streptomyces strains GGUI10#19, GGUI5#24 and GGUM1#26 had a low *gyrB* gene sequence similarity to their closest relative, *S. chumphonensis* K1-2^{T} (93.52%, 94.31% and 94.08% *gyrB* gene sequence similarity, respectively, over a length of ~300 nucleotides). As there is a genome sequence for *S. chumphonensis* K1-2^{T}, the genomes of *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUM1#26 could be sequenced for whole-genome

comparisons. If the ANI values between the genomes of these strains and *S. chumphonensis* K1-2^{T} are below 96% and the dDDH values are below 70% they can be considered to represent a distinct genomic species (Auch *et al.*, 2010; Ciufo *et al.*, 2018). If this is the case, phenotypic characterisation can be carried out to fully classify *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUM1#26. Whole-genome comparison of these isolates will also determine whether they are strains of the same species or distinct species from each other.

Analysis of the partial 16S-rRNA and *gyrB* gene sequences of *Nocardia* strain GGUI10#20 revealed its closest relative to be the type strain of *N. otitidiscaviarum* (99.43% 16S- rRNA gene sequence similarity over 883 nucleotides; 99.52% *gyrB* gene sequence similarity over 414 nucleotides). *N. otitidiscaviarum* is a rare opportunistic pathogen that was first isolated in 1924 from the middle ear of a guinea pig and has since been isolated from multiple human clinical specimens. This species has also been isolated from crude oil contaminated seawater in India (Vyas and Dave, 2011). In this study, *Nocardia* strain GGUI10#20 was isolated from a sediment sample from a rock pool at Gericke's Point. Interestingly, a persistent contaminant was found among the strains isolated from this site that belonged to the genus *Alcanivorax*, which are marine bacteria also found in high abundances in oil-polluted environments (Gericke's Point is not an oil-polluted site). The results obtained from the culture-dependent investigation in this study are limited by the fact that only partial 16S-rRNA and *gyrB* gene sequences were used (in addition to the existing limitations associated with utilizing single-gene genetic analyses in bacterial taxonomy). Therefore, since a genome sequence is available for the type strain of *N. otitidiscaviarum*, whole-genome analyses should be performed to definitively establish whether *Nocardia* strain GGUI10#20 represents a new species in the genus *Nocardia*.

Data generated from amplicon-based NGS of eDNA (performed by the BTB research group) was used to investigate the actinobacterial composition of sediment samples in each sub-site (sub-tidal zone, beach and rock pool) at Gericke's Point. Since the metagenomics analyses were culture-independent, they were able to reveal the actinobacterial biodiversity of the Gericke's Point sediment that was not seen in the culture-dependent part of this study. The following orders were identified in all sub-sites using partial 16S-rRNA gene sequence reads: *Candidatus* Microtrichiales, *Candidatus* Actinomarinales, *Mycobacteriales*, *Frankiales*, *Geodermatophilales, Micrococcales*, *Micromonosporales* and *Pseudonocardiales.* Additional orders were found in the beach and rock pool sediment including *Euzebyales, Streptomycetales* and *Streptosporangiales*. Strains belonging to the order *Bifidobacteriales*

were found in the rock pool sediment only, indicating the rock pool sediment contains the most actinobacterial diversity among all the sediment samples investigated in this study (though not by far). The samples from the sub-tidal zone contained the lowest actinobacterial diversity. This could be because the ocean is a more complex environment that is only habitable for certain bacteria. Interestingly, even though both strains isolated from the ocean sub-site were identified as *Streptomyces*, no members of this genus were found in the sub-tidal zone sediment samples in the metagenomic analysis. It is possible that, although *Streptomyces* strains are present in this environment, their 16S-rRNA gene sequences were not amplified using the primers in this study. There is also a small chance that, if amplified, the region of the 16S-rRNA gene (~270 nucleotides) obtained for these strains was too distinct to be classified correctly using the SILVA database and the cutoff of 97% 16S-rRNA gene sequence similarity. The genus *Streptomyces* is well-described and members of this genus are easy to cultivate, as they grow very well on laboratory media (Seong *et al.*, 2001). This is suggested as an explanation for why, even though members of this genus made up an extremely small part of the actinobacterial diversity discovered in the Gericke's Point sediment samples, the majority of strains isolated in the culture-dependent investigation were identified as *Streptomyces*.

As mentioned previously, the majority of actinobacteria identified in the Gericke's Point sediment samples could not be assigned to a known genus (according to the SILVA 16S-rRNA database). This not only indicates the abundance of novel actinobacterial diversity yet to be revealed, but also highlights the limitations of reference-based taxonomic profiling. Despite the majority of actinobacteria identified in the samples from this marine environment being seemingly unculturable, the results showed that members of some very interesting (previously cultured) actinobacterial genera were also present. These include members of the genera *Salinispora*, *Euzebya*, *Marmoricola* and *Isoptericola*. All of these genera are considered rare as they contain only a few described species with validly published names and would be good targets for metagenomic-guided selective isolation. Additionally, members of the genus *Frankia* were found in the sediment samples obtained from the rock pool and subtidal zone at Gericke's Point. This is surprising as *Frankia* strains are predominantly terrestrial and, to date, have not been found in marine environments. Therefore, marine strains of this genus would be interesting to investigate and would also be good targets for selective isolation.

The screening of the actinobacterial isolates for antimycobacterial antibiotic activity showed that *Streptomyces* strains GGUI#1, GGUM1#2, GGUM1#4, GGUI10#6, GGUI#7, GGUI10#8, GGUI#9, GGUI#10, GGUM1#11, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI#23, GGUI5#27, GGUI#28, GGUM1#29, GGUI5#30, GGUI5#31, GGUI5#33 and unidentified strain GGUI5#21 displayed strong and very strong activity (ZOIs > 2000mm^2) against *M. aurum* strain A+, suggesting that they produce strong antimycobacterial compounds. Data from mass spectrometry analyses performed by Dr. Daniel Watson revealed that the antimycobacterial compounds produced (on at least one screening medium) by strains GGUI#1, GGUM1#4, GGUI10#6, GGUI#7, GGUI10#8, GGUI#9, GGUI#10, GGUI#12, GGUI10#15, GGUI#16, GGUI5#18, GGUI5#21, GGUI#23, GGUI5#27, GGUI#28, GGUI5#30, GGUI5#31 and GGUI5#33 had no structural match in the Global Natural Products Social Molecular Networking (GNPS) database. This indicates that almost 70% (18 of 26) of the *Streptomyces* strains isolated in this study can produce potentially novel antibiotics. This is not surprising, as the majority of novel secondary metabolites isolated from marine actinobacteria have been produced by marine *Streptomyces* strains (Manivasagan *et al.*, 2014). Notably, *Streptomyces* strain GGUI5#33 also produced a compound that could be an analogue of the rare antibiotic, abyssomycin L, which could also be worth investigating further. This compound was originally isolated from a marine actinobacterium from the genus *Verrucosispora* (now *Micromonospora*). All potentially novel antimycobacterial compounds found in this study are worth investigating further as candidates for development as new anti-TB drugs. Future work would include the confirmation of antimycobacterial activity by re-testing against *M. aurum* strain A+ using agar overlays, followed by the determination of antitubercular activity by testing against *Mtb* and then structural elucidation of promising antitubercular compounds.

In addition to *M. aurum* strain A+, *Streptomyces* strain GGUI10#6 also displayed very strong activity (on all screening media) against *S. aureus* strain ATCC 25923 (another Gram-positive bacterium). This is relevant as novel antibiotics are also needed to combat the increasing incidence of MRSA. *S. aureus* is an opportunistic pathogen commonly causing skin and respiratory infections. Since it was active against both Gram-positive bacteria included in this study, further screening against other test bacteria could be carried out to definitively establish the full spectrum of antibiotic activity of strain GGUI10#6.

Overall, the results obtained in this study emphasize the marine environment's potential as a source of novel actinobacteria and novel bioactive compounds. Furthermore, although emphasis has been put on the isolation of novel actinobacteria (i.e. new genera) in antibiotic discovery, this and previous studies have shown that novel strains of well-characterised actinobacterial genera that have been isolated from the marine environment can be a great source of novel bioactive compounds as well.

REFERENCES

Auch, A.F., von Jan, M., Klenk, H.P. and Göker, M. (2010) 'Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison', *Standards in genomic sciences*, *2*(1), pp. 117–134.

Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., Brover, S., Schoch, C. L., Kimchi, A., DiCuccio, M. (2018) 'Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI', *International Journal of Systematic and Evolutionary Microbiology*, *68***,** pp. 2386-2392.

Goodfellow, M. and Fiedler, H. P. (2010) 'A guide to successful bioprospecting: Informed by actinobacterial systematics', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *98*(2), pp. 119–142.

Hozzein, W. N. and Goodfellow, M. (2007) '*Streptomyces synnematoformans* sp. nov., a novel actinomycete isolated from a sand dune soil in Egypt', *International Journal of Systematic and Evolutionary Microbiology*, *57*(9), pp. 2009–2013.

Jose, P. A., Maharshi, A. and Jha, B. (2021) 'Actinobacteria in natural products research: Progress and prospects', *Microbiological Research*, *246*, p. 126708.

Manivasagan, P., Venkatesan, J., Sivakumar, K. and Kim, S. K. (2014) 'Pharmaceutically active secondary metabolites of marine actinobacteria', *Microbiological Research*, *169*(4), pp. 262–278.

Seong, C. N., Choi, J. H. and Baik, K. S. (2001) 'An Improved Selective Isolation of Rare Actinomycetes from Forest Soil', *The Journal of Microbiology*, *39*, pp. 17-23.

Vyas, T. K. and Dave, B. P. (2011) 'Production of biosurfactant by *Nocardia otitidiscaviarum* and its role in biodegradation of crude oil', *Int. J. Environ. Sci. Tech*, *8*(2), pp. 425–432.

World Health Organization (2021) 'WHO global lists of high burden countries for tuberculosis (TB), TB / HIV and TB (MDR / RR-TB)', pp. 2021–2025.

APPENDIX

Table A1. Morphological characteristics of 30 actinobacterial marine isolates. All strains were grown on ISP4 agar medium for 14 days at 30°C, except GGUI#22 whose characteristics were observed on ISP2 agar medium at room temperature (~22 °C). N/A, not applicable (no diffusible pigment produced).

Isolate	Colour of spore mass	Colour of substrate mycelium	Colour of diffusible pigment	Genus identification
OCEAN SITE				
GGUI#1	White	Light brown	Pale Pink	*Streptomyces*
GGUI10#6	White	Dark Brown/Black	Pale reddish brown	*Streptomyces*
DRY SITE				
GGUM1#2	Light brown	Brown	Brown	*Streptomyces*
GGUI#3	Cream white	Cream white	N/A	*Streptomyces*
GGUM1#4	White/Grey	Yellowish cream	N/A	*Streptomyces*
GGUI10#15	Cream/Grey	Brown	Brown	*Streptomyces*
GGUM1#29	Charcoal grey	Yellowish brown	N/A	*Streptomyces*
ROCKY SITE				
GGUM1#5	White	Yellowish cream	N/A	*Streptomyces*
GGUI#7	Dark grey/white	Yellow	Light Brown	*Streptomyces*
GGUI10#8	Cream /Grey	Brown	Light Brown	*Streptomyces*
GGUI#9	White	Brown	Dark Brown	*Streptomyces*
GGUI#10	Cream	Brown	Dark Brown	*Streptomyces*
GGUM1#11	Cream White	Brown	Brown	*Streptomyces*
GGUI#12	Cream/Grey	Brown	Brown	*Streptomyces*
GGUI#16	White	Light Brown	Light Brown	*Streptomyces*
GGUI5#18	Cream/Grey	Light Brown	Light Brown	*Streptomyces*
GGUI10#19	White	Cream white	N/A	*Streptomyces*
GGUI10#20	Pale pink	Pale orange	Yellowish brown	*Nocardia*
GGUI5#21	Yellow cream	Yellowish cream	N/A	Not determined
GGUI#22	White	Dark Brown	Brown	Not determined
GGUI#23	Light brown	Dark brown	Brown	*Streptomyces*
GGUI5#24	White	Brown	Dark Brown/Black	*Streptomyces*
GGUI5#25	Grey	Reddish brown	Dark Brown	*Streptomyces*
GGUM1#26	White	Yellowish cream	Dark Brown	*Streptomyces*
GGUI5#27	White/Grey	Light Brown	N/A	*Streptomyces*
GGUI#28	White	Orange	Light Brown	*Streptomyces*
GGUI5#30	White/Grey	Brown	Light Brown	*Streptomyces*
GGUI5#31	White/Grey	Brown	Light Brown	*Streptomyces*
GGUI5#32	White	Cream white	N/A	Not determined
GGUI5#33	Cream/Grey	Yellowish brown	Light brown	*Streptomyces*

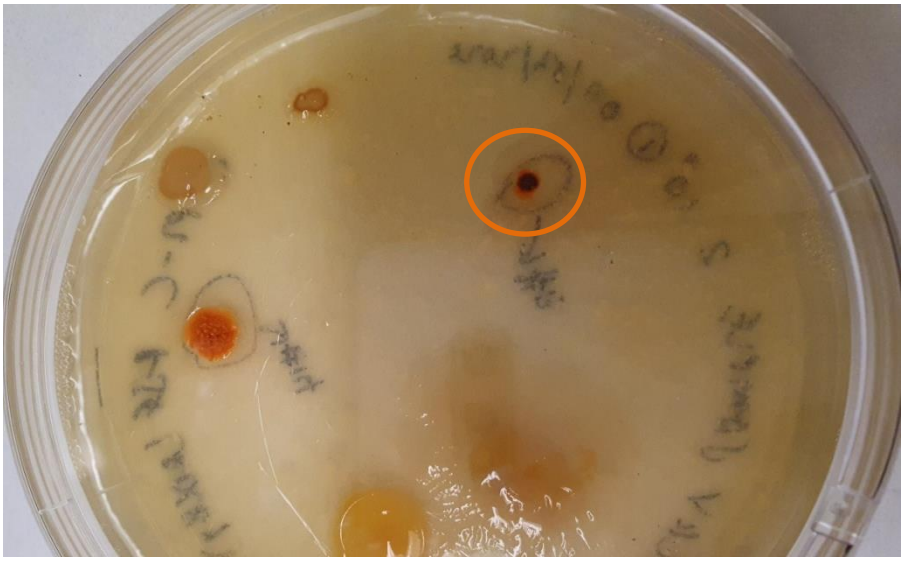

Figure A1. Isolation plate containing isolate GGUI#13 (circled in orange). Colony is orange with black spore mass (typical of the genera *Micromonospora* and *Salinispora).* This strain was, unfortunately, lost as a result of persistent fungal contamination.

Table A2. Details on all contaminants encountered in this study. Identification was based on EzBioCloud analyses of partial 16S-rRNA gene sequences. Sequence length indicates the number of base pairs (bp) in the query sequence. Percentage (%) identity indicates the degree of similarity between the query sequence and the top hit strain gene sequence.

Contaminant identification	% identity	Sequence length (bp)	Actinobacterial isolate affected	Notes on contaminant
Gram-negative				
Pseudomonas moraviensis	99.89	918	GGUI5#21	Widespread occurrence in terrestrial and aquatic habitats. Non-spore forming.
	99.89	945	GGUI#23	
	99.89	911	GGUM1#26	
	99.89	938	GGUI#28	
	99.89	931	GGUI5#33	
***Alcanivorax* species**	97.55	449	GGUI5#21	Marine bacterium found in high abundances in oil-polluted environments. Non-spore forming.
	99.54	875	GGUI#22	
	97.67	258	GGUI5#32	
Pantoea septica	99.33	898	GGUM1#4	Yellow-pigmented. Non-spore forming
Delftia tsuruhatensis	99.63	821	GGUI10#19	Prefers environments with minimal to no salt for growth. Non-spore forming.
	100	816	GGUI5#31	
Gram-positive				
Staphylococcus saprophyticus* subsp. *saprophyticus	100	919	GGUI#32	Non-spore forming.

Figure A2. Morphology of selected *Streptomyces* strains GGUI10#19, GGUI5#24 and GGUM1#26 on ISP4 agar medium after 14 days of incubation at 30 °C

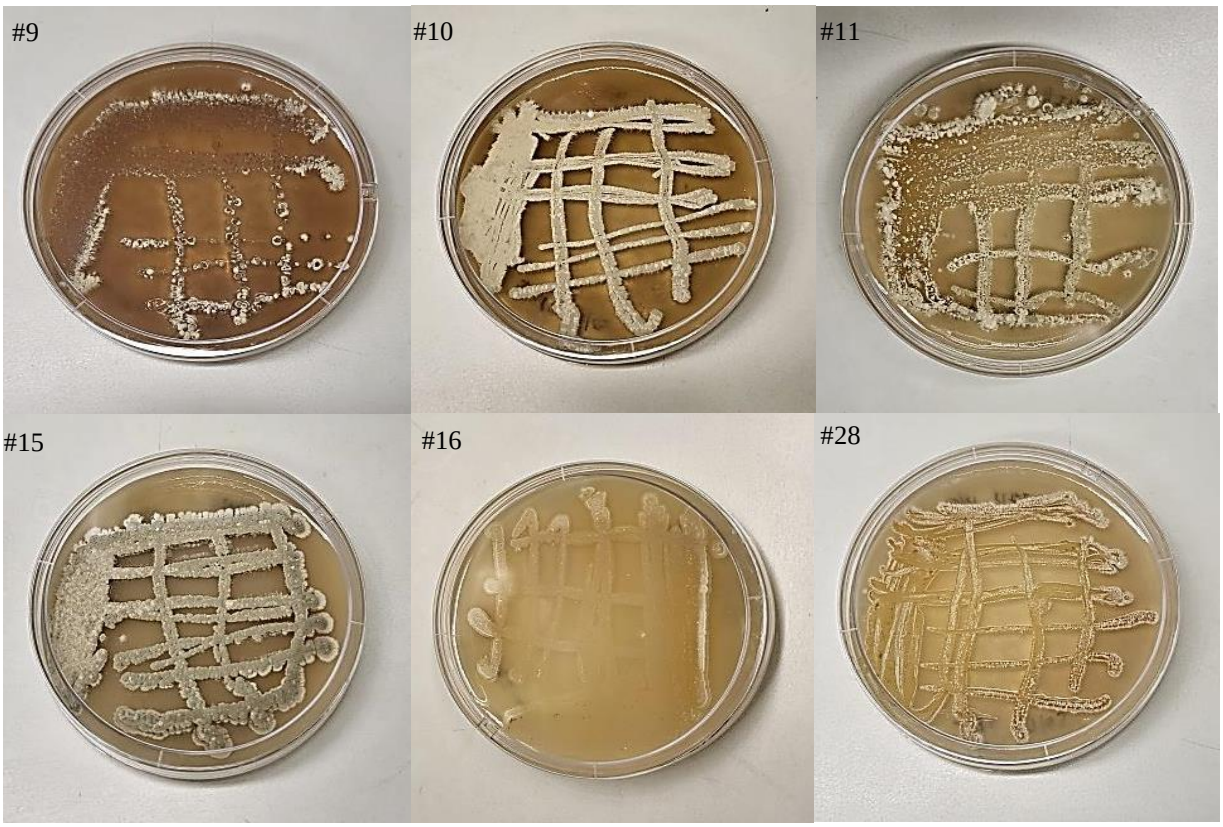

Figure A3. Morphology of selected *Streptomyces* strains GGUI#9, GGUI#10, GGUM1#11, GGUI10#15, GGUI#16 and GGUI#28 on ISP4 agar medium after 14 days of incubation at 30 °C.