Screening of marine bacteria as a source of bioactive secondary metabolites

By

Jatro Kulani Mhlongo

- A thesis submitted in partial fulfilment of the requirements for the
- degree of Doctor of Philosophy (PhD)

- Institute for Microbial Biotechnology and Metagenomics, Department
- of Biotechnology, Faculty of Natural Science
- University of the Western Cape,

Supervisor: Prof Marla Trindade

Co-supervisor: Prof Denzil Beukes

Date: 7-10-2020

Declaration

I, Jatro Kulani Mhlongo, declare that "Screening of marine bacteria as a source of bioactive secondary metabolites" is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Date: 7-10-2020

Signature:……..

Acknowledgements

I would like to thank my supervisor Prof Marla Trindade for her leadership, guidance and for providing a pleasant working space for the duration of the journey.

I would like to thank my supervisor co-supervisor Prof Denzil Beukes for his advice and guidance.

Dr Lonnie van Zyl for the technical support and Dr Bronwyn Kirby-McCullough for assistance with whole genome sequencing.

Thanks to Robin and Leo for their contributions and support.

The National Research Foundation (NRF), Oceanmedicine and Pharmasea for funding this work

Thanks to Prof Marcel Jaspars, Dr Jioji Tabudravu and Dr Rainer Ebel for allowing me in their laboratory (Marine Biodiscovery Centre, University of Aberdeen) and training on different natural product purification systems, as well as characterisation using techniques such as nuclear magnetic resonance spectroscopy, high performance liquid chromatography (HPLC) and mass spectrometry (MS).

Thanks to all the IMBM members for making this experience enjoyable and endurable.

Lastly, I like to thank my friends and family for words of encouragement, support and believing in me.

Thank You

Abstract

Marine invertebrate associated (MIA) bacteria are an important source of bioactive secondary metabolites with the potential to address the current anti-microbial resistance crisis experienced globally. Secondary metabolites (SM) have historically yielded several compounds with pharmaceutical applications such as anti-viral, anti-microbial, anti-cancer, anti-inflammatory and anti-parasitic. This study aimed to use bioassay and genomic approaches in the identification of MIA bacteria isolated from South African marine invertebrates as a source of bioactive compounds and the characterisation of the produced SMs using analytical techniques. A total of 23 MIA bacteria were cultured under different conditions (one strain many compounds approach (OSMAC)) to evaluate their ability to produce anti-microbial compounds against a panel of indicator strains namely *Escherichia coli* 1699, *Bacillus cereus* ATCC10702, *Pseudomonas putida* ATCC12633, *Mycobacterium aurum* A^{+}, *Staphylococcus epidermidis* ATCC14990, *Aspergillus fumigatus* MRC and *Candida albicans* NIOH. The isolates were also evaluated for their ability to confer anti-inflammatory activity. Anti-microbial activity was detected in all the studied isolates with 48 % showing anti-microbial activity against *E. coli* 1699 which is genetically engineered to be resistant to over 50 commercially available antibiotics. The majority of isolates showing bioactivity indicate that the South African marine environment represents an ecological environment niche that harbours uncharacterised metabolites.

Three MIA bacteria (PE08-149B, PE14-07 and PE14-63) with desired anti-microbial profiles were selected for whole genome sequencing. Genome sizes of ~5.9 Mbps, ~5.8 Mbps and ~5.7 Mbps for PE08-149B, PE14-07 and PE14-63, respectively, were observed. The genome data was further subjected to bioinformatics tools such as Bagel, antiSMASH and PRISM for detection of SM biosynthesis gene clusters (BGCs). Multiple gene clusters were identified per strain with a minimum of 5 BGCs. The detected gene clusters included non-ribosomal peptides (NRPs), and ribosomally synthesized and post-translationally modified peptides (RiPPs), siderophores and terpenes. RiPPs were the abundant BGCs across all strains followed by the siderophore and betalactone BGCs.

A total of three compounds were identified and isolated from PE14-07 fermentation broth using High-pressure liquid chromatography (HPLC) and characterized with nuclear magnetic resonance

(NMR) spectrometry and high-resolution liquid chromatography-mass spectrometry (HRLC-MS). Extraction from PE08-149B resulted in the isolation and identification of two compounds active against *P. putida* using the above methods in combination with bioassay-guided isolation.

The identification of potential new SMs using bioinformatics tools, HRLC-MS, bioassays and HPLC indicates that the South African coastline is a rich source of SMs with the potential to contribute towards mitigating the anti-microbial resistance crisis. Furthermore, the identification and isolation of SMs with anti-microbial activity against some of the indicator strains from PE14-07 and PE08-149B through OSMAC strategy indicates bioprospective potential. The identification of uncharacterised BGCs in the isolates further indicates their importance in drug discovery. This also indicates the importance of combining the strategies for drug discovery and source prioritization.

Keywords: Marine environment, marine invertebrate, MIA bacteria, anti-microbial, genome, bioassay, secondary metabolites, NMR, LC-MS, antiSMASH, Bagel 3 and PRISM

Table of content

List of Tables

List of Figures

List of Abbreviations

16s rRNA - 16S ribosomal Ribose nucleic acid gene

1D - One-dimensional

2D - Two-dimensional

ACM -Activated charcoal medium

ACP-domain - Acyl carrier protein domain

ACP-domain - Acyl carrier protein domain

A-domain - Adenylation domain

AMT-domain - Aminotransferase domain

AntiSMASH - Antibiotics and Secondary Metabolite Analysis Shell

APE - aryl polyene

AT-domain - Acyl-transferase domain

BLAST - Basic local alignment and search tool

BLASTn - Basic local alignment and search tool for nucleotide

BLASTp - Basic local alignment and search tool for protein

bp - Base pairs

CAF - Central analytical facility

CC - Column chromatography

C-domain - Condensation domain

CI - Chloroform isoamyl alcohol

Contig - Contiguous

COSY - Correlation spectroscopy

CTAB - Cetyl trimethylammonium bromide

CV - column volume

DCM - Dichloromethane

DEREP-NP – Dereplication of natural product

DH – Dehydratase

DH_20 - Distilled water

DMSO - Dimethyl sulfoxide

DNA - Deoxyribose nucleic acid

E - Epimerization

ER – Enoylreductase

EtOAc - ethyl acetate

FDA - Food and drug administration

g - Earth's gravitational force

gDNA - genomic deoxyribose nucleic acid

GNP – Genomes to natural products platform

GYM – Glucose, yeast extract and malt extract media

h - hour

HCl – Hydrochloric acid

HGT - horizontal gent transfer

HIV - Human immunodeficiency virus

HMBC - Heteronuclear multiple-bond correlation

HPLC - High-pressure liquid chromatography

HRLC-MS – High-resolution liquid chromatography-mass spectrometry

Hserlactone - Homoserine lactone

HSQC - Heteronuclear single quantum coherence

IR – Infrared spectroscopy

Kb - Kilobase pairs

KR - Ketoreductase

KS - Ketosynthase

LA - Lysogeny agar

LB - Lysogeny broth

LC - Liquid chromatography

LC-MS – Liquid chromatography-mass spectrometry

Mb - Megabase pairs

MeOH - Methanol

MIA - Marine invertebrate-associated

min - minutes

MS - Mass spectrometry

NaCl - Sodium chloride

NaOH - sodium hydroxide

NaOH – Sodium hydroxide

NaPDos - Natural Product Domain seeker

NCBI - National Center for Biotechnology Information

NMR- Nuclear magnetic resonance

NP - Natural product

NRP - Non-ribosomal peptide

NRPS – Non-ribosomal peptide synthase

OD - Optical density

OSMAC - One Strain Many Compounds

PCA -Principal component analysis

PCI - Phenol / chloroform / isoamyl alcohol

PCP-domain - Peptidyl carrier domain

PDA – Potato dextrose aga

PDB - Potato dextrose broth

PK - Polyketide

PKS - Polyketide synthase

PRISM - Prediction Informatics for Secondary Metabolomes

PVET-SPE-X - Pseudovibrio ethyl acetate- solid-phase extraction -X%

RiPP - Ribosomally synthesized and post-translationally modified peptides

RNA - Ribose nucleic acid

RO - Reverse osmosis

RP - Reverse phase

RPM - Revolutions per minute

s - seconds

SA - South Africa

SDS - Sodium dodecyl sulfate

SMs - Secondary metabolites

SPE - Solid-phase extraction

TB - Tuberculosis

TDA - Tropodithietic acid

TE - Thioesterase

TEA - Tris (2,3-dibromopropyl) phosphate-acetate- Ethylenediaminetetraacetic acid

TFA- Trifluoroacetic acid

TPSA - Topological Polar Surface Area

T-RFLP - Terminal restriction fragment length polymorphism

TSA - Trypticase soy broth

TSB – Trypticase soy broth

UV - Ultraviolet

WGS - Whole genome sequencing

ZBA – Zobell agar

ZBB – Zobell broth

1. Chapter one: Literature review

1.1 Drug discovery

Drug discovery is a process whereby a compound is identified from a wide range of sources to assist in the treatment of a new or an existing medical condition that does not have suitable treatment as graphically illustrated in Figure 1.1 (Hughes et al., 2011). The process can take over 15 years for one drug to be available on the market (Ciani and Jommi, 2014). Secondary metabolites (SMs) are major contributors of compounds in drug discovery, followed by synthetic compounds. This is evident since ~70 % of the current drugs on the market are linked to SMs with the number promising to rise, while ~30 % are purely synthetic (Molinari, 2009; Newman and Cragg, 2016, 2012; Zivanovic, 2012).

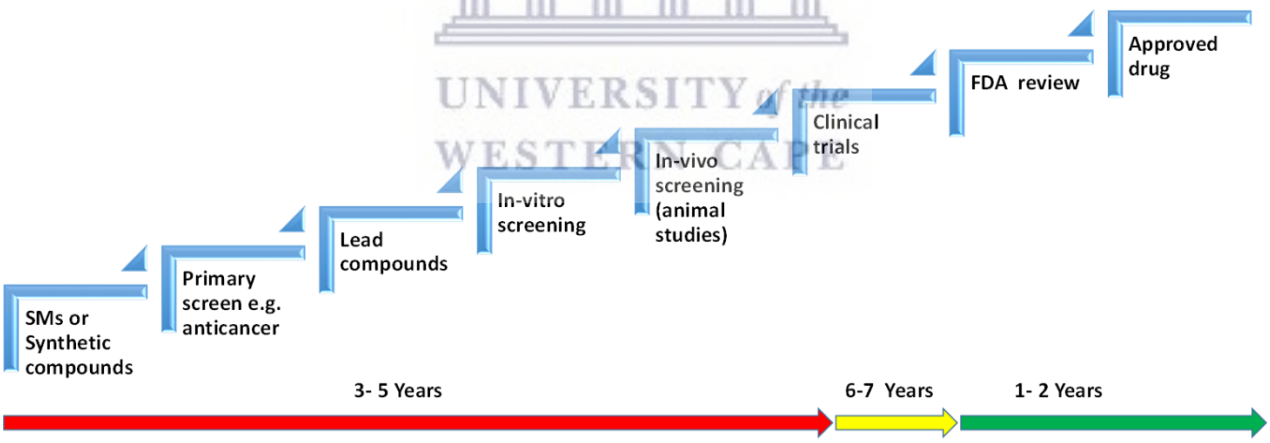


Figure 1.1. Graphical representation of the drug discovery process with timelines estimations (Hughes et al., 2011; Rocha-Roa et al., 2018).

The discovery of unique SMs involves the extraction, characterisation and isolation of new compounds from a variety of natural sources (Figure 1.2) such as plants, animals, fungi, marine invertebrates and bacteria (Ojima, 2008). However, SM discovery is facing a few challenges with re-discovery of known compounds being the main one. Other reasons include challenges in

structure elucidation, purification of SMs from crude extracts and providing enough quantity of the SM to support early-stage drug development (Phillips et al., 2013). These have led pharmaceutical companies to lose interest in continuing SM discovery efforts, resulting in the withdrawal of funding.

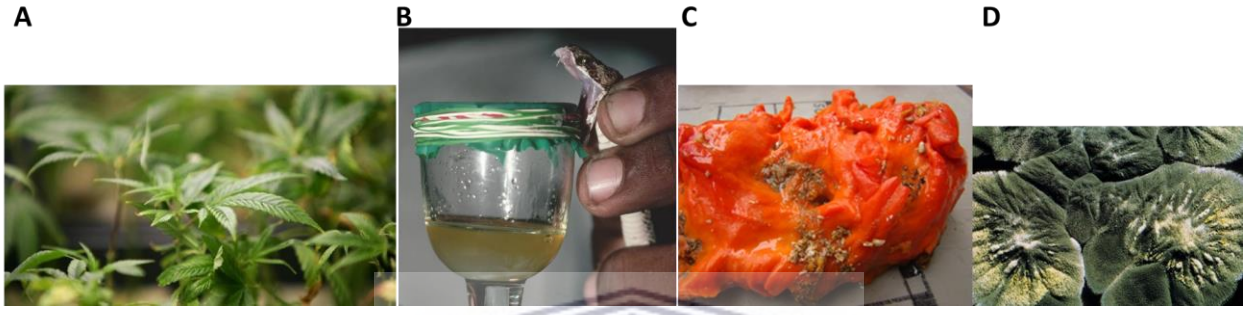


Figure 1.2. Examples of secondary metabolite sources. A: plant: *Cannabis sativa* (Oyon, 2018), B: snake: *Bothrops jararaca* (Nightingale, 2019), C: marine invertebrate: *Hamacantha* (*Vomerula*) *esperioides* and D: microorganism: *Penicillium notatum* (The Editors of Encyclopaedia Britannica, 2019)

In response, pharmaceutical companies have tried to develop alternative methods to aid in drug discovery such as combinatorial and synthetic chemistry; however, SMs are still found to be irreplaceable and unique (Phillips et al., 2013). This is because nature produces compounds with unmatched chemical diversity compared with any laboratory process developed thus far and are biochemically more target-specific (Bertrand et al., 2014; Zivanovic, 2012). This is despite chemical techniques being cheaper and faster compared to biological synthesis methods (Machado et al., 2015).

Research has also found that over 50 % of the SM or SM-derived drugs are not in violation of the Lipinski rules (i.e., an octanol-water partition coefficient logP, not more than 5, a molecule with a molecular mass not more than 500 Da, a maximum of 5 hydrogen bond donors, and a maximum of 10 hydrogen bond acceptors in the molecule); which is mostly used in the evaluation of "drug-like" compounds (Goodwin et al., 2017; Newman and Cragg, 2012). Nature, therefore, continues to be the greatest source of new compounds with a new mode of action that can be used as potential drugs in the treatment of emerging and re-emerging diseases (Newman and Cragg, 2012; Schinke

et al., 2017). Additionally, SMs are not only important for the pharmaceutical industry but also in industries developing cosmetics, nutritional supplements, molecular probes, biocatalysis, and agrichemicals.

1.2 Why do we need new drugs?

Due to the high rate of antibiotic administration and incorrect use of antibiotics, pathogenic microorganisms have developed and continue to develop resistance to currently available drugs. This together with the lack of cure for certain illnesses and the search for better drugs with better efficacy and minimal side effects has intensified the global search for new drugs (Schinke et al., 2017). For instance, microorganisms develop resistance to antibiotics easily as was observed in the case of penicillin whereby, 14 % of *Staphylococcus* species were already resistant shortly after its introduction into the market (Finking and Marahiel, 2004). Other factors that contribute to antibiotic resistance include poor sanitation, improper infection and disease prevention or control measures in the health care and farms, lack of affordable health care services (i.e., diagnosis, medicine and vaccines) and lack of legislation or enforcement(WHO, 2020a).

The above mentioned events lead to pathogenic microorganisms using resistance mechanisms such as minimising the uptake of the drug, modifying the binding target of the drug, enzymatic breakdown of the drug and active efflux pumps (Reygaert, 2018). Gram negative bacteria use all the above mechanisms while Gram positive bacteria mainly rely on minimising drug intake into the cell (Chancey et al., 2012). Bacterial drug resistance can be natural or acquired at some stage in the life cycle. Natural resistance is independent of previous exposure to the antibiotic while acquired resistance refers to the acquisition of genetic material encoding resistance to certain antibiotics through horizontal gene transfer or gene mutation (Martinez, 2014; Reygaert, 2018).

The WHO has classified anti-microbial resistance as one of the 10 critical threats to public health. In 2019 the WHO identified only 32 antibiotics in preclinical trials aiming to address the priority pathogens (i.e., *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, Enterobacteriaceae,

Enterococcus faecium, *Staphylococcus aureus*, *Helicobacter pylori*, *Campylobacter* sp., *Salmonellae*, *Neisseria gonorrhoeae*, *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Shigella* sp.). Commonly used antibiotics are rendered useless against these pathogens (WHO, 2017, 2020a).

Furthermore, the South African (SAn) health department reported the isolation of *Klebsiella pneumonia* in 2011 resistant to all documented antibiotics making it the most extreme case in SA history (Mendelson, n.d.). *Klebsiella pneumonia* together with other ESKAPE (*Enterococcus faecalis*, *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter cloacae*, and *Escherichia coli*) pathogens are responsible for nosocomial infections around the world and have developed resistance to most antibiotics used for their treatment (Benkő et al., 2020; Santajit and Indrawattana, 2016).

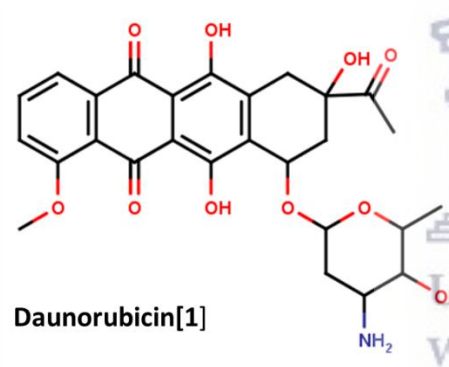
This rapid and extreme development of drug resistance means there exists a need for constant replenishment of new drugs to combat various diseases in an effort to manage public health, hence, the continued search for new chemicals leads from nature. Although microbial resistance to existing antibiotics is the main reason for the hunt for new SMs, SMs also serve as lead compounds in the development of drugs for the treatment of other infectious diseases, inflammation, immune suppressors for use in organ transplants, immunological disorders and perhaps in the fight against cancer (Ojima, 2008; Zhao, 2011).

1.3 Secondary metabolites

Secondary metabolites refer to any chemical compounds produced by living organisms which are not essential for the development of the organism (e.g., growth) and are also termed natural products or specialised metabolites. The role of these compounds in nature is not yet fully understood; however, there is evidence suggesting their use as signalling or communication molecules (N-acyl homoserine lactones), serving as scavengers for nutrient ions (siderophores), protection against infections and predation of the producing organism or their host (proteases,

hydrogen cyanide, polyketide/peptide/terpenoid compounds) (Clavico et al., 2006; Crawford and Clardy, 2011; Firn and Jones, 2000; Pawlik et al., 2002).

For example, daunorubicin **[1]** and other DNA-intercalating SMs have been shown to prevent bacteriophage infection in *Streptomyces*, *E. coli* and *Pseudomonas aeruginosa* by inhibiting bacteriophage replication, suggesting that these classes of SMs are protecting the bacteria from infection (Kronheim et al., 2018). It is worth noting that most of the compounds inhibiting bacteriophage infection possess anti-cancer properties. This character can be used as an assay to screen for anti-cancer SM in early development laboratories whereby there are no tissue culture facilities.



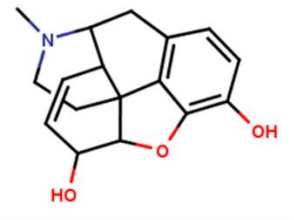
Daunorubicin[1]

The use of SMs as medicines dates back to ancient times with countries such as Egypt, South America, China, Rome, Iraq, Greece, and India, which are widely known for this practice (Phillipson, 2001). One of the earliest recorded instances of SMs being used for medicinal purposes originates from ancient Mesopotamia, circa 2600 BC which was recorded on clay tablets. The recording contained about 1000 plant extracts used for different illnesses, with many of them still being used today (Ji et al., 2009). For example, the *Cedrus* species and *Cupressus sempervirens* L are used to treat cough / flu and *Glycyrrhiza glabra* L is used in the treatment of diseases such as hyperdipsia, respiratory disorders, epilepsy, paralysis, sexual debility, fever, rheumatism, stomach ulcers, haemorrhagic diseases, skin diseases, and jaundice (El-Saber Batiha et al., 2020; Levatin, n.d.; Newman et al., 2000; Orhan and Tumen, 2015).

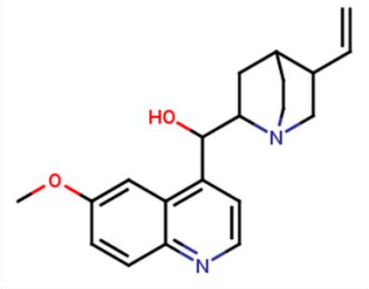
This formed the foundation of the pharmaceutical industry. Approximately 60 % of drugs used in the treatment of cancer (e.g., paclitaxel, atazanavir, trabectedin, ledipasvir and masoprocol) can be traced back to their natural origin, indicating that nature is a great source of medicines. Besides, according to the world health organisation (WHO), two-thirds of the human population consumes plant-derived traditional medicine for healthcare purposes (Jantan et al., 2015; Newman and Cragg, 2016).

Micro- and macro-organisms are regarded as bio-factories for the synthesis of primary and secondary metabolites with diverse chemical structures enabling them to have a wide range of functions (Gonzalez et al., 2003). Due to their diverse chemical structures and biological functions, SMs are of importance to biologists and chemists as they can be used for several applications such as biocatalysis and pesticides, in addition to drug development (Harvey, 2008; Zhao, 2011).

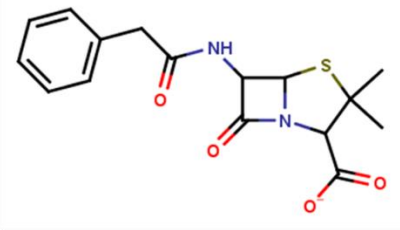
Morphine **[2]**, quinine **[3]** and penicillin **[4]** were some of the first SMs to be commercially available as medicinal drugs. Morphine was isolated in 1805 from a plant called *Papaver somniferum* and was the first naturally derived product (drug) to be available commercially (Laux-Biehlmann et al., 2013). Quinine, an anti-malaria agent was isolated from *Cinchona* bark in 1820 by Caventou and Pelletier; however, its use dates back to early 1600 and was previously referred to as "Jesuits" bark (Achan et al., 2011). Penicillin was discovered by Alexander Fleming in 1928 from *Penicillium notatum* and used as an antibiotic against Gram-positive bacteria. The discovery of penicillin led to microorganisms being recognised as an attractive source of new therapeutic agents for both human and animal diseases, while plants were the main source prior to this discovery (Challis, 2008; Ramachandran et al., 2012).



Morphine[2]

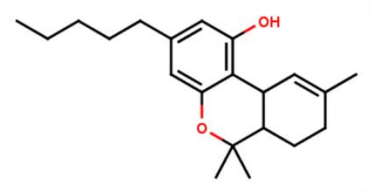


Quinine [3]

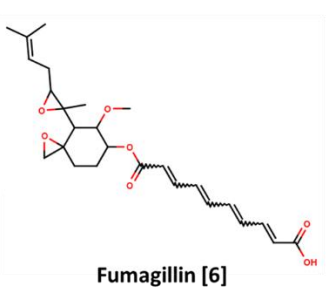


Penicillin [4]

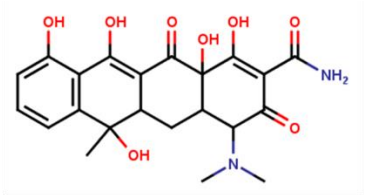
Secondary metabolites and their derivatives constituted up to ~27 % of the compounds approved by the United States Food and Drug Administration (FDA) between 1981 and 2006. Of this, 68 % and 34 % of these compounds have antibiotic and anti-cancer properties, respectively, with some compounds having multiple bioactivities (Newman and Cragg, 2007; Phillips et al., 2013). In the period between 1981 - 2019, of the 1394 FDA approved small molecule drug entities, 441 (32 %) were of natural origin, with 71 being unmodified NPs. Additionally, out of the 185 small molecules approved anti-cancer drugs during this period 62 molecules were of natural origin (Newman and Cragg, 2020). A number of SMs are currently on the market such as dronabinol **[5]** (pain treatment), fumagillin **[6]** (anti-parasitic), tetracycline **[7]** (anti-bacterial), pleuromutilin **[8]** (anti-bacterial), romidepsin **[9]** (oncology) and vancomycin **[10]** (anti-bacterial) for application in different therapeutic areas (Mishra and Tiwari, 2011).



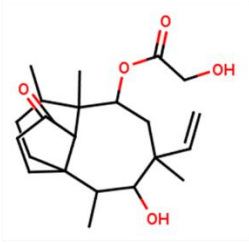
Dronabinol [5]



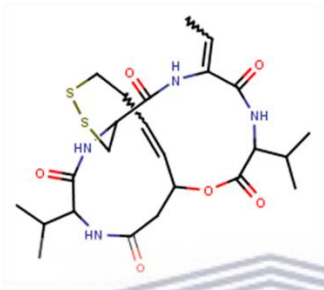
Fumagillin [6]



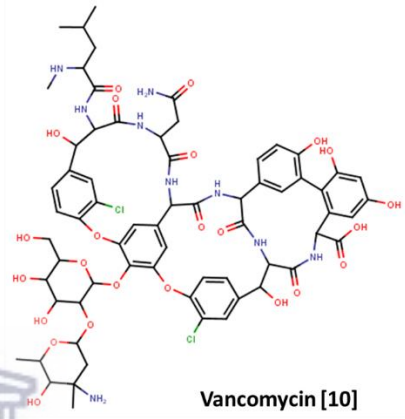
Tetracycline [7]



Pleuromutilin [8]



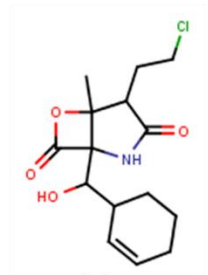
Romidespin [9]



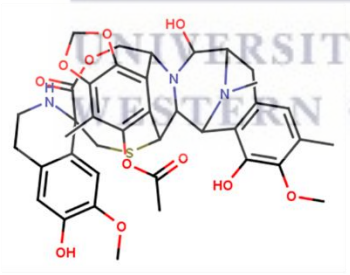
Vancomycin [10]

Despite this success, as mentioned before targeting SMs in drug discovery is facing a challenge in that many compounds are re-discovered, especially from terrestrial environments. Hence, researchers have turned to under-explored extreme environments such as the marine environment as a potential source of new SMs (Malve, 2016). Marine organisms are metabolically different from the terrestrial strains, and therefore, they produce unique compounds not found in terrestrial organisms (Mondol et al., 2013). For instance, halogenation of the SM is rare in terrestrial compounds compared to the marine environment (Jensen and Fenical, 1994). Salinosporamide A **[11]** is a well-known marine halogenated SM produced by the bacterium *Salinospora* sp., and the bacterium is only found in the marine environment. It possesses potent activity against HCT-116 human colon cancer cells and is currently in phase II clinical trials (Feling et al., 2003; Pérez and Fenical, 2017).

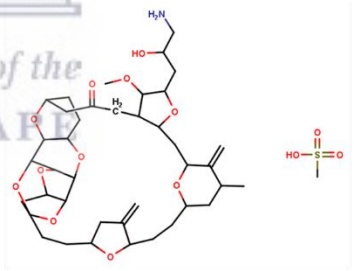
The marine environment has yielded several compounds which have been approved by the United States FDA or European Medicines Agency (EMA). Some of these include; trabectedin **[12]** and eribulin mesylate **[13]** for treatment of tumours; and iota-carrageenan **[14]** for viral infections (Martins et al., 2014). In the last 15 years over 167 marine SMs have shown activity against drug-resistant pathogens such as vancomycin-resistant enterococci, penicillin-resistant *Streptococcus pneumonia*, MRSA, vancomycin-resistant *S. aureus* and meticillin-resistant *Staphylococcus epidermidis*. For example, an anthracimycin **[15]** from a Marine *Streptomyces* sp. CNH365 showed activity against MRSA and vancomycin-resistant *S. aureus* while a cyclic peptide (PM181104) **[16]** from a sponge-associated actinobacterium *Kocuria* sp. MTCC 5269 showed activity against MRSA, vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* (Liu et al., 2019). These show that the marine environment is home to unique and new bioactive compounds with the potential to address anti-microbial resistance. These compounds are isolated from different marine sources such as algae, bacteria, fungi, sponges, seaweeds, molluscs and corals (Ghareeb et al., 2020).



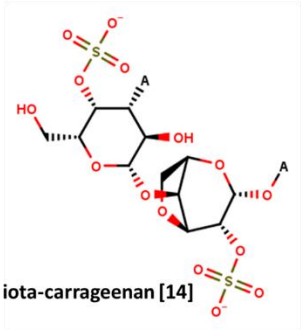
Salinosporamide A [11]



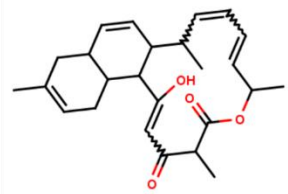
Trabectedin [12]



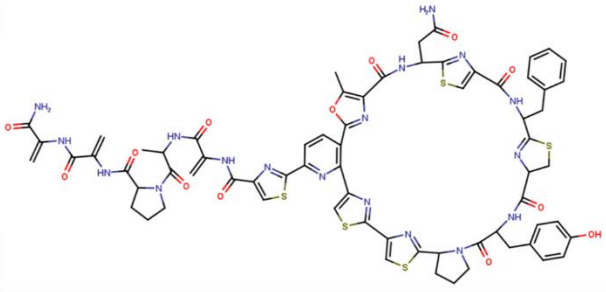
Eribulin mesylate [13]



iota-carrageenan [14]



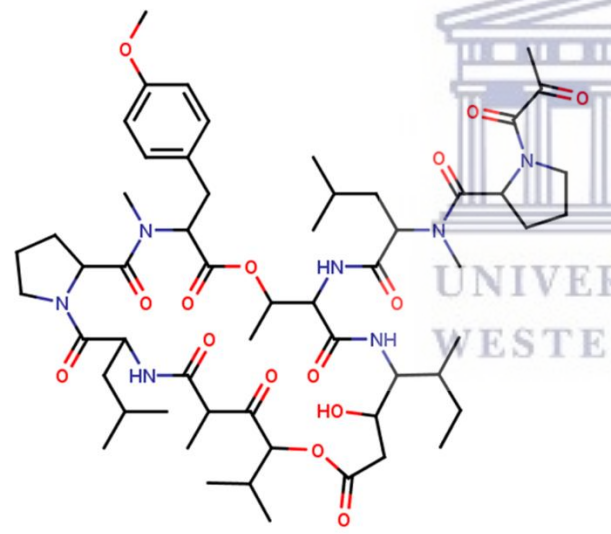
Anthracimycin [15]



cyclic peptide (PM181104)[16]

1.3.1 Secondary metabolite mode of action

Secondary metabolites display bioactivity by targeting either a protein, cell membrane or nucleic acid and result either in cell death or growth inhibition (Wink, 2015). In terms of antibiotic bioactivity, the SMs can be further subclassified as bacteriostatic or bactericidal, with the former meaning that the compound functions by arresting growth (prevent cell multiplication) while the latter entails killing the cell or organism (Nemeth et al., 2015). The compound can be specific to one target (e.g. one protein or cell membrane) or it can have multiple targets (Wink, 2015). For example, plitidepsin **[17]** is an anti-cancer compound that functions by blocking the cell cycle to induce the cell cycle. Table 1.1 contains more examples of compounds and their mode of action.



Plitidepsin 17

New targets or modes of action are vital in addressing anti-microbial resistance. Unfortunately, target screening is proving challenging as the assay may result in positive hits which are not active against the whole cell assay. This phenomenon was observed by GlaxoSmithKline (GSK) in a quest to discover anti-microbial compounds whereby 70 samples were target screened; however, none of the positive hits showed anti-bacterial activity when tested against bacteria (Payne et al.,

2007). This could be because it is not possible to determine parameters such as off-targets (e.g. different kinase not involved in cancer), membrane permeability, enzymatic breakdown and non-physiological properties of the compounds (Newman, 2017; Padmalayam, 2016). Hence, the need to lean towards phenotype-based screening which also enables possible discovery of new targets and un-hypothesised mode of action (Wagner, 2016), despite target screening having a quicker turnaround time (Newman, 2017).

However, due to dangers presented by and the availability of facilities handling pathogenic microorganisms, the initial screening is usually done with less pathogenic microorganisms. The microorganisms used should have similarities to the target pathogenic microorganism. For example, in our study *M. aurum* is used to represent *Mycobacterium tuberculosis* which causes tuberculosis (TB). *M. aurum* has shown similar antibiotic sensitivity to *M. tuberculosis* and it is fast growing, making it a good candidate / surrogate to identify compounds with suitable anti-tuberculosis properties (Chung et al., 1995; Gupta et al., 2009). Furthermore, genome analysis showed that the *M. aurum* genome contains orthologous genes which are responsible for drug resistance within *M. tuberculosis* (Phelan et al., 2015).

Table 1. 1. Example of SMs (drugs) of marine origin and their mode of action including natural sources

Secondary metabolite	Source	Disease/application	Mode of action	Reference
Plitidepsin	*Aplidium albicans.*	Anti-cancer	blocking the cell cycle	(Ghareeb et al., 2020; Jaspars et al., 2016)
Tetrodotoxin	*Hapalochlaena maculosa*	Pain killer	Blocks sodium channels.	(Ghareeb et al., 2020)
Iota-carrageenan	*Rhodophyceae seaweeds*	Anti-viral	Creates a preventative nasal spray	(Ghareeb et al., 2020)
cytosine arabinoside	*Tectitethya crypta*	Anti-cancer	Suppresses of DNA polymerase and DNA synthesis	(Ghareeb et al., 2020)
ω-conotoxin MVIIA /	*Conus magnus*	Severe and chronic pain	N-type voltage-gated calcium channels	(Adams and Berecki, 2013, p.)
Alkaloid ET-743	*Ecteinascidia turbinata*	Anti-tumour	DNA modification	(Erba et al., 2001)
Salinosporamide A	*Salinospora tropica*	Anti-cancer	20S protease inhibitor	(Ghareeb et al., 2020; Gulder and Moore, 2010; Newman and Cragg, 2014)
Bryostatin	*Candidatus Endobugula sertula*	Anti-cancer	modulation protein kinase Cs (PKCs)	(Sun and Alkon, 2006)

1.4 The marine environment

The marine environment covers more than 70 % of the earth's surface and harbours a huge number of diverse micro- and macro-organisms (Zhao, 2011). A low number of marine plants, fungi, bacteria and animals have yielded approximately 12000 new chemical entities according to Sabdono and Radjasa (2011). The increase from 20 compounds in 1984 to approximately 1000 compounds yearly by 2014 from marine environments, totalling about 20000 compounds in the past 50 years (Blunt et al., 2014), clearly demonstrates the ocean's potential as a source of bioactive compounds (Blunt et al., 2014; Choudhary et al., 2017).

The oceans have remained underexplored as a source of bioactive SMs despite the ocean being home to the majority of phyla of life compared to land. Some marine scientists speculate that the chances of discovering new drugs from marine sources are a thousand times higher than the terrestrial sources (Bull and Stach, 2007; Debbab et al., 2010; Sabdono and Radjasa, 2011). In 2014, about 25 marine SMs were either in clinical trials or preclinical trials as anti-cancer agents (Newman and Cragg, 2014). This demonstrates that this environment is a potential source of new drug leads. The success of finding new chemical structures and employing marine derived SMs as treatments are the main reason this study focused on the exploitation of the marine environment as a source of new SM.

1.4.1 Marine bacteria

Marine bacteria are widely distributed in the ocean as they are found free-floating, attached to surfaces and living within marine eukaryotes. They are presented with unique and challenging environmental conditions such as high pressure, salinity, a range of temperatures, areas with little to no light exposure and where they are challenged by a diversity of other microbes and their metabolites (Zhao et al., 2011). These conditions have enabled marine bacteria to develop features of adaptability and survival, resulting in them harbouring rich biological, chemical and genetic

diversity which can be exploited for human health applications (Debnath et al., 2007). For example, marine bacteria produce docosahexaenoic acid (DHA) **[18]** in response to increased pressure and low temperature; however, the metabolite is used for the treatment of high cholesterol and heart disease (de Carvalho and Fernandes, 2010) and bisucaberin **[19]** a naturally iron scavenging compound is studied for anti-tumour activity isolated from *Alteromonas haloplanktis* and *Vibrio salmonicida* (fish pathogen) (de Carvalho and Fernandes, 2010; Kameyama et al., 1987; Winkelmann et al., 2002). Some of these marine bacteria live in an endosymbiotic relationship with marine invertebrates such as sponges, tunicates and bryozoans; and these have become a particular focus for drug discovery (Piel, 2006).

Docosahexaenoic acid [18] **Bisucaberin [19]**

1.4.2 Marine invertebrate symbionts

Marine invertebrates such as sponges, bryozoans, fish and tunicates are known to be a rich source of new and chemically diverse SMs with biomedical value (Sanchez et al., 2012). However, for the purpose of this study, only marine invertebrates will be reviewed in detail. Marine invertebrates are multicellular invertebrates that have been in existence for more than 800 million years. Sponges filter high volumes of seawater which contain a wide variety of microorganisms, providing them with the opportunity to colonise sponge tissue and thus forming an endosymbiotic relationship (Sabdono and Radjasa, 2011). Recent research has demonstrated that many of the SMs isolated

from invertebrates and other marine animals are not produced by the host, but are produced by bacterial strains living in association with these marine organisms (Harvey, 2008; Piel, 2006; Sabdono and Radjasa, 2011; Zarins-Tutt et al., 2016).

This was revealed by identifying that some of the secondary metabolites from some marine invertebrates are very similar to those identified from marine endosymbionts / bacteria (Table 1.2) (Devi et al., 2010). For example, 3,5-dibromo-2-(3',5'-dibromo-2'-methoxyphenoxy) phenol **[20]** was initially thought to be produced by a *Dysidea* sp. until it was isolated from a *Vibrio* sp. associated with the invertebrate (Elyakov et al., 1991). The recently clinically approved Ecteinascidin 743 (ET-743, Yondelis) initially isolated and thought to be from a tunicate (*Ecteinascidia turbinata*) was discovered to be produced by a bacterium (*Candidatus* Endoecteinascida frumentensis) using metagenomic techniques (Schofield et al., 2015). Conversely, Mimosamycin **[21]** was initially isolated from *Streptomyces lavendulae* No. 314 and later discovered from the marine invertebrates *Reniera* sp. and *Xestospongia* sp. (Kelecom, 2002). These findings show that the discovery can be vice versa whereby if the former is the case, the true producer will have to be verified. A more comprehensive review on this matter was published by Piel and co-workers in 2004 and 2009.

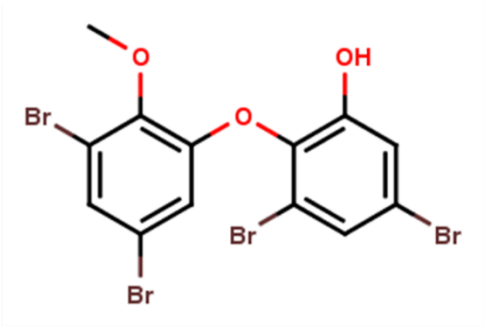
Table 1. 2. Marine invertebrate SMs with structural relation to bacterial SMs from marine invertebrate endosymbionts.

Marine invertebrate metabolite	Host Invertebrate	Bacterial metabolite	Producer (bacteria)
Jaspamide	*Jaspis* sp.	Chondramide	*Chondromyces crocatus*
Onnamides	*Theonella swinhoe*	Onnamides	*Candidatus* Entotheonella factor/gemina.
Swinholide A	*T. swinhoei*	Tolytoxin	*Tolypothrix sp.*

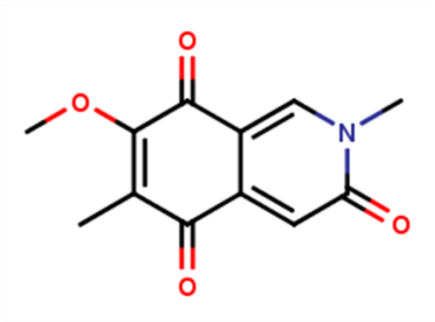
Kasumigamide	*Discodermia calyx*	Kasumigamide	*Entotheonella sp.*
Salicylihalamide	*Haliclona* sp.	Apicularen A	*Chondromyces sp.*
Manzamine A	*Acanthostrongylophora* sp.	Manzamine A	*Micromonospora* sp.
Andrimid	*Hyatella* sp.	Andrimid	*Vibrio* sp.
Mimosamycin	*Petrosia* sp.	Mimosamycin	*Streptomyces lavendulae*

Adapted from (Atta-ur-Rahman, 2017; Nakashima et al., 2016; Piel, 2009, 2004; Sabdono and Radjasa, 2011).

Marine invertebrates are therefore now known as an attractive source of microorganisms with the capacity to produce bioactive SMs and have become a focus of research in the last decade. The fact that microorganisms are responsible for the production of many of these SMs, could solve another bottleneck associated with drug discovery as they represent a sustainable and scalable supply of new SMs (Sabdono and Radjasa, 2011). As microorganisms can be scaled up for large fermentation and are easily preserved for a longer period through appropriate storage protocols, this will conserve the environment as there is no need for continuous sampling of the macroorganisms which are more difficult to harvest and to maintain as stocks.



3,5-dibromo-2-(3',5'-dibromo-2'-methoxyphenoxy) phenol [20]



Mimosamycin [21]

Microorganisms found in these invertebrates can constitute up to 40% of the total invertebrate biomass (Anand et al., 2006). Marine invertebrates such as *Phorbus tenacior* (Dupont et al., 2014), *Haliclona (gellius) sp.* (Sipkema et al., 2011), *Haliclona simulans* (Phelan et al., 2012), *Pseudoceratina clavata* (Kim and Fuerst, 2006), *Raspailia ramosa* and *Stelligera stuposa* (Jackson et al., 2012) are known to harbour a wide range of marine bacteria. Endosymbiosis happens in two ways, namely through vertical or horizontal transmission. The vertical transmission means that the bacteria were transferred to the embryo by the parent during development while horizontal means that the bacteria were acquired during filter feeding (Garate et al., 2017; Versluis et al., 2018).

A range of microorganisms belonging to Chlamydiae (Kjeldsen et al., 2010), Proteobacteria (Skariyachan et al., 2014), Planctomycetes, Verrucomicrobia, Deltaproteobacteria (Sipkema et al., 2011), Bacteroidetes (Jackson et al., 2012), Firmicutes (Jackson et al., 2012) and Actinobacteria (Dupont et al., 2014; Xi et al., 2012) have been found in association with marine invertebrates. Altogether, 26 phyla have, thus far, been shown to have an association with marine invertebrates indicating a rich diversity of microbial species with the marine invertebrates (Jackson et al., 2012).

1.4.3 South African marine research

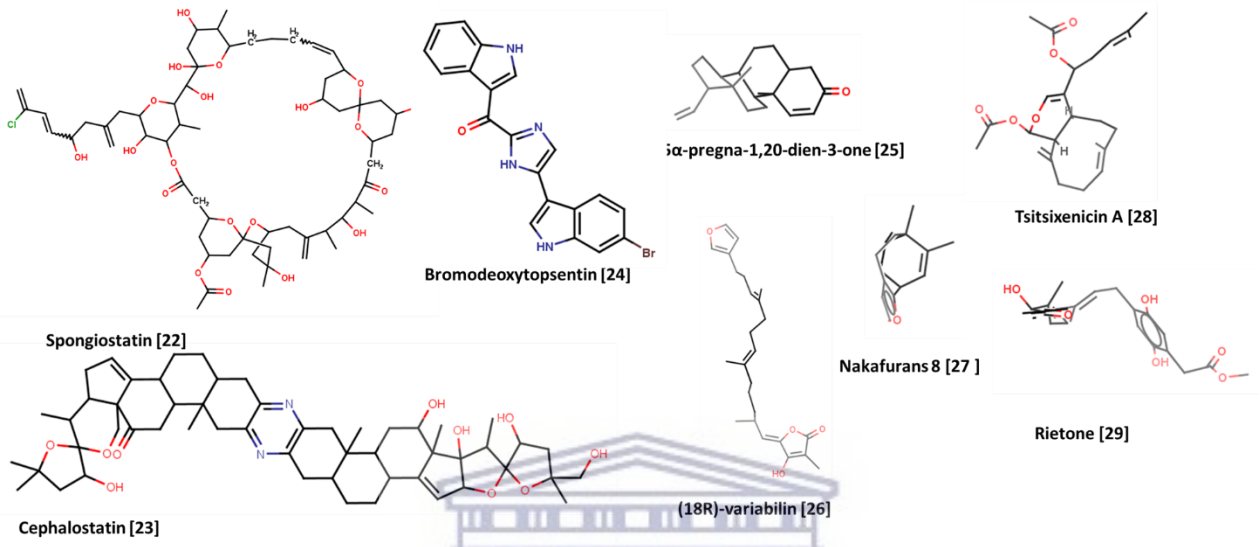
The SA marine environment encompasses a 3650 km long coastline which represents a rich source of marine biodiversity (molluscs, soft corals, ascidians, algae, sponges and microorganisms) (Davies-Coleman and Beukes, 2005; Griffiths et al., 2010). According to Griffiths and co-workers (2010), 12 914 marine biota species across multiple families and a number of undocumented small-bodied invertebrates were recorded to exist in South African oceans. This indicates unimaginable bioprospecting potential for SA marine researchers. However, only a few invertebrates have been exploited for bioprospective purposes thus far (Davies-Coleman and Veale, 2015). The three biogeographic zones (warm temperate southeast coast, cool temperate west coast and subtropical temperature east coast) in the South African marine environment present biodiversity that harbours a wide range of SM, which is currently underexplored (Griffiths et al., 2010).

Genetic studies have indicated that these different biogeographic zones harbour certain species that are endemic to each one (Jooste et al., 2018). This means that each biogeographic zone is a potential source of different chemicals produced by the same or different marine organisms. As stated in Section 1.4.1 these environmental conditions and other parameters present in the marine environment forces organisms to produce unique metabolites to survive. Therefore, the South African marine environment possesses the potential to be a source of macro- and micro-organisms with the ability to produce specialised metabolites based on the biogeographic zone of origin.

The South African marine environment showcased its potential as a source of bioactive compounds in 1979, with the discovery of spongiostatin **[22]** and cephalostatin **[23]** as anti-cancer compounds from marine invertebrate *Spirastrella spinispirulifer* (Pettit et al., 1993) and tubeworm *Cephalodiscus gilchristi* (Pettit et al., 1988), respectively. Despite the two discoveries being a significant moment in the history of South African marine research an act of biopiracy was conducted during the collection of these invertebrates with 450kg and 2409 kg of material for *Cephalodiscus gilchristi* and *Spirastrella spinispirulifer*, respectively, collected. Despite these huge amounts their development was hampered due to lack of sufficient compound from the sponges. Studies later discovered the same compounds from other invertabrates sourced in other environments suggesting a possibility of microbial origin (Davies-Coleman, 2005). This highlights the need to accelerate the identification of the true producers as this has the potential to prevent overharvesting of sponges (and other organisms) from the environment as microorganisms are a more sustainable solution compared to other alternatives such as aquaculture and chemical synthesis.

Since then, several bioactive compounds have been isolated from the SA coastline such as bromodeoxytopsentin **[24]** (anti-microbial), 5α-pregna-1,20-dien-3-one **[25]** (anti-inflammatory), (18R)-variabilin **[26]** (anti-microbial), nakafurans 8 **[27]** (anti-microbial), tsitsixenicin A **[28]** (anti-inflammatory) and rietone **[29]** (anti-viral) (Davies-Coleman and Beukes, 2005; Hatherley et al., 2015). Despite the four decades of research on marine invertebrates as a source of SMs, it is worth noting that only two, namely *C. gilchristi* (compound **[23]**; anti-cancer) and *Topsentia*

pachastrelloides (compound [**24]**; anti-MRSA) have provided SMs whose bioactivity are of interest to pharmaceutical companies (Davies-Coleman and Veale, 2015).



Despite the high number of invertebrates recorded, studies evaluating the potential of South African marine macro- / micro-organisms are limited (Matobole et al., 2017). It is worth noting that none of the above mentioned compounds were isolated from marine bacteria or fungi. To our knowledge only one published study exists outside of the Institute for Microbial Biotechnology and Metagenomics (IMBM) showcasing the potential of South African marine bacteria as a source of bioactive compounds; however, the study only focused specifically on actinomycetes (Cwala et al., 2011). This highlights the need for the South African marine environment to be explored as a source of marine bacteria with the potential to provide new SMs.

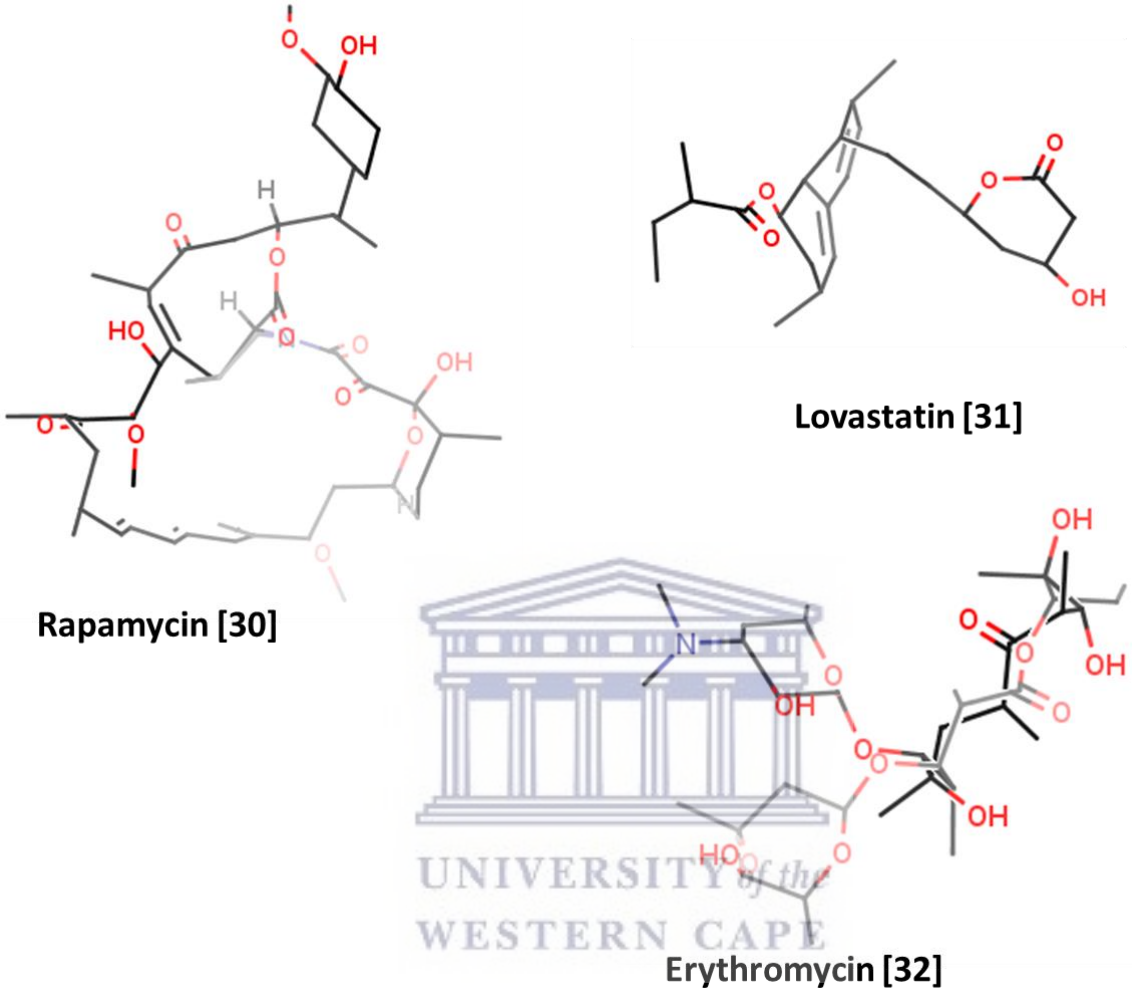
The institute (IMBM) was established in 2007 and is focused on generating scientific knowledge in the field of biotechnology. The aim is to develop a valuable biotechnological product for application in areas such as pharmaceutical, food and agriculture using genomic, microbiology and molecular biology. The institute has a bacterial collection consisting of over 3900 bacterial strains isolated from the terrestrial and marine environment from a variety of sources (e.g., marine invertebrates). Additionally, there are several metagenomic libraries generated from different environments.

1.5 Types of microbial secondary metabolites

Microorganisms produce a wide range of SMs which are grouped into different classes based on the enzymes used in their biosynthesis and chemical/biological properties. The classes of SMs include polyketide (PK), non-ribosomal peptides (NRP), ribosomally-synthesised and post-translationally modified peptides (RiPPs), siderophores, terpenes and hybrids of these classes (Hanssen, 2014). The precursors for the synthesis of SMs are obtained from the primary metabolite and their intermediates. Below is a discussion of each class and its biosynthesis.

1.5.1 Polyketides

Polyketides (PK) are a class of biological heteropolymers with high structural diversity and are biosynthesised from acyl-CoA (e.g. carboxylic acid) precursors utilising polyketide synthases (PKS) (Gao et al., 2013; Jenke-Kodama and Dittmann, 2009). Polyketides are very important as most of them and their derivatives are known to make good drug candidates, of which rapamycin **[30]**, tetracycline **[7]**, daunorubicin **[1]**, lovastatin **[31]** and erythromycin **[32]** are good examples. Polyketide-derived pharmaceuticals represented 20 % of the top selling drugs in 2005 according to a report by Weissman and Leadlay, which in terms of monetary value amounts to over UK £10 billion per year worldwide (Gao et al., 2013). However, more work needs to be done in evaluating the exact ecological and biological role of PKs in nature as is the case with many other types of SMs (Jenke-Kodama and Dittmann, 2009), which represents a gap in SM research.



1.5.1.1 Types of polyketides

Polyketides are divided into three types namely; type I, type II and type III, based on the type of PKS enzymes utilised in their synthesis (Cheng et al., 2009). Type I PKs are formed from an enzyme (type I PKS) with a modular or iterative mode of action. Iterative mode relates to the same domains (refer to Section 1.5.1.2) being used more than once during the biosynthesis. Type II is catalysed by multienzymes like the type I iterative PKS with the difference being the ketosynthase

(KS) domain. Type II PKSs have two KS domains (KS_α and KS_β) with the KS_β being responsible for controlling the polyketide length and the β-keto generated by the KS domain is modified at the end of the reaction. Type III PKs are formed from type III PKS. Type III PKSs are different to type I and II because they lack multiple catalytic domains and utilize an acyl carrier protein (ACP)-independent mechanism (Cheng et al., 2009; Hochmuth and Piel, 2009). The organisation of these enzymes is similar to the animal fatty acid biosynthesis pathway; hence they are named similarly.

1.5.1.2 Biosynthesis of polyketides

Polyketide synthases (PKS) are multifunctional proteins consisting of three main domains namely; ketosynthase (KS), acyltransferase (AT), and the acyl carrier protein (ACP). Other optional domains include the ketoreductase (KR), dehydratase (DH), methyltransferase (MT) and enoyl reductase (ER), responsible for the modification of the produced PK core structure (Chan et al., 2009; Jenke-Kodama and Dittmann, 2009). Commonly used precursors for the biosynthesis of all types of PKs are acetyl-CoA, malonyl-CoA, benzoyl-CoA, isobutyryl-CoA, and acetoacetyl-CoA (Kirimura et al., 2016). The KS domain is responsible for condensation of the ACP linked to a malonate intermediate or an ACP linked acyl thioester to increase the length of the PK molecule. An AT domain is required for the recognition of an extender unit to be incorporated into a growing chain. Figure 1.3 demonstrates the biosynthesis of PK (Chan et al., 2009).

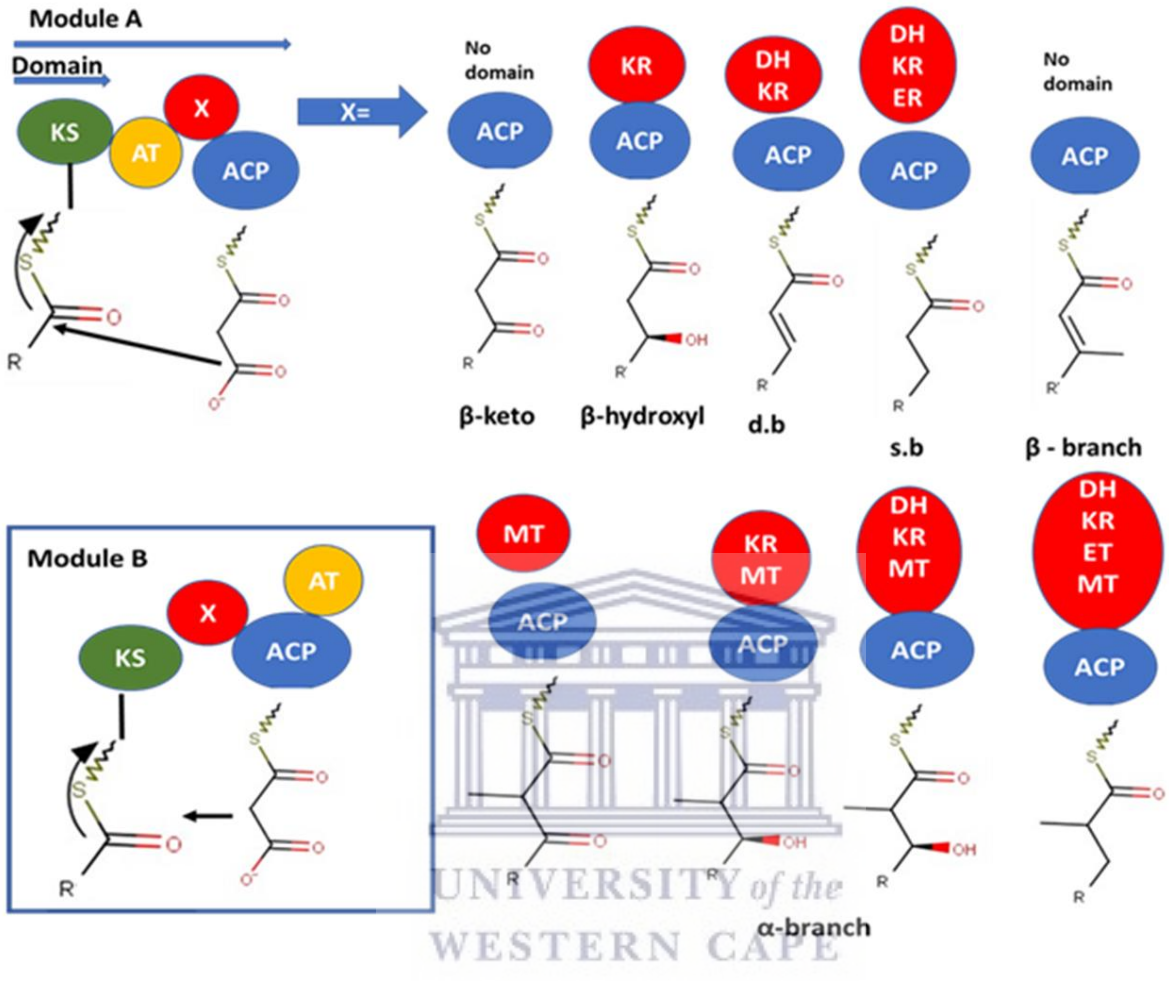
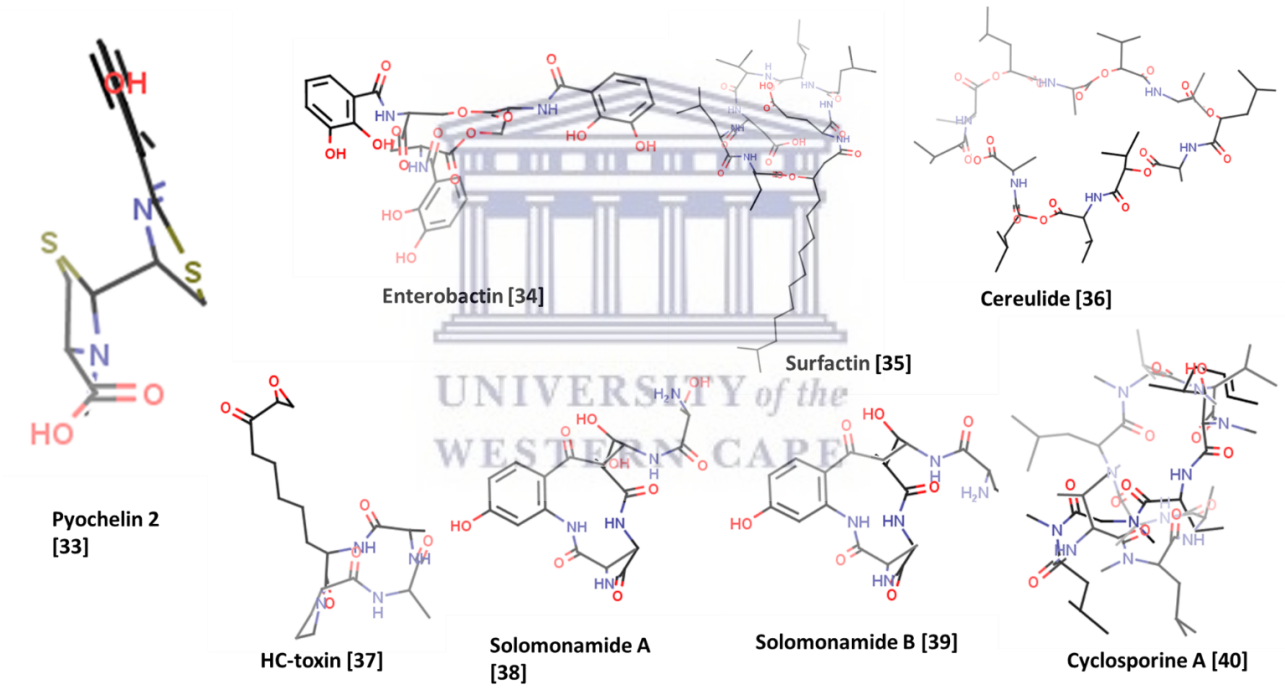


Figure 1.3. A simple arrangement of PKS domains responsible for the synthesis of a polyketide is shown above, using one modular adapted from (Hochmuth and Piel, 2009). Legend: ketosynthase (KS), acyltransferase (AT), an acyl carrier protein (ACP), ketoreductase (KR), dehydratase (DH), methyltransferase (MT) and enoyl reductase (ER). Module A: *cis*-AT Polyketide synthases (PKS) and module B: *trans*-AT PKS.

1.5.2 Non-ribosomal peptides

Non-ribosomal peptides (NRP) are a class of compounds synthesised from both non-proteinogenic and proteinogenic amino acids using large multimodular enzymes similar to PKS, known as non-ribosomal peptide synthases (NRPS). These peptides have a broad range of biological/chemical activity and structural diversity (Finking and Marahiel, 2004; Neilan et al., 1999). NRPs include a wide range of important compounds such as antibiotics (vancomycin **[10]** and penicillin **[4]**),

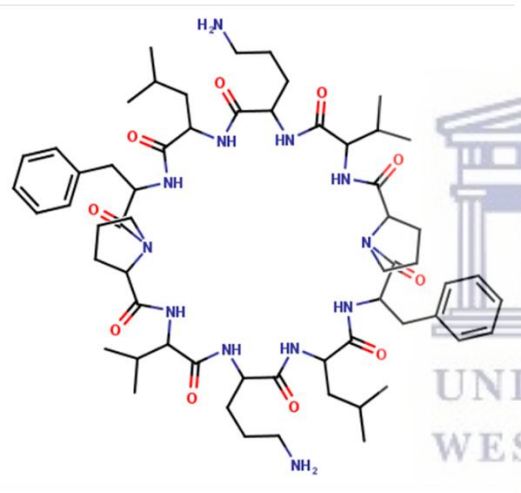
siderophores (pyochelin 2 **[33]** and enterobactin **[34]**), biosurfactants (surfactin **[35]**), toxins (cereulide **[36]** and HC-toxin **[37]**), anti-inflammation agents (solomonamide A **[38]**-B **[39]**) and immunosuppressants (cyclosporine A **[40]**) (Agrawal et al., 2016; Caboche et al., 2007; Horwood, 2004; Walsh, 2004; Zhang et al., 2009). Non-ribosomal peptides are small peptides when compared with ribosomal-synthesised proteins and were previously thought to be produced by actinomycetes and bacilli only. However, studies from both marine and terrestrial bacteria have shown the distribution of NRPs in several bacterial taxonomic classes (Caboche et al., 2007; Horwood, 2004).



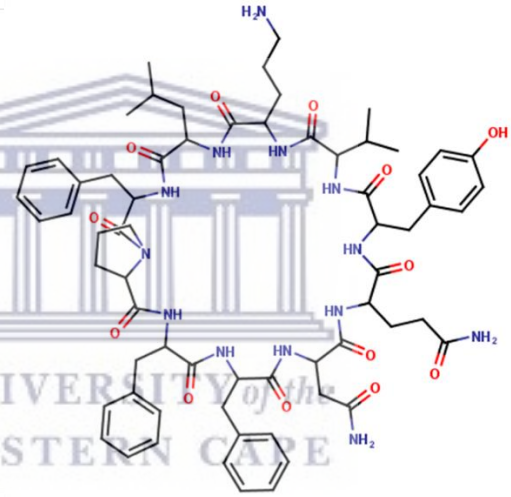
1.5.2.1 Biosynthesis of non-ribosomal peptides

The biosynthesis gene cluster (BGC) encoding the biosynthesis of a NRP is arranged as an operon similar to PKs and other SMs (Figure 1.4). The NRPS are organised as multi-modular proteins containing a number of domains for the biosynthesis of the final product (Agrawal et al., 2017; Aleti et al., 2015). Precursors for NRP include the common 20 proteinogenic amino acids, non-

proteinogenic amino acids, sugars, fatty acids and α-hydroxyl acids which can add up to approximately 500 different monomers. A wide range of precursors and modifications performed on these structures contribute toward the chemical diversity and the associated biological activities observed amongst the NRP class. The mechanism for biosynthesis was first described in 1971 during a study to understand the synthesis of gramicidin S and tyrocidin antibiotics (Caboche et al., 2007).



Gramicidin S [41]



Tyrocidine [42]

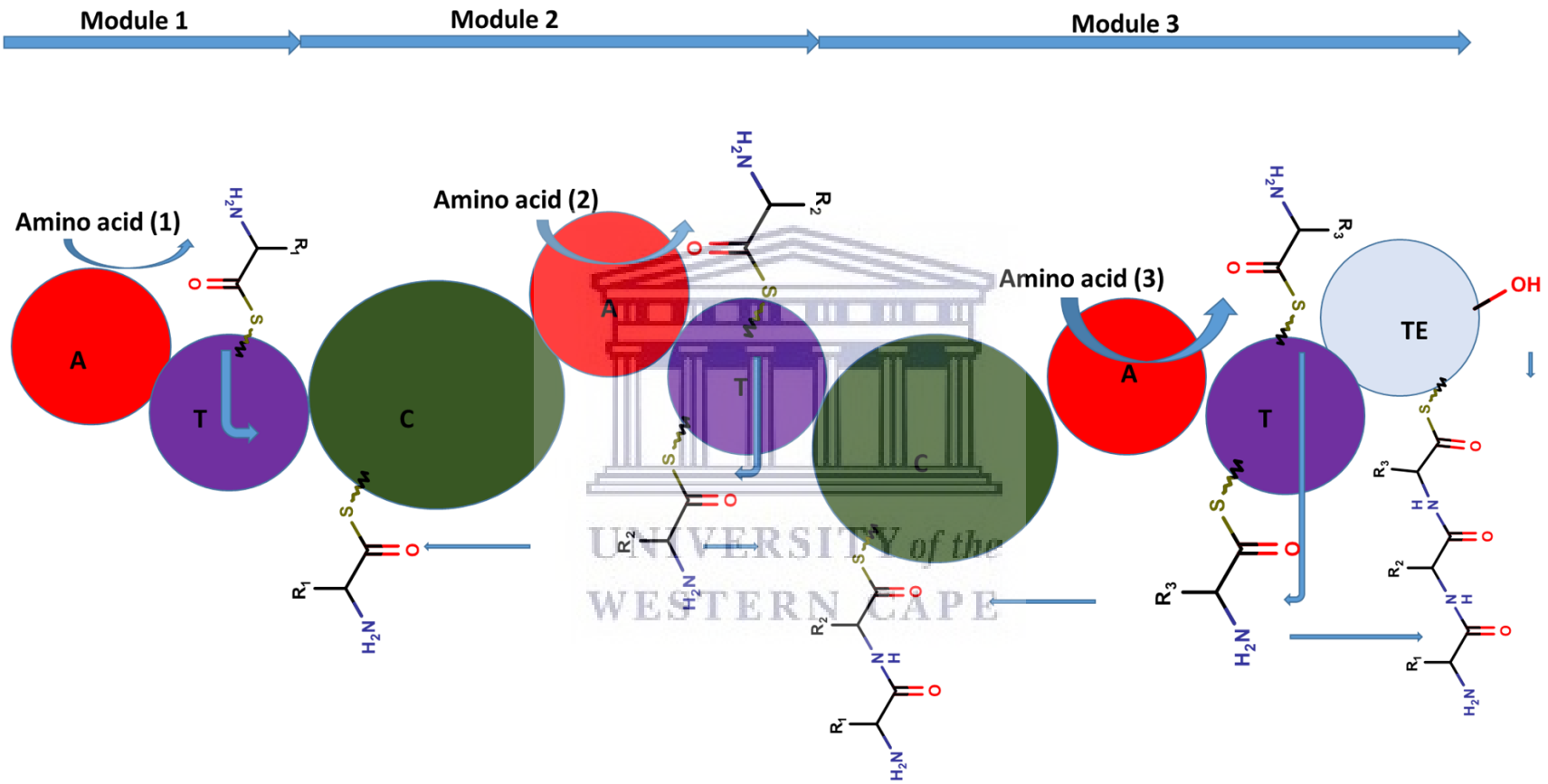


Figure 1.4. A simple arrangement of NRPS domains responsible for the synthesis of a NRP was adopted (Winn et al., 2016). Legend: adenylation domain (A), thiolation domain (T), condensation domain (C) and thioesterase domain (TE).

Three primary domains must be present in a module for the synthesis of a NRP; and these include the adenylation (A) domain for the activation of the amino acid, peptidyl carrier protein (PCP) domain for the transportation and delivery of the peptidyl substrate to the peptide chain, also known as the thiolation domain (T), and condensation (C) domain for the condensation of the activated amino acid. Additionally, a thioesterase (TE) domain is required for the release of the synthesised NRP by hydrolysis or a macrocyclization reaction, and housekeeping by removing incorrectly incorporated amino acids (Strieker et al., 2010; Winn et al., 2016).

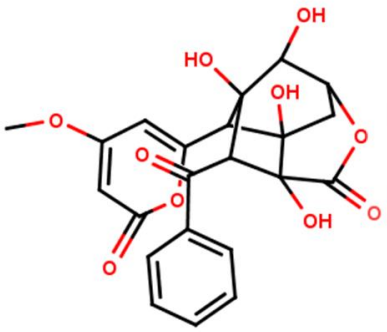
Modification domains such as the epimerization (E) and methyltransferase (MT) are responsible for catalysing the epimerization of the α-carbon from the amino acid and methylation of the nitrogen from the amine group, respectively, which are optional within gene clusters (Challis and Naismith, 2004). NRPs are subjected to similar modifications as PKs and other SMs.

1.5.3 Ribosomally-synthesised and post-translationally modified peptides (RiPPs)

RiPPs are peptides with diverse biological activity and are biosynthesised either by micro- and macro-organisms (Wang, 2012). This class of compounds was previously classified based on bioactivity or the producing organism (e.g. bacteriocins active against bacteria) until the coining of the term RiPP in 2011 (Arnison et al., 2013; Joerger, 2003; Letzel et al., 2014).

These peptides are lethal to other bacterial strains but not the producing strain and are usually highly specific to their targets. Their target-specific property makes them ideal candidates for the treatment of infections leaving the natural microbiota unaffected, as this is known to promote secondary infections (Letzel et al., 2014). RiPPs can also operate using similar mechanisms as those of known antibiotics, by disruption of RNA or DNA and protein synthesis; however, the most common mechanism is the disruption of the cell membrane causing pores/openings resulting in cytoplasm leakage (Joerger, 2003).

RiPPs are usually a short peptide chain consisting of up to 36 amino acids which often gains its anti-microbial property due to post-translation modifications (Joerger, 2003). RiPPs are classified based on the starting unit or key reaction during synthesis (Stepper et al., 2011). For example, members of the thiopeptide family are characterised by a central azole substituted nitrogen-containing six-membered ring (Shen et al., 2019). RiPPs encompass compounds such as bacteriocins, microcins, lanthipeptides, microviridin, lasso peptides, thiopeptides, sactipeptides, and linear azol(in)e-containing peptides (LAPs) (Arnison et al., 2013; Joerger, 2003). Lactic acid bacteria are well-known for their ability to produce these anti-microbial peptides and are the best-studied example (Parada et al., 2007). For example, two bacteriocins were isolated from *Enterococcus durans* A5-11B, namely enterocins A5-11A and A5-11B **[43]** (Batdorj et al., 2006).



Enterocin [43]

1.5.3.1 Biosynthesis of anti-microbial RiPPs

Unlike NRPs, ribosomes are responsible for the synthesis of RiPPs. Similarly, to the genes involved in the biosynthesis of NRPs and PKs, those responsible for RiPP biosynthesis are also arranged in clusters (operons). Genes responsible for regulation, biosynthesis and modification are found adjacent to each other (Figure 1.5) (Arnison et al., 2013). Precursor peptide chains (typical 20 to 110 amino acid residues) produced during the biosynthesis of RiPPs contain a conserved N-terminal leader sequence with a double glycine motif that is cleaved at the C39 peptidase domain (Arnison et al., 2013; Wang, 2012).

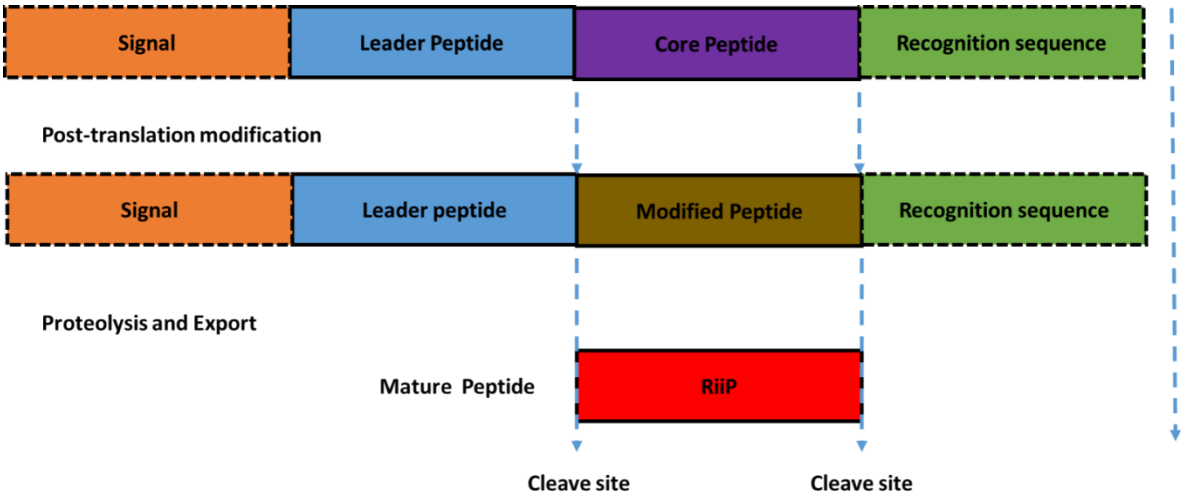


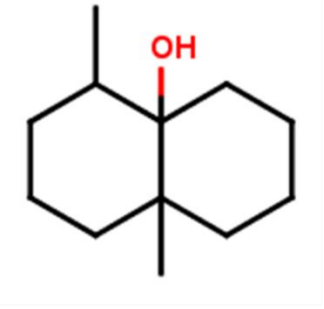
Figure 1.5. Typical biosynthesis of RiPPs from precursor peptides (adapted from (Wang, 2014))

The N-terminal leader sequence guides the peptide to a specific section within the cell for modification and secretion. In the rare instance where the leader peptide is attached to the C-terminal, it is known as a "follower" peptide (Huo et al., 2012), and C-terminal recognition sequences are responsible for cleaving and cyclisation. Upon cleavage, a modified core peptide is produced and becomes known as the RiPP. Other precursors used during the synthesis of RiPPs are nitrile hydratase-like leader peptides and leader regions resembling Nif11 nitrogen-fixing proteins (Arnison et al., 2013). Tailoring such as lanthionine formation, macrocyclization, dehydration and heterocyclisation are common amongst RiPPs (Cox et al., 2015; Oman and van der Donk, 2010).

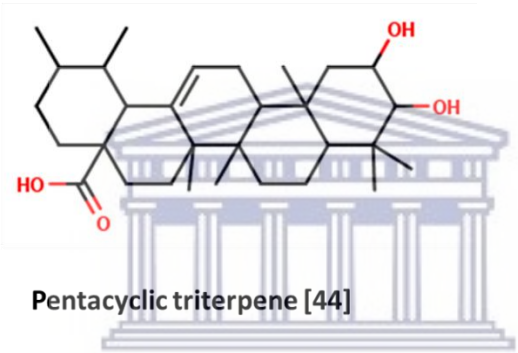
1.5.4 Terpenoids

Terpenoids are a class of SM which are biosynthesised from 5 carbon isoprene units and represent the largest SM class with over 50 000 compounds having been reported thus far (Matsuda et al., 2015; Peralta-Yahya et al., 2011). Terpenoids are also known as terpenes. Terpenoids are further categorized into different subclasses, namely mono-, sesqui-, di- and sester-terpenoid. The high structural diversity observed lies in the variation within the terpene cyclase enzymes responsible for their biosynthesis (Matsuda et al., 2015; Yamada et al., 2015). Terpenoids

were previously thought to be mostly common to plants and fungi. However, their presence within bacteria, first discovered in 1971 with the discovery of pentacyclic triterpene **[44]** (Rosa et al., 1971) and later through developments in genome mining, revealed that terpenoids are also common in bacteria (Cane and Ikeda, 2012; Yamada et al., 2012). The best known bacterial terpenoid is geosmin **[45]** produced by *Streptomyces*, *Cyanobacteria*, *Myxobacteria* and other microorganisms, which is responsible for the characteristic earthly fragrance following a rain event (Yamada et al., 2015).



Geosmin [45]



Pentacyclic triterpene [44]

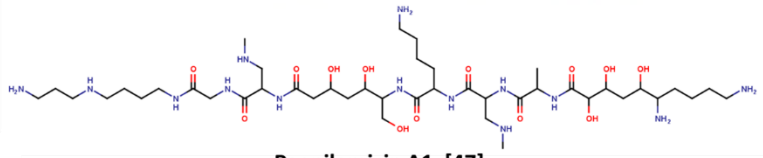
1.5.4.1 Biosynthesis of Terpenoids

Terpenoids are biosynthesised by terpene synthases and the majority of the bacterial terpenoids/ terpene are cyclic, hence the enzymes are sometimes referred to as terpene cyclases (Cane and Ikeda, 2012; Matsuda et al., 2015). Bacterial and fungal terpene synthases are different to plant terpene synthases as their enzyme complex contains one α-helical catalytic domain while the plant enzymes contain two (N-terminal and C-terminal domains). The domain is responsible for the binding of magnesium ion (Mg^{2+}) by the "DDxxD" motif and the "NSE/DTE" motif, which are essential in the cyclisation of commonly known substrates, such as farnesyl diphosphate, isopentenyl diphosphate, geranyl diphosphate, and dimethylallyl diphosphate. These precursors are supplied by primary metabolic pathways, specifically the mevalonate and deoxyxylulose 5-phosphate pathways (Cane and Ikeda, 2012; Peralta-Yahya et al., 2011). The synthesised backbone from the precursor (e.g. isopentenyl pyrophosphate or its isomer, dimethylallyl pyrophosphate) is

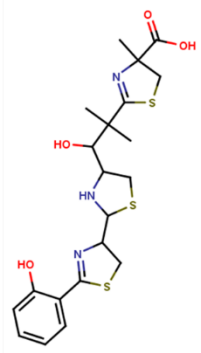
subject to modifications such as acylation, peroxidation, methylation and oxidation. These subsequent metabolic reactions generate compounds with huge structural complexity and diversity with interesting pharmaceutical properties (Odom, 2011; Pattanaik and Lindberg, 2015). Carotenoids are a good example of terpenoids produced by a wide range of micro-organisms.

1.5.5 Hybrid SMs

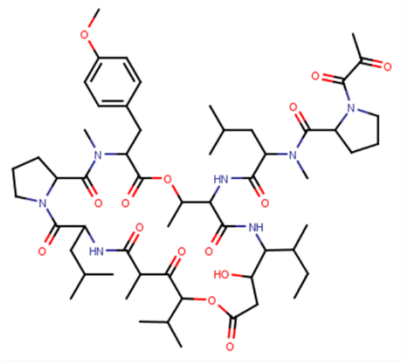
A hybrid SM is catalysed by enzymes from a combination of two or more classes. For example, two iron chelating siderophores known as yersiniabactin **[46]** (Pfeifer et al., 2003) and paenilamicin A1 **[47]** (Müller et al., 2014) are biosynthesised as NRP-PK hybrids**.** The starter unit can either be of PK or NRP origin depending on the BGC arrangement, and their biosynthesis combines features of PKs and NRPs or other classes. The marine environment has yielded NRP-PK hybrids such as aplidine **[48]** and salinosporamide A **[11]** (Lorente et al., 2014). The genes responsible for the biosynthesis of SMs can be mined from genomes using bioinformatics tools as they contain conserved regions from the individual clusters (de Jong et al., 2006; Walsh and Fischbach, 2010; Weber et al., 2015) (Hammami et al., 2007). Genome mining is elaborated in Section 1.6.5.



Paenilamicin A1 [47]



Yersiniabactin [46]



Aplidine 48

1.6 Strategies for secondary metabolite production in bacteria

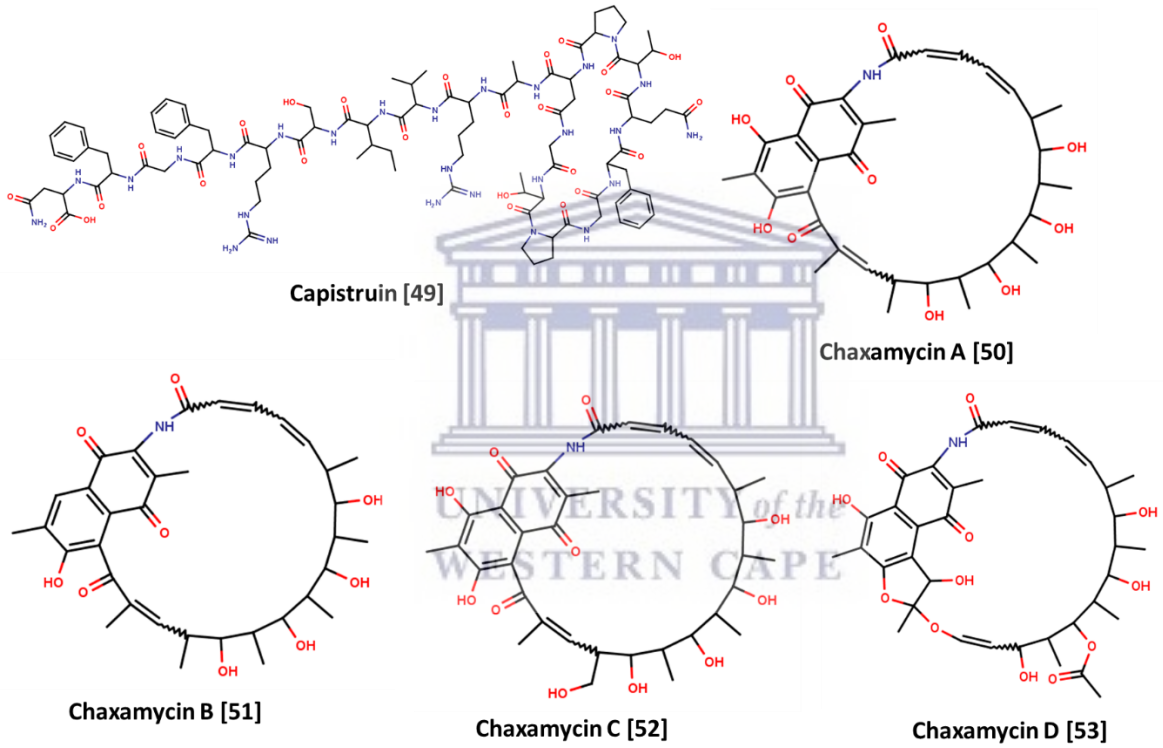
Secondary metabolite discovery is hindered by some challenges such as low or ineffective expression of the BGC under laboratory culture conditions resulting in low product titre. The inability to culture the majority (99 %) of the microorganisms on earth also serves as a major bottleneck in the discovery of new SMs (Galm and Shen, 2006; Nah et al., 2017). Due to the evidence, different strategies are used for the production of SMs in bacteria, such as one strain many compounds, co-culturing and host genetic engineering, as described in the following sections. Alternatively, the SM gene cluster is cloned into a host for heterologous production of the desired metabolite (Zarins-Tutt et al., 2016). Traditionally, SM production is achieved by employing fermentation and biological-guided assays with no genetic knowledge of the producing strain. With the genomics revolution, genome-guided isolation and genetic manipulations have also been utilised in facilitating the production of SMs (Machado et al., 2015; Weber et al., 2015; Wilkinson and Micklefield, 2007; Zarins-Tutt et al., 2016).

1.6.1 One strain many compounds

The one strain many compounds (OSMAC) approach is a process where the culturing conditions are changed in an attempt to trigger the expression of cryptic BGC or to increase product titre (Bode et al., 2002; Scherlach and Hertweck, 2009). In this approach, one bacterial strain is relied upon to produce more than one SM when challenged or when conditions are adjusted. Due to the fact that different conditions are required for the expression of different SM gene clusters, varying culturing conditions such as temperature, pH, culturing vessel (culture aeration or size), and media or media composition (carbon source, nitrogen source and salt concentration) can increase or trigger the expression of certain SM gene clusters during a defined period in its growth (Scherlach and Hertweck, 2009).

Capistruin **[49]** is an excellent example of a SM that has been produced from *Burkholderia thailandensis* E264 strain, using the OSMAC approach, by Knappe and co-workers. This was

achieved by changing the culturing conditions (medium and temperature) from M9 media at 37 ^{o}C to M20 media at 42 ^{o}C (Knappe et al., 2008). Another example of the OSMAC approach is the production of chaxamycin A **[50]** and B **[51]** from a *Streptomyces* sp. when cultured in ISP2 medium; however, when glucose was substituted with glycerol two new chaxamycins C **[52]** and D **[53]** were produced whereas chaxamycin A **[50]** and B**[51]** were no longer produced (Rateb et al., 2011; Zarins-Tutt et al., 2016).

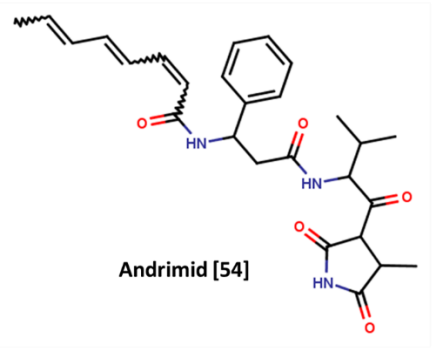


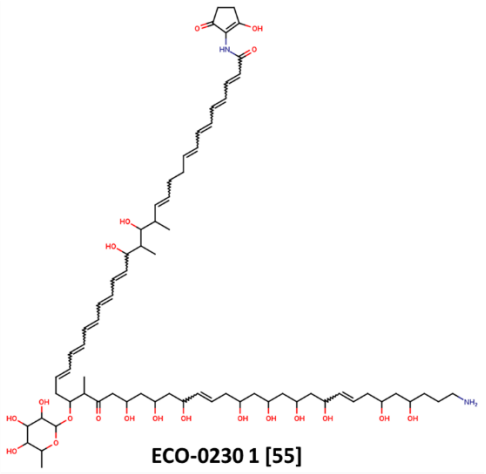
The marine *Vibrio coralliiltycus* S2052 was demonstrated to produce double the amount of the antibiotic andrimid **[50]** when grown in chitin as a carbon source compared to glucose; however, when grown on alginate it didn't produce the andrimid (Wietz et al., 2011). Andrimid **[54]** has anti-microbial activity against a wide range of bacteria such as *E. coli*, MRSA, *Klebsiella pneumoniae*, *Candida albicans* and *Bacillus subtilis* by interfering with fatty acid synthesis (Singh et al., 1997; Wietz et al., 2011). The OSMAC approach also enables the discovery

of new SM independent of the pathway knowledge, enabling the possibility of discovering new classes.

The marine *Pseudoalteromonas piscicida* PG-02 isolated from the Persian Gulf was intensely studied with an OSMAC approach whereby parameters such as temperature, nitrogen source, carbon source, NaCl % and pH to optimise the production of MRSA active antibiotic. The following conditions 28 °C (temperature), pH 7, 0.5 % NaCl, tryptone (nitrogen) and glucose (carbon) were established to be optimum for antibiotic production. The zone of clearing increased from 16 mm to 38 mm when comparing extracts of initial culturing parameters and optimized conditions, respectively (Darabpour et al., 2012). The bioprocess intensification of the product is very important as this will eliminate product shortage during development into a chemotherapeutic drug (Darabpour et al., 2012; Marwick et al., 1999). In the case of *P. piscicida* PG-02 it will assist in the production of enough material for the purification, identification and structural elucidation of the active peptide (Darabpour et al., 2012).

Genome mining of *Streptomyces aizunensis* NRRL B-11277 by McAlpine and co-workers resulted in the identification of a new BGC, and the production was induced by screening 50 different fermentation media. With the utilisation of genomic data together with knowledge of a known SM, the structure of a new SM was predicted which also assisted in purification. The compound was identified from the fermentation broth using the predicted monoisotopic mass and UV-absorption. The isolated compound was a polyene polyketide, now known as anti-fungal ECO-02301 **[55]** (Baker, 2015; McAlpine et al., 2005) and its discovery revealed the importance of combination strategies.

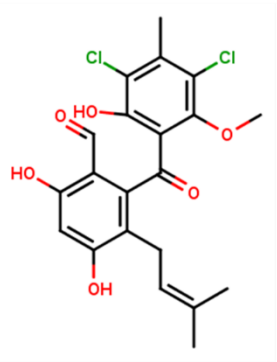




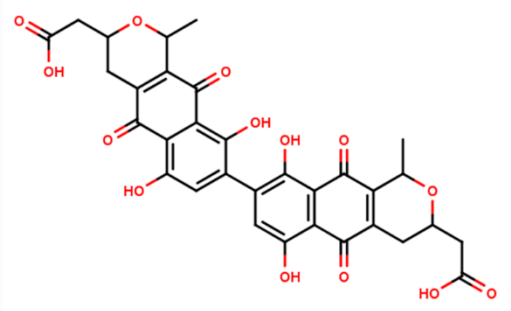
1.6.2 Co-culturing

Co-culturing involves the culturing of two or more organisms in an attempt to induce production of SM by mimicking Nature's small molecule communication (quorum sensing) system (Dashti et al., 2014a; Zarins-Tutt et al., 2016). Co-culturing can also involve culturing microbes with living or dead test/challenger strains in an attempt to induce or increase product yield (Zarins-Tutt et al., 2016). A challenger strain refers to the strain against which the production strain has shown or is known to have anti-microbial activity.

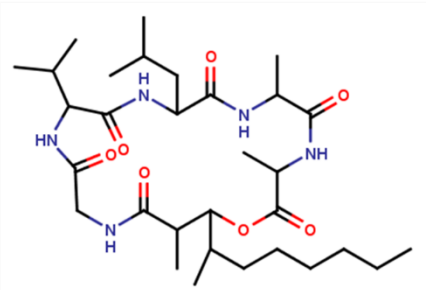
Examples of SMs produced with this approach include compounds such as chlorinated benzophenone pestalone **[56]** from *Pestalotia sp.* strain CNL-365 co-cultured with marine α-proteobacterium strain CNJ-328 (Cueto et al., 2001). Another example is a 20-fold increase in the production of actinorhodin **[57]** observed when *S. coelicolor* was co-cultured with *Myxococcus xanthus* (Pérez et al., 2011). Co-culturing is not only limited to strains from the same Kingdom, as was demonstrated with the production of cyclic depsipeptides emericellamides A **[58]** and B **[59]** when *Emericella sp.* (CNL-878) (fungus) and *Salinispora arenicola* (bacterium) were co-cultured (Dashti et al., 2014b; Oh et al., 2007).



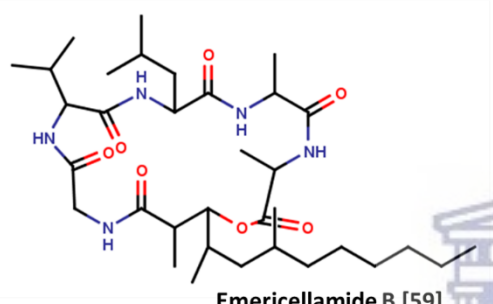
Pestalone [56]



Actinorhodin [57]



Emericellamide A [58]



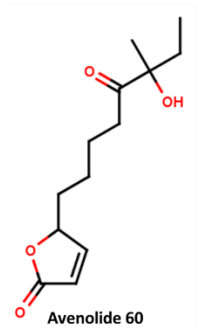
Emericellamide B [59]

1.6.3 Genetic engineering

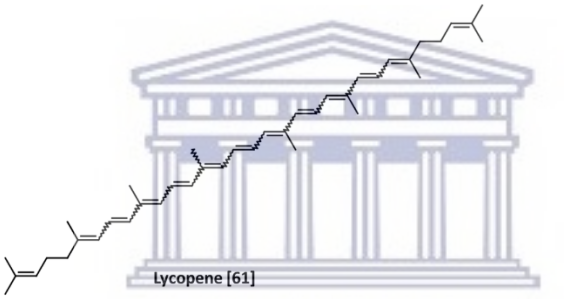
Genetic engineering involves processes used in activating the expression of cryptic pathways or improving the product titres of the already produced SM using molecular biology techniques. One approach relies on the use of mutagenesis strategies using radiation or mutagenic compounds (Unsin et al., 2013; Zarins-Tutt et al., 2016). As an example, mutagenesis was utilised in the activation of actinorhodin **[57]** expression by *Streptomyces lividans*. The activation of actinorhodin **[57]** was accomplished by mutating the *rpsL* gene responsible for streptomycin resistance (Okamoto-Hosoya et al., 2003). Mutagenesis was also utilised in the production of avenolide **[60]** from *Streptomyces albus* J1074. This was accomplished with the use of transposon mutagenesis (Ahmed et al., 2017).

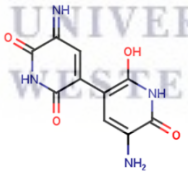
Combinatorial synthesis forms part of genetic engineering, whereby different gene clusters are combined to produce unique products through heterologous production. Other engineering

techniques involve inactivation of enzymes and manipulating the wild-type strain genetic cluster regulation, also referred to as refactoring (Unsin et al., 2013). Altering the regulatory system has been successfully used in the over-production of lycopene **[61]** in *Streptomyces avermitilis* (Bai et al., 2015). The Bode group managed to replace the natural promoter for the overproduction of GameXPeptides and indigoidine **[62]** in *Photorhabdus luminescens*, despite the production of the compounds being regulated by an unknown mechanism. Briefly, this was achieved by replacing the native prompter with the arabinose-inducible promoter P_{BAD} enabling production activation when required (Bode et al., 2015; Zarins-Tutt et al., 2016). The disadvantage of genetic engineering is that it requires a high level of skills and knowledge of the desired pathway.



Avenolide 60


Lycopene [61]



Indigoidine [62]

1.6.4 Heterologous production

Heterologous production is a technique whereby, a gene or pathway is expressed in a foreign host with the necessary mechanisms to express, synthesise and modify the product as it would have been in the native host (Wenzel and Muller, 2005). Heterologous expression has an advantage over some of the other approaches as heterologous hosts are usually phenotypically and genetically well characterised. It also helps in addressing challenges such as slow growth posed by the native strains, uncultivable microorganisms and eliminates genetic manipulation of naïve strains which

is particularly challenging as most strains do not have any or well developed genetic systems (Fu et al., 2012; Gomez-Escribano and Bibb, 2011; Wenzel and Muller, 2005).

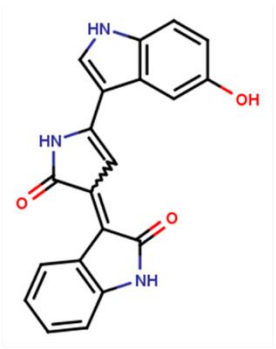
Heterologous expression also suffers drawbacks; the main challenges being efficient overexpression of the desired SM and proper structural modifications. These challenges are due to limited hosts' availability to biosynthesise some of these SMs. The hosts are inadequately equipped with systems to deal with final product modifications, synthase enzyme folding and different codon usage (Nah et al., 2017; Wilkinson and Micklefield, 2007). It is worth noting that some of these drawbacks can be addressed with a process such as codon-optimized synthetic gene clusters and co-expression of chaperones to address folding issues; however, such an exercise requires a high level of expertise and knowledge of the gene cluster and biosynthesis mechanism.

Heterologous production encompasses processes such as cloning of gene clusters from a single isolate, metagenomics, transformation-associated recombination (TAR) and "direct" cloning for larger gene clusters (Wilkinson and Micklefield, 2007). The classical cloning approach is not suitable for cloning of the majority of BGC as the vectors utilised can only accommodate small inserts (up to 40 kb) (Nah et al., 2017) and some SM BGC can be as big as 120 kb (Nah et al., 2017; Yamanaka et al., 2014). A few SMs, mainly belonging to the bacteriocin and rhamnolipids classes, have been cloned successfully with this approach in *P. putida* KT2440 (Wittgens et al., 2018).

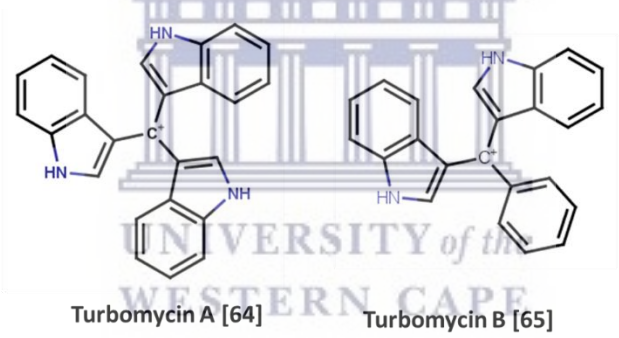
Over 98 % of microorganisms are uncultivable which represent a rich source of bioactive SM and remain unreachable using classical microbiology techniques (Streit et al., 2004). Metagenomic techniques facilitate the discovery of SM from uncultivable microorganisms and involve the extraction of DNA from a wide range of sources in the environment and the construction of fosmid /cosmid libraries. The generated library is either screened for bioactivity (functional metagenomics) or sequenced for BGC identification (sequence-based metagenomics). Functional metagenomics means a phenotypical trait should be observed while sequence-based metagenomics

involves looking for certain genes within the metagenomic DNA through next-generation sequencing (NGS).

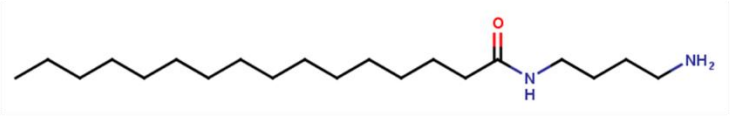
Despite the size limitations, some compounds have successfully been discovered through metagenome screening and include violacein **[63]** (6.7 kb) (Brady et al., 2001), turbomycin A **[64]** & B **[65]** (25 kb) (Gillespie et al., 2002), palmitoylputrescine **[66]** (9.3 kb) (Brady and Clardy, 2004), and vibrioferrin **[67]** (7.9 kb) (Fujita et al., 2011). To effectively discover larger BGCs from uncultured microorganisms, a BGC scattered across different cosmids is stitched together by identifying overlapping sequences (Kim et al., 2010). However, the process is complicated and labour intensive, and not possible if the library does not represent the complete genomes of all the community members, which is a common limitation.



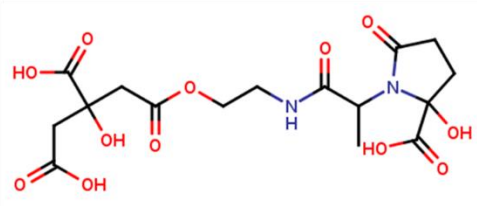
Violacein [63]



Turbomycin A [64] Turbomycin B [65]



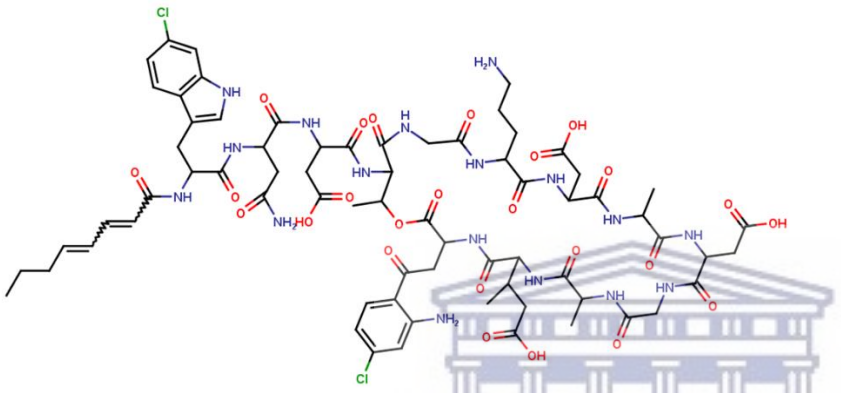
Palmitoylputrescine [66]



Vibrioferrin [67]

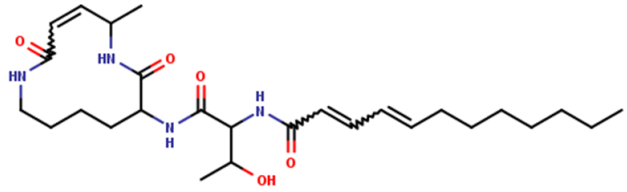
Transformation-associated recombination (TAR) cloning is particularly suited to isolate large pathways (up to 120 kb) and involves using the yeast *Saccharomyces cerevisiae* as cloning host, as direct cloning of fragments >40 kb is difficult to achieve (Yamanaka et al., 2014). This method employs the natural *in vivo* homologous recombination ability of *Saccharomyces cerevisiae* to

directly capture large genomic loci (Lee et al., 2002). Taromycin A **[68]** from *Saccharomonospora* sp. CNQ-490 was produced using this technique with *S. coelicolor* M512 being used as the heterologous expression host. Yamanaka and co-workers designed a plasmid capable of being shuttled between *E. coli* and actinobacteria and manipulated in *S. cerevisiae* (Yamanaka et al., 2014).

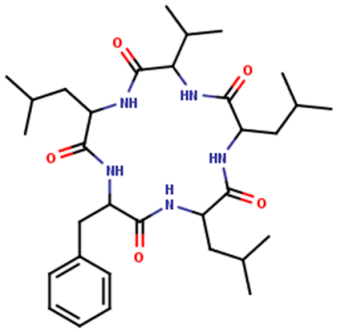


Taromycin [68]

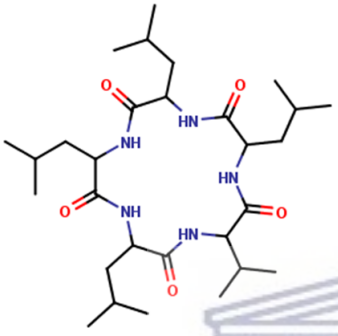
"Direct cloning" bypasses the construction and screening of libraries, utilising enzymes such as Redα and RecE from lambda phage and red prophage, respectively, in *E. coli* (Fu et al., 2012). These enzymes promote homologous recombination between foreign linear genomic fragments and *E. coli* chromosomal DNA following a similar process to the yeast-based TAR system. Luminmycin A **[69]** and luminmide A **[70]** / B **[71]** have been successfully heterologously expressed using this technique (Cobb and Zhao, 2012; Fu et al., 2012). Both direct and TAR cloning techniques are affected by repetitive DNA sequences within the genome as they interfere with the homologous recombination of the desired fragments and they require a high level of skills.



Luminmycin A [69]

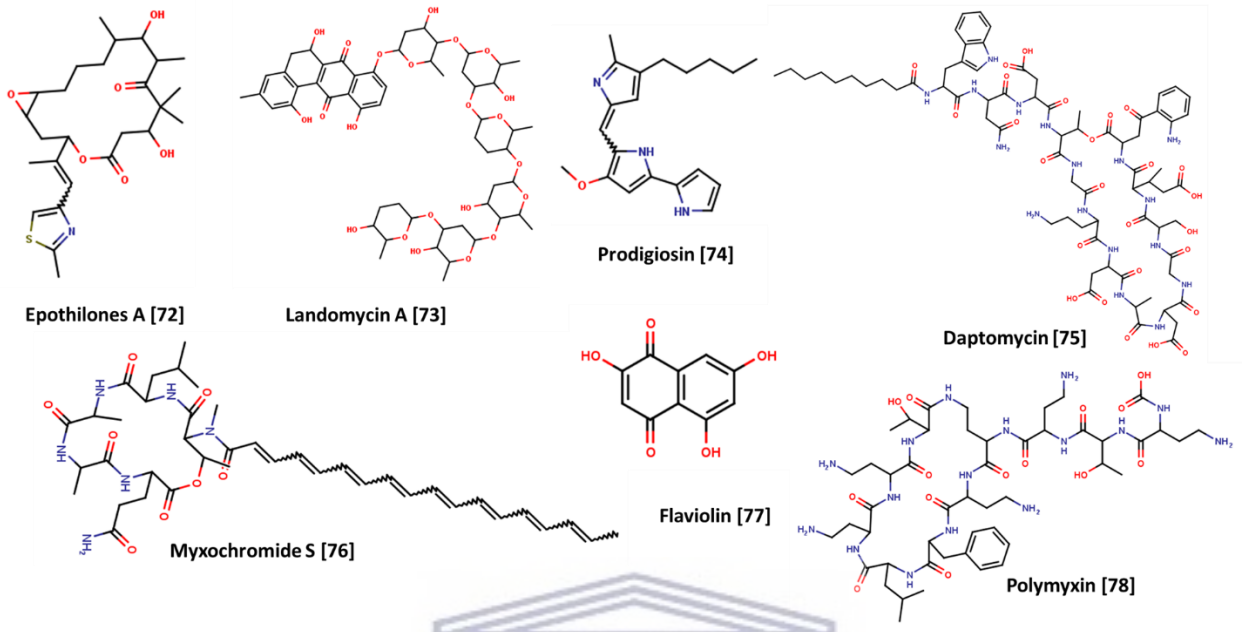


Luminmide A [70]



Luminmide B [71]

Other bacterial strains which have been used successfully for heterologous production of SM gene clusters are *Myxococcus xanthus* (e.g. epothilones A **[72]**) (Julien and Shah, 2002), *Streptomyces fradiae* (e.g. landomycin A **[73]**) (von Mulert et al., 2004), *Erwinia carotovora* (e.g. prodigiosin **[74]**) (Thomson et al., 2000), *Streptomyces roseosporus* (e.g. daptomycin **[75]**), *Pseudomonas putida* (e.g. myxochromide S **[76]**) (Wenzel et al., 2005), *Pseudomonas syringae* (e.g. flaviolin **[77]**) (Gross et al., 2006; Wenzel et al., 2005) and *Bacillus subtilis* (polymyxin **[78]**) (Choi et al., 2009). This indicates the availability of diverse bacterial hosts for wide application as the choice of host is very critical in the success of heterologous expression; however, some genera are still not represented.



1.6.5 Genome informed discovery

Microbial SM sources were historically chosen at random or based on taxonomy (16S rRNA) based on association (e.g., *Streptomyces*) which resulted in the rediscovery of known compounds. This was later rectified with the use of genome mining and metabolomics to access the strains in order to eliminate strains with only known BGCs (Ayuso et al., 2004; Ritacco et al., 2003). Genome mining is a process whereby genome sequences (from a single organism or metagenomes) are analysed for SM gene clusters or any other potentially useful genes using bioinformatics tools. Genome informed discovery involves the bioinformatics characterisation of the SM BGC of interest, gathering information to be used in fermentation medium design to help track or trigger expression of the SM of interest, and heterologous production (Gross, 2007).

The drop in sequencing costs and the availability of whole genome sequences for a large number of bacterial strains in public databases together with advances in SM bioinformatics tools has led to the utilisation of genome mining as a tool for the discovery of new SMs (Aleti et al., 2015; Machado et al., 2015). Bioinformatics tools such as antiSMASH (Weber et al., 2015), Bagel

(van Heel et al., 2013), PRISM (Skinnider et al., 2015), CLUSEAN (Weber et al, 2009) and BACTIBASE (Hammami et al., 2007) are employed in the mining of gene clusters which are responsible for SM production.

Some of these tools are able to search for more than one type of SM gene cluster (antiSMASH) while others are specific for a certain class, as is the case with Bagel (bacteriocins) (de Jong et al., 2006). The identification of these BGCs by bioinformatics tools is made possible due to conserved regions (sequences) in the BGC within the genomes of these microorganisms encoding certain enzymes used in the biosynthesis of SMs (Machado et al., 2015). Table 1.3 contains examples of some of the common genome mining tools, and more tools that can be accessed on the secondary metabolite bioinformatics portal website (http://www.secondarymetabolites.org/).

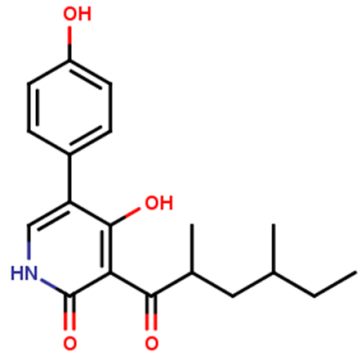
Table 1. 3. List of bioinformatics tools utilised for BGC genome mining

Programme name	Class of metabolite	Reference
antiSMASH	NRP, PK, RiPP, terpenoids, etc	(Blin et al., 2019; Weber et al., 2015)
PRISM	PKs and NRPs	(Skinnider et al., 2015)
Bagel	bacteriocins (RiPP)	(de Jong et al., 2006)
NaPDos	PKs and NRPs	(Ziemert et al., 2012)
CLUSEAN	NRPs and PKs	(Weber et al., 2009)
ClusterFinder	NRP, PK, RiPP, terpenoids, etc	(Cimermancic et al., 2014)
SMURF	PK, NRP, hybrid-PK/NRP and terpenoids (fungal)	(Khaldi et al., 2010)

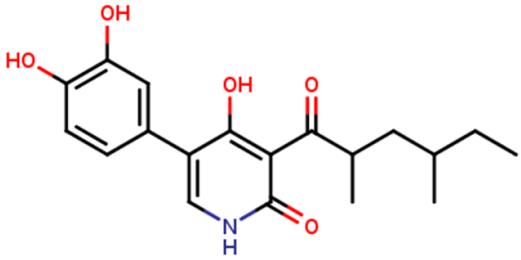
EvoMining	NRP, PKs, etc	(Cruz-Morales et al., 2016)
eSNaPD	NRPs, Pks, (metagenomic)	(Reddy et al., 2012)
CASSIS and SMIPS	PKs, NRPs and dimethylallyl tryptophan synthases (DMATs) (eukaryotic)	(Wolf et al., 2016)

A stunning discovery made through genome mining was the demonstration that the ability of microorganisms to produce SMs was underestimated by 90 % even in well-studied bacterial species. It was found that many of the pathways were silent under laboratory culturing conditions and they are referred to as cryptic / ophan BGC (Scherlach and Hertweck, 2009; Wilkinson and Micklefield, 2007; Yamanaka et al., 2014). This finding suggests that even the most exploited microorganism may still have the potential to yield new SMs.

An example of how genome mining can lead to the discovery of new compounds is that of aspyridones A **[79]** and B **[80]** from *Aspergillus nidulans*. Analysis of the genome revealed the presence of a cryptic SM gene cluster responsible for the biosynthesis of a PK-NRP hybrid when induced. The cluster was confirmed to be cryptic after HPLC-MS analysis of 40 different culturing conditions indicated the absence of a PK-NRP hybrid compound. The BGC was activated by overexpression of the putative activator gene (*apd*R) found within the cluster. The overexpression was accomplished by generating an *A. nidulans* mutant with an alcohol dehydrogenase promoter (*alc*Ap) driving expression of the *apd*R gene on a plasmid (Bergmann et al., 2007).



Aspyridone A [79]



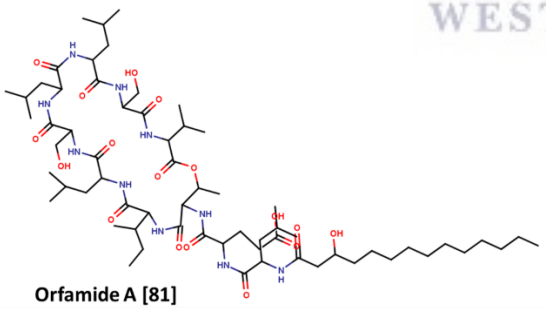
Aspyridone A [80]

Another example where genome mining has demonstrated its usefulness is the discovery of the pathways responsible for the production of compounds that have long since been known to be produced by microorganisms. The BGC responsible for the production of the sordarin compounds was discovered later while the first compound in the family was isolated in 1969, *Sordaria araneosa* Cain ATCC 36386 is one such example. The discovery of the pathway has led to the identification of enzymes responsible for the biosynthesis of sordarin and the characterisation of two enzymes GDP-6-deoxy-D-altrose:sordaricin glycosyltransferase and cycloaraneosene synthase. The characterisation of the enzymes confirmed that the correct pathway was identified (Kudo et al., 2016).

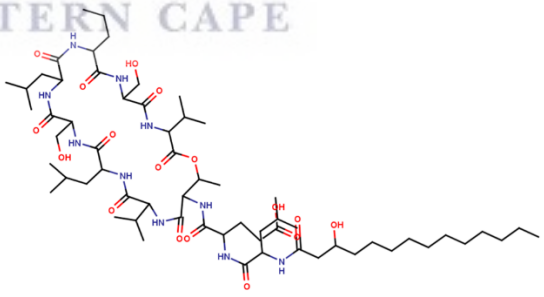
Additionally, genome mining has also revealed that 31 % of the members in the phylum Firmicutes have the ability to produce NRPs and PKs, with 70 % encoding for NRPs while PKs and hybrid metabolites make up the remaining 30 % (Aleti et al., 2015). Genome mining permits an alternative means of studying and inducing the expression of cryptic genes to produce the respective SMs, through cloning and heterologous expression of the BGC. Furthermore, the discovery of the BGC will enable the development of strategies to increase product yield such as promoter replacement. Increasing product yield is vital for the downstream processing of the SM.

The naspamycin BGC was heterologously expressed in *S. coelicolor* M1154 after being detected by a PCR probe from *Streptomyces sp.* DSM5940. The PCR probes were designed from a putative uridylpeptide BGC identified from the mining of *Streptomyces roseosporus* NRRL15998 genome (Kaysser et al., 2011). The heterologous expression of gene clusters facilitates the linking of the SM to the BGC responsible for their production and fast tracks structural characterisation of the produced metabolite. BGC mining can also assist in dereplication of known SMs based on BGC similarity to known clusters (Kaysser et al., 2011).

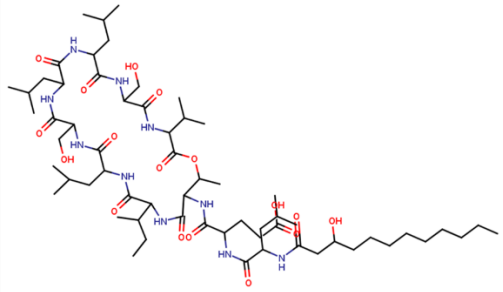
Orfamide A **[81]**, B **[82]** and C **[83]** were discovered using this approach from *Pseudomonas fluorescens* Pf-5. Briefly, the genome of the strain was mined for the presence of SM BGCs which resulted in the identification of a NRP gene cluster. The BGC was intensively characterised using bioinformatics tools to identify the main precursor which happened to be leucine. After culturing of the *Pseudomonas fluorescens* Pf-5 in several fermentation media, expression of the BGC was evaluated with a real-time polymerase chain reaction (RT-PCR). Once a suitable expression medium was identified, radiolabelled ^{15}N-leucine was used for the tracking of the SM during purification with ^{15}N-HMBC NMR (Gross et al., 2007; Paulsen et al., 2005).



Orfamide A [81]

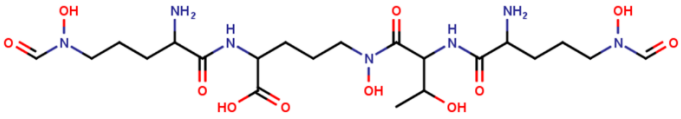


Orfamide B [82]



Orfamide C [83]

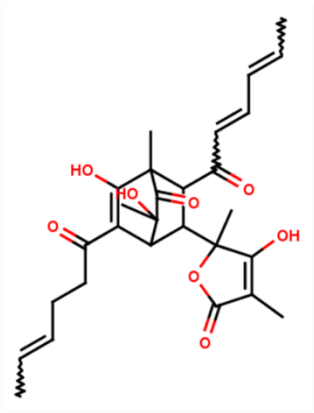
The siderophore, coelichelin **[84]** was also discovered using this approach. Genome mining of *S. coelicolor* M145 resulted in the identification of an orphan BGC responsible for its biosynthesis. After bioinformatics characterisation BGC the predicted structure was identified to have a hydroxamic acid an iron chelating function group. This discovery influenced the fermentation of the strain in an iron-deficient media to trigger or enhance the expression (Challis et al., 2005; Gross, 2007).



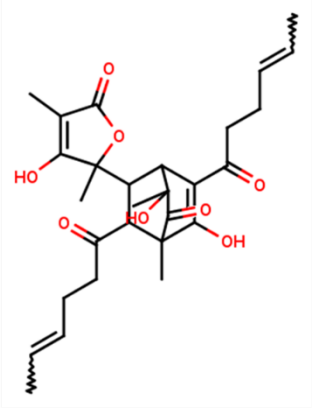
Coelichelin [84]

Genome mining was also utilised by Scherlach and Hertweck (2006) for the discovery of four new prenylated quinolone-2-one alkaloids from *Aspergillus nidulans*. The analysis of the BGC revealed the presence of a protein similar to anthranilate synthase which is involved in the synthesis of anthranilic acid, which is a monomer in alkaloid (quinazoline, quinolone and acridine) and tryptophan biosynthesis. The researchers focused on HR-SM crude sample analysis with a focus on nitrogen-containing compounds (Scherlach and Hertweck, 2009, 2006).

The overexpression of a global fungal regulated gene (*PdLae*A) from the marine fungus *Penicillium dipodomyis* YJ-11 resulted in the discovery of two new compounds (10,11-dihydrobislongiquinolide **[85]** and 10,11,16,17-tetrahydrobislongiquinolide **[86]**) and four known sorbicillin analogues. The regulator was discovered after analysis of the fungus genome sequence (Yu et al., 2019)



10,11-dihydrobislongiquinolide [85]



10,11,16,17-tetrahydrobislongiquinolide [86]

1.7 Dereplication of secondary metabolites

Dereplication is a process whereby known compounds (or BGCs as a proxy for the capacity to produce known compounds) are identified to minimise rediscovery of known products (Nielsen et al., 2011). Dereplication is performed to reduce cost, reduce the time needed for structural elucidation and allow prioritising of promising samples (Zink et al., 2005). It is a process that utilises both the analytical and biological information accompanied by database searches to evaluate the novelty of the sample composition or BGC (Gaudêncio and Pereira, 2015; Weber et al., 2015). Dereplication parameters include data from mass spectrometry, bioassays, nuclear magnetic resonance (NMR) spectroscopy, taxonomic identification, chemical structure, proteomics, genomics, and a combination of these parameters (Gaudêncio and Pereira, 2015). Figure 1.6 demonstrates a typical dereplication workflow.

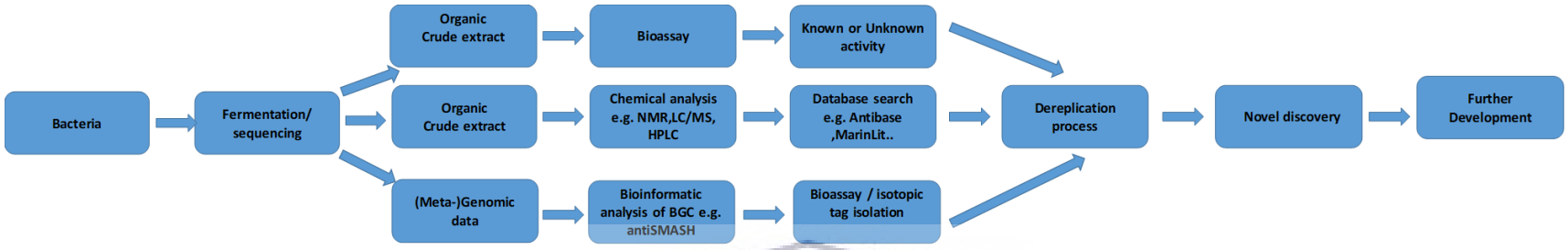
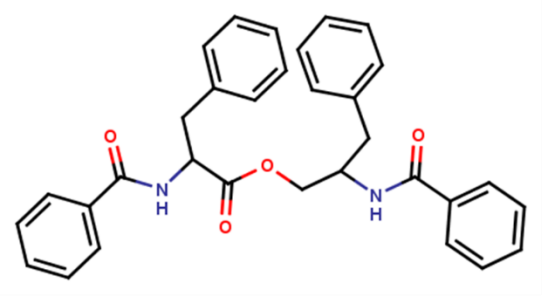


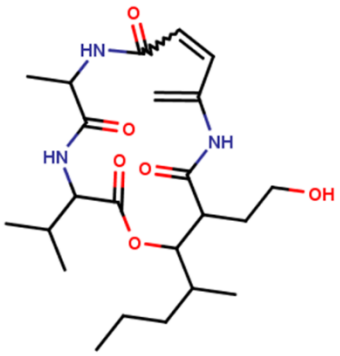
Figure 1.6. Illustration of the bioassay, chemical and genomic guided dereplication process used in secondary metabolite discovery to prioritise promising samples. The image was created using information from Gaudêncio and Pereira, 2015

Analytical dereplication entails subjecting the fermentation broth to a variety of analyses, including high-resolution mass spectrometry (HR-MS), ultraviolet-visible spectrophotometry (UV-VIS), nuclear magnetic resonance (NMR), infra-red (IR) spectroscopy and high pressure / performance liquid chromatography (HPLC). The generated data is subjected to database searches to identify already known compounds. A number of public, in-house and commercial databases are available for dereplication such as MarinLit, Antibase, Dictionary of NPs, StreptomeDB, PubChem, GNPS, Bactibase, SciFinder, NAPROC-13, Norine (Caboche et al., 2007; Gaudêncio and Pereira, 2015; Nielsen et al., 2011).

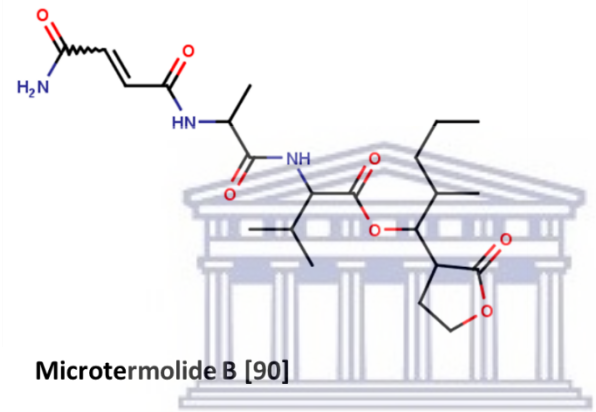
The discovery of four new isomers of a known anti-cancer compound (asperphenamate**[87]**), from *Penicillium bialowiezense*, were facilitated using HRLC-MS data analysis as a dereplication tool (Kildgaard et al., 2014). Hou and co-workers combined liquid chromatography-mass spectrometry LC-MS and principal component analysis (PCA) in the effort to dereplicate and chemically prioritise microbial strains for the production of new metabolites. The study successfully demonstrates that 16S rRNA classification is not sufficient to prioritise species as the same species can produce different SMs (Hou et al., 2012). In a similar study, 30 *Streptomyces* species metabolomes were examined to select unique producers from the same ecological niche using HPLC-MS-PCA. The PCA analysis resulted in *Streptomyces* sp. MspM5 identified as producing different metabolites, which on further probing using Antibase resulted in the isolation of two NRP-PK hybrid metabolites, namely microtermolide A **[88]** and B **[89]** (Carr et al., 2012).



Asperphenamate [87]

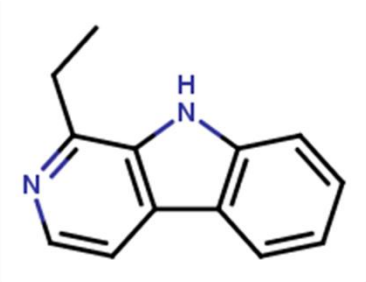


Microtermolide A [89]

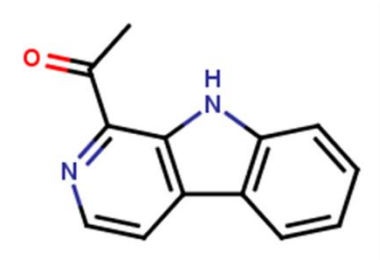


Microtermolide B [90]

Nuclear magnetic resonance (NMR) spectroscopy, an analytical method mainly used for structural elucidation, was utilised in SM chemical profiling for dereplication (Kumar, 2015; Macintyre et al., 2014). The database known as DEREP-NP was developed by Zani and Carroll utilising NMR (^{1}H, HSQC, and/or HMBC) and/or MS (*m*/*z*) data and contains 65 structural fragments derived from 229 358 SMs found in microorganisms, plants and other sources. However, this technique can only be applied to pure compounds or fractionated crude extracts. It successfully identified the presence of 1-ethyl-9H-β-carboline **[91]** and 1-(9H-β-carboline-1-yl) ethanone **[92]** from a bryozoan fraction (Zani and Carroll, 2017).



1-ethyl-9H-β-carboline [91]



1-(9H-β-carboline-1-yl) ethenone [92]

NMR was also utilised by Macintyre and co-workers with a combination of statistics (PCA) to dereplicate bacterial strains based on their crude extract chemical composition. Briefly, 77 bacterial extracts from marine endosymbionts were subjected to HRMS, NMR, taxonomical identification and bioassay experiments. The data generated was used for dereplication purposes with PCA analysis. This resulted in *Bacillus* sp. 44117, *Rhodococcus* sp. ZS402 and *Vibrio splendidus* strain LGP32 being selected for upscaling and further studies as they were chemically distinctive from the other strains (Macintyre et al., 2014).

Genome mining also serves as strain dereplication tools utilising database searches reviewed in section 1.6 (NCBI, NaPDos, antiSMASH, PRISM, Bagel). Dereplication is made possible because known BGCs are linked to known SMs enabling prioritisation of strains with unknown BGCs. Additionally, the database has algorithms to identify unknown BGCs of major classes of SM. However, it is worth noting that this method cannot detect completely new pathways, which is evident with multiple SMs with no known pathways (Weber and Kim, 2016). Despite this limitation, there are still countless known pathways with no known products discovered using genome mining making this method an important part of the drug discovery process (Scherlach and Hertweck, 2009). It is worth noting that, since this approach is *in-silico* based, it is difficult to establish a similarity cut-off for indicating same or different compound production by similar synthases. Therefore, it is not possible confidently speculate the novelty of the compound based on the sequence as highly similar synthases can synthesise / catalyse a formation of same SM and vice versa is also true.

Coupling genomic information with chromatographic and spectroscopic analysis at an early stage enables early identification of both known and unknown SMs in the crude sample. High throughput genome mining of 10000 actinomycetes for phosphoric acid-containing SMs resulted in the discovery of 55 new SMs (Ju et al., 2015). The integration of these techniques can facilitate the development of progressive drug discovery pipelines, thereby eliminating the discovery of known SMs. The analysis of the genomic potential of microorganisms to produce SMs will enable prioritisation of strains.

1.8 The aim of the thesis

The work performed during this study formed part of an international collaborative project known as PharmaSea. The umbrella project focus was on biodiversity research and identification of new lead bioactive compounds from marine microorganisms with the potential of being developed as new therapeutic agents. In this study, a combination of bioassays, chemical and genomic approaches were used as tools for strain prioritisation and dereplication to identify potential new anti-microbial producing marine bacteria using small scale one strain many compounds approach (OSMAC). The marine bacteria studied were isolated from marine invertebrates collected from Algoa (Port Elizabeth) and Kalk Bay (Cape Town), SA.

The main objectives of the study were:

- Bioassay based dereplication of antibiotic producing marine bacteria on liquid and solid media utilising the OSMAC approach
- Whole genome sequencing of selected antibiotic producing bacteria
- Genome mining of selected antibiotic producing marine bacteria for new secondary metabolite biosynthetic gene clusters
- Bioinformatic characterisation of BGCs
- Extraction, purification and characterisation of metabolites from selected marine bacteria

2. Chapter 2: Materials and Methods

2.1. Reagents

Unless stated otherwise, all chemicals and reagents used were either purchased from Sigma Aldrich (St. Louis, Missouri, United States of America), Merck (Kenilworth, New Jersey, United States of America), New England Biolabs (Ipswich, Massachusetts, United States of America), Thermo Fisher Scientific (Waltham, Massachusetts, United States of America), Biolabs and Life Technologies (Waltham, Massachusetts, United States of America) with no further purification. Primers were synthesized by Inqaba Biotech (Johannesburg, South Africa).

2.2. Cultivation of the isolated marine bacteria and test strains

A total of 23 marine invertebrate-associated (MIA) bacteria were selected from a previously established culture collection of over 3000 endosymbiotic bacteria isolated from marine invertebrates and were used in this study. The marine invertebrates were collected from Algoa Bay (coordinate: 33°50′S; 25°50′E), and Kalk Bay (coordinate: 34°07′40″S; 18°26′54″E), along the South African coastline (Appendix V). The MIA bacteria were cultured and screened for their potential to produce bioactive secondary metabolites (SM) under different culturing conditions. The marine bacteria were cultured and maintained on glucose, yeast and malt (GYM) medium for the duration of the study. The media used in this study are outlined in Appendix I. The isolates studied in this study were selected based on initial screening of the isolates using the agar overlay assay, with isolates showing bioactivity being selected for further interrogation.

In this study, a broad spectrum of microorganisms including Gram-negative, Gram-positive and fungi were selected as indicator strains for assessing anti-microbial activity. Prior to testing for anti-microbial activity, bacterial test strains (*Escherichia coli* 1699 - Cubist Pharmaceuticals, *Bacillus cereus* ATCC10702, *Pseudomonas putida* ATCC12633, *Mycobacterium aurum* A^{+}, *Staphylococcus epidermidis* ATCC14990) were cultured in lysogeny broth (LB) at 37 ^{o}C for

approximately 16 h and maintained in lysogeny agar (LA) for the duration of the study and sub-cultured weekly. Fungal test strains (*Aspergillus fumigatus* MRC and *Candidas albicans* NIOH) were cultivated in potato dextrose broth (PDB) at 30 °C for 24 h and maintained on potato dextrose agar (PDA).

2.3. Anti-microbial and anti-inflammatory activity screening

All the bioassays used in this study are phenotypic based assays. Three methods were employed for bioactivity screening, which includes i) well diffusion, ii) soft agar overlay and iii) high throughput screening (HTS) for anti-bacterial, anti-fungal, anti-yeast, and anti-inflammatory activity. The HTS was performed by our collaborators (Medina; Granada, Spain and MarBio; The Arctic University of Norway, Norway). MarBio is an analytical platform specialising in screening, isolation and identification of secondary metabolites mainly from the Arctic marine environment. The platform relies on a highly automated HTS for several bioactivities such as anti-bacterial, anti-cancer, anti-inflammatory, anti-diabetic, ant-oxidant and anti-immunostimulant activity. Medina is a microbial drug discovery platform utilising HTS system similar to MarBio. However, Medina's focus is on oncology, neurodegeneration, infectious- and parasitic diseases. Both platforms screen for bioactive SMs using clinical isolates.

2.3.1. Well-diffusion assays

A well-diffusion assay is an anti-microbial susceptibility test, whereby wells are punched into solid media to act as a chamber for the tested extracts. Prior to the bioassay, the bacterial test strains were cultured as stated in Section 2.2. The 16 h old cultures were diluted to achieve a cell suspension of approximately 1.5 x 10^8 cells / ml (0.5 McFarland standard) before the well-diffusion assay. This step was performed to standardise the inoculum for each bioassay performed. This was achieved by diluting the culture to an optical density (OD_{600nm}) of between 0.08-0.1. The fungal test strains were standardised by adjusting the fungal suspension to an absorbance of approximately 0.5 at 600 nm.

MIA bacteria were cultured in a 10 ml GYM medium and incubated at room temperature (± 25 °C) with shaking at 150 rpm for 2 days to be used as inoculum for the fermentation process. The cultures were then used to inoculate 10 ml of Zobell broth (ZBB), trypticase soy broth (TSB), activated charcoal media (ACM) and GYM media in McCartney bottles and incubated for 14 days under the same incubation conditions as for inoculum preparation. After 14 days, the cultures were centrifuged at 6000 x *g* for 10 min to remove cells and the supernatant was added into a sterile container. The pelleted cells were lysed through the addition of MeOH and shaken for 30 min. After 30 min, the mixture was centrifuged to remove cell debris and the supernatant pooled with fermentation broth. The fermentation broth was concentrated to half the original volume (5 ml) using a CentriVap centrifugal vacuum concentrator at 40 °C (Labconco, United States of America).

100 µl of each test strain were spread onto LA in a 90 mm petri dish and after spreading, the plates were partially dried under sterile conditions (laminar flow). Thereafter, wells were made in the agar, using a sterile micropipette tip (1000 µl) and 100 µl of the concentrated fermentation supernatant was added into each well. The plates were then left under the laminar flow hood to allow for pre diffusion of the supernatant into the agar. After 2 h, the plates were incubated at 37 ºC for 16 h. Zones of inhibition were observed as the absence of bacterial growth around the wells and results were recorded.

Anti-fungal bioassays with *A. fumigatus* and *C. albicans* were conducted similarly to the bacterial strains. After 48 h incubation of the fungal test strain in PDB, each fungal strain was adjusted to an OD_{600nm} of 0.5 and spread-plated on PDA containing petri dishes and partially dried in a level two biosafety cabinet. Wells were made on the sterile agar and 100µl of the fermentation broth was used for the well-diffusion assay. The plates were incubated for 48 h at 30 °C. Anti-fungal activity was observed as a clear zone around the well.

2.3.2. Agar overlay assays

The agar overlay assay is an anti-microbial susceptibility assay which requires the application of a homogeneous bacterial lawn over an established bacterial colony on solid medium. This method was selected to evaluate the ability of the MIA bacteria to produce bioactive SM under solid

fermentation conditions. The MIA bacterial isolates were spotted on agar plates (ZB, TSB, ACM and GYM), and incubated for 14 days at room temperature inside sterile plastic bags that were not fully sealed, to permit airflow. The test strains were cultured as described in Section 2.2 and the inoculum for overlay assays were standardised using the following equation:

OD_{600} × Xµl = 4 for *E. coli* 1699

OD_{600} × Xµl = 160 for the other test strains.

Using the above equation, the calculated volume of the test strain was mixed with 6 ml soft agar and poured over the marine bacteria spotted agar plates. The plates were left to stand in the laminar flow for 30 min to allow complete solidification of the soft agar, followed by incubation at 37 ºC for 16 h. Clear zones around the spotted colonies when observed was scored as anti-microbial activity. The anti-fungal assay was not conducted using this assay.

2.3.3. High-throughput screening assays by PharmaSea partners

For anti-microbial (Martín et al., 2013; Monteiro et al., 2012) and anti-inflammatory activity (Svenson, 2013) screening, semi-fractionated extracts were used. Marine isolates were cultured in 10ml TSA, ZB, GYM and ACM media for 14 days as described in Section 2.2. The extracts were prepared through acetone extraction, where an equal volume of acetone was added to the marine bacterial culture and incubated on an orbital shaker for 1 h. The acetone was evaporated in a fume hood before the crude extract was applied to fractionation columns. Fractionation columns were prepared by puncturing holes into the bottom of a 15 ml falcon tube and then covering it with glass wool. The tube was filled with 10 ml of SP207ss resin (Sigma-Aldrich, United States of America). The extracts were prepared in our laboratory and shipped to the collaborators as dry extracts.

The columns were equilibrated by washing with 50 ml of acetone, followed by a 50 ml MeOH wash and a final wash with 100 ml reverse osmosis (RO) water. After crude extracts were added to the columns, the columns were washed with 50 ml RO water. Upon eluting the organic extract with 50 ml acetone, the eluate was collected, and the acetone evaporated in a CentriVap centrifugal vacuum concentrator. The dried extract was re-suspended in distilled water (dH_2O) and screened

against *Aspergillus fumigatus* ATCC 46645 (Monteiro et al., 2012), MRSA-MB5393 and *Candida albicans* MY155 (Martín et al., 2013) using the HTS system. Extracts were also assessed for anti-inflammatory using tumour necrosis factor-*α* (TNF-α) as an indicator (Svenson, 2013). Anti-inflammatory assay activity was measured using techniques such as enzyme-linked immunosorbent assay to quantify the presence of pro-inflammatory mediators, such as TNF-α, COX-2, IL-6 and IL-1β, from a cell line. Fractions with above 50 % inhibitions of TNF-α detection were defined as active based on the MarBio standards.

2.4. Molecular biology techniques

This section of the thesis covers genomic work conducted in the study which includes genomic DNA extraction, sequencing, DNA analysis and genome mining of SM BGCs.

2.4.1. Genomic DNA extraction

Genomic DNA (gDNA) was extracted from MIA bacteria which were cultured in 50 ml GYM medium for 2 days at room temperature (± 25 ºC) with shaking at 150 rpm. The cells were centrifuged at 3200 × *g* for 10 min at 25 ºC. After centrifugation, the supernatant was discarded and the pellet was re-suspended in 9.5 ml Tris-(hydroxymethyl)-aminomethane-ethylenediaminetetraacetic acid (TE buffer), 0.5 ml of 10 % sodium dodecyl sulfate (SDS), 50 µl of 20 mg/ml proteinase K, and 1 µl RNAse (10 mg/mL), then mixed by inverting the tube a few times. The mixture was incubated for 1 h at 37 ºC and 1.8 ml of 5 M sodium chloride (NaCl) was added and mixed. Thereafter, 1.5 ml of 10 % cetyltrimethylammonium bromide (CTAB) / 0.7 M NaCl was added to the mixture, mixed, and then incubated at 65 ºC for a further 20 min.

After the incubation at 65 ºC, equal volumes of chloroform / isoamyl alcohol (CI) at a ratio of 24:1 was added to the mixture and then further mixed by inverting the tube several times. The tube containing the mixture was then centrifuged at 16000 × *g* for 10 min at 25 ºC for separation of the two liquid phases. The aqueous (upper) phase was transferred into a new 50 ml falcon tube and

equal volumes of phenol / chloroform / isoamyl alcohol (PCI) was added. Thereafter the tube was centrifuged at 16000 × *g* and this step was repeated twice. The aqueous phase was transferred into a clean 50 ml falcon tube. Thereafter, 0.6 volumes of 100 % isopropanol was added to the aqueous solution to precipitate the DNA. Using a sterile, bent glass pasture pipette, the precipitated DNA was hooked and transferred to 1 ml of 70 % ethanol and centrifuged at 16000 × *g* for 5 min. The supernatant was decanted, and the pellet air-dried. The air-dried pellet containing the gDNA was dissolved in 4 ml TE buffer and left to stand at room temperature until the solution turned clear indicating the complete dissolution of the DNA pellet.

2.4.2. Nucleic acid quantification and storage

The concentration and purity of genomic DNA was determined using the Nanodrop ® ND 1000 (Nanodrop Technologies, Inc., United States of America (USA)) and Qubit™ fluorometer (Invitrogen, USA). Extracted DNA was aliquoted and stored at 4 °C, then at -20 °C and lastly frozen at -80 °C until further use.

2.4.3. Gel electrophoresis

Gel electrophoresis was used to separate, analyse and evaluate genomic DNA (gDNA) and plasmid DNA integrity and PCR amplicons. Agarose gels between 0.5-1.5 % were prepared depending on the application. Briefly, agarose was prepared by weighing the required amount of agarose powder and dissolved in 1 × Tris-acetate-EDTA (TAE) buffer. The mixture is then boiled in a microwave at high heat until the agarose has completely dissolved. Once the agarose had cooled, it was mixed with ethidium bromide (0.5 µg/ml) and poured into a gel casting tray. The samples were mixed with 3 µl 10 × loading dye before they were loaded onto the gel.

The gel was electrophoresed in an electrophoresis chamber filled with 1 × TAE buffer at 100 V for 1 h. The size of the DNA was estimated using either PstI digested λ DNA or 1 Kb DNA ladder (New England Biolabs, USA). DNA was visualised using an Alpha Imager® HP 2000 (Alpha

Innotech, USA) under ultraviolet (UV) light at a wavelength of 302 nm. To prevent DNA damage, the duration of DNA exposure to UV light was kept to a minimum. DNA to be excised for cloning was visualized for approximately 4 s at 360 nm.

2.4.4. Molecular identification of MIA bacterial isolates

Isolated MIA bacteria were identified using PCR amplification and sequencing of the 16S rRNA gene using universal primers E9F (GAGTTTGATCCTGGCTCAG (Farrelly et al., 1995) and U1510R (GGTTACCTTGTTGTTACACTT) (Reysenbach et al., 1992). The PCR amplification conditions were: 95 ºC for 3 min for one cycle, then 30 cycles of 95 ºC for 30 s, 55 ºC for 35 s, 72 ºC for 1 minute, and a final extension for 15 min at 72 ºC using DreamTaq polymerase (Thermo Scientific, USA) in a total reaction volume of 50 µl. The expected 1.5 kb amplicon was visualised and then purified with a NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany). Purified DNA products were sequenced by the Central Analytical Facility (CAF) at Stellenbosch University using an ABI PRISM 377 automated sequencer. The sequences were processed using Chromas lite / pro version 2.6.5 software (Technelysium Pty Ltd, Australia) for alignment and manual editing of sequences, after which the sequences were compared with those available on the GenBank database using the NCBI BLASTn tool. All NCBI BLAST analyses were performed with NCBI default settings.

2.4.5. Genome sequencing, assembly and annotation

Those MIA bacteria that showed desirable anti-microbial profiles were selected for whole genome sequencing. The genomes were sequenced at the University of the Western Cape's sequencing unit, using the Illumina MiSeq sequencing platform. Sequencing libraries were prepared using one nanogram (1ng) of extracted gDNA using the Nextera XT kit and sequencing was performed using a MiSeq 500 cycle v2 kit resulting in paired (forward and reverse) 250 bp reads. Sequence data were assembled using CLC Genomics Workbench 6.5, after which the reads were merged and mapped to the phiX174 genome prior to *de novo* assembly to remove contaminating sequence as

it was used as a positive control for the sequence run. After removal of reads mapping to phiX174, reads were trimmed using the following settings: quality score 0.05, maximum 3 ambiguous bases and sequencing adapters and short reads (50 bp) were removed. *De novo* assembly was performed using the following parameters namely mismatch cost 2, insertion cost 3, deletion cost 3, length fraction 0.90, similarity fraction 0.90 and a minimum contig length of 1 kb.

Following *de novo* assembly the draft genome contigs were uploaded to the antiSMASH version 3-4 webserver (Blin et al., 2019; Weber et al., 2015), Bagel 3 (van Heel et al., 2013), and PRISM 3 (Skinnider et al., 2017, 2015) for *in silico* mining of secondary metabolite biosynthesis gene clusters (BGC) using the default settings. The similarity to characterised clusters was identified from minimum information about a biosynthetic gene cluster (MIBiG) database using antiSMASH built-in function. The identified NRP pathways were further characterised for substrate specificity (A-domain) using NRPSpredictor (Rottig et al., 2011). The C-domains of the identified NRPSs were analysed for novelty using Natural Product Domain Seeker (NaPDos) (Ziemert et al., 2012). Average nucleotide identity of the sequenced genomes was measured against 14 publicly available genomes using JSpecies v1.2.1 with default parameters (Richter and Rosselló-Móra, 2009). The genomes were selected from the first 14 16S rRNA BLASTn results with available genome data.

The PE14-07 genome was also mined for tropodithietic acid (TDA) biosynthesis genes as it is one of the common and characterised (chemically and biologically) SM in the *Roseobacter* family. The 6 genes utilised in the biosynthesis of TDA from marine *Pseudovibrio* sp. P12 (Raina et al., 2016) were used as a query against the draft genome of *Pseudovibrio* sp. PE07-14 isolate with NCBI BLAST + (Camacho et al., 2009). Sequences used for the identification of the TDA genes within PE14-07 are reported in Appendix VII. The results were recorded and reported in Table format.

2.5.Secondary metabolite extraction

2.5.1. Inoculum preparation

To purify and concentrate secondary metabolites produced by MIA bacteria PE14-07 and PE08-149B, the following procedures were used. For large scale fermentations, a 50 ml inoculum culture was prepared for 1-2 days in GYM and used to inoculate large scale GYM fermentation (25 % volume) in multiple 2 L flasks.

2.5.2. Activation and regeneration of resins (amberlite IR 120 and IRA 910)

Amberlite IR 120 resins were activated by shaking in 200 ml of 15 % hydrochloric acid (HCl) for 30 min at room temperature. The resins were washed twice with 200 ml RO water and stored in 100-150 ml RO water until further use. The amberlite IRA 910 resins were activated similar to the amberlite IR120 resins, but with the use of 0.1 M sodium hydroxide (NaOH), instead of HCl. Both resins were regenerated using the same protocol to activate them and were reused more than once throughout the study. However, the resins were only reused with extracts from the same strains to prevent possible cross-contamination.

2.5.3. Extraction of SMs

The compounds were extracted using amberlite IR 120, amberlite IRA 910 and EtOAc. Marine isolates were fermented for 14 days before extractions of SMs with both amberlite IRA 910, amberlite IR120 resins and EtOAc. Figure 2.1. illustrates the extraction workflow.

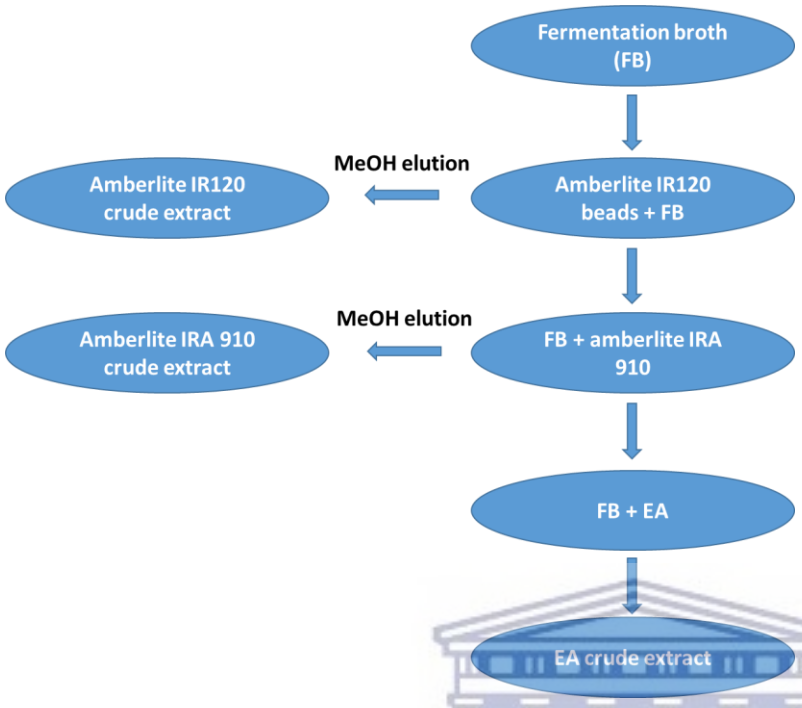


Figure 2. 1. Chronological illustration of the extraction process followed.

Amberlite IRA 120 resins were added to a final concentration of 5 % (w/v). To obtain both intra- and extracellular metabolites from the fermentation, cells were collected by centrifugation at 6000 × *g* for 15 min at 25 °C and the cell pellets were treated with 50 ml MeOH to lyse the cells, thereby releasing the intracellular metabolites. The mixture was shaken at room temperature for an h and the debris separated by centrifugation at 6000 × *g* for 10 min or by vacuum filtration using a sterilised filtration system. The intracellular extracted metabolite and fermentation broth were pooled together before extraction with the resins.

Amberlite IR 120 resins were added to the beaker containing the fermentation broth and shaken at room temperature for 2 h, to allow enough contact between the SMs and the resins. After 2 h, the broth was decanted, and the resins were washed twice with 200 ml RO water. Thereafter, 500 ml of MeOH was added to the resins to elute the SMs from the resins. The MeOH extract containing natural products of interest were separated from the amberlite IR 120 resin using a 0.45 µm filter,

under vacuum or a separating funnel as illustrated in Figure 2.2. The MeOH was evaporated using a rotatory evaporator at 40 ^{o}C.

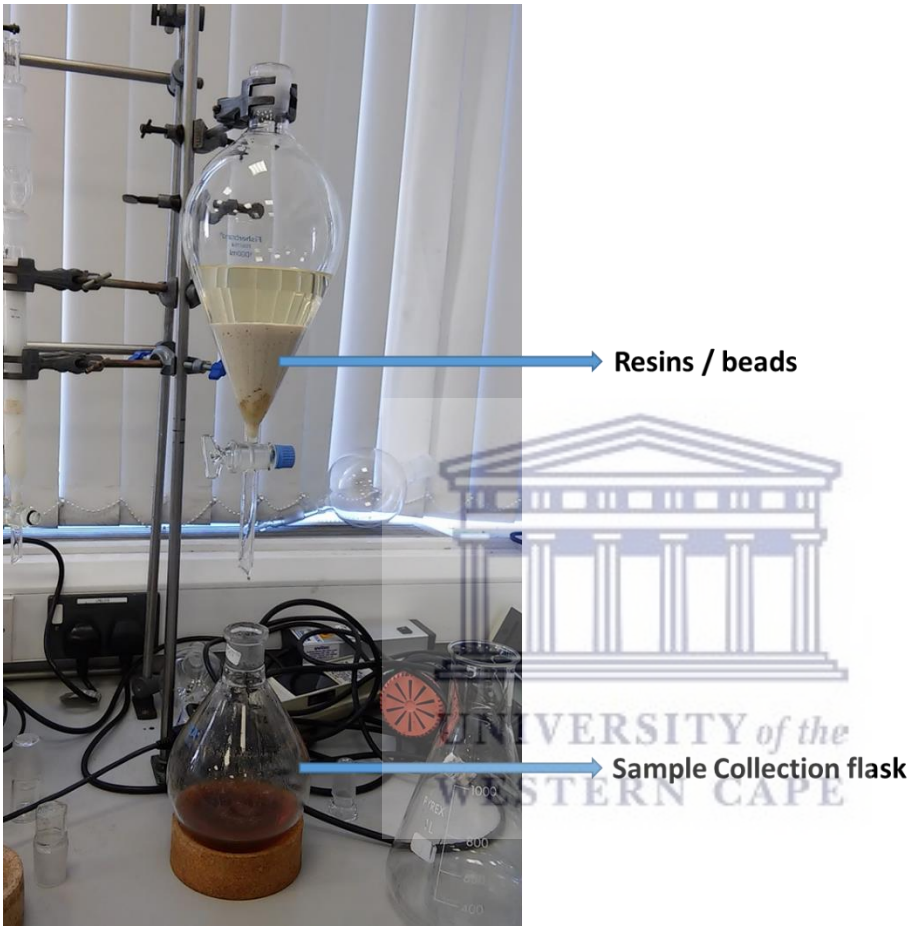


Figure 2. 2. Illustration of how MeOH containing SMs was separated from the resins.

Extraction of SM with amberlite IRA 910 resins was carried out in the same manner as amberlite IR 120. An equal volume of EtOAc was added into the fermentation broth. The mixture then was shaken for 2 h at 150 rpm. The mixture was then left to settle and separated using a separation funnel. The EtOAc was then evaporated using a rotary evaporator at 40 ^{o}C. After extraction anti-microbial assay was conducted using a well-diffusion assay as explained in Section 2.3.1. Crude extracts showing high levels of anti-microbial diversity or interesting chemistry were selected for further purification. Prior to the anti-microbial assay, the crude extracts were either dissolved in dH_2O or 80 % dimethyl sulfoxide (DMSO).

2.5.4. Purification of natural products

The pre-fractionated crude extracts were further purified to obtain pure compounds for compound characterisation and structural elucidation.

2.5.4.1. Solid-phase extraction

The extracted crude material from amberlite IR120 resin and EtOAc extracts were subjected to reverse phase-solid phase extractions (RP-SPE). The setup of the RP-SPE system is shown in Figure 2.3. A 5 g RP-SPE cartridge (Sigma-Aldrich, USA) packed by the manufacturer was conditioned with 4 column volumes (CV) of 100 % MeOH. The column was then washed with 4 CV of RO water to remove all the MeOH and 500 mg of the sample was dissolved in 100 % RO water before being loaded onto the column.

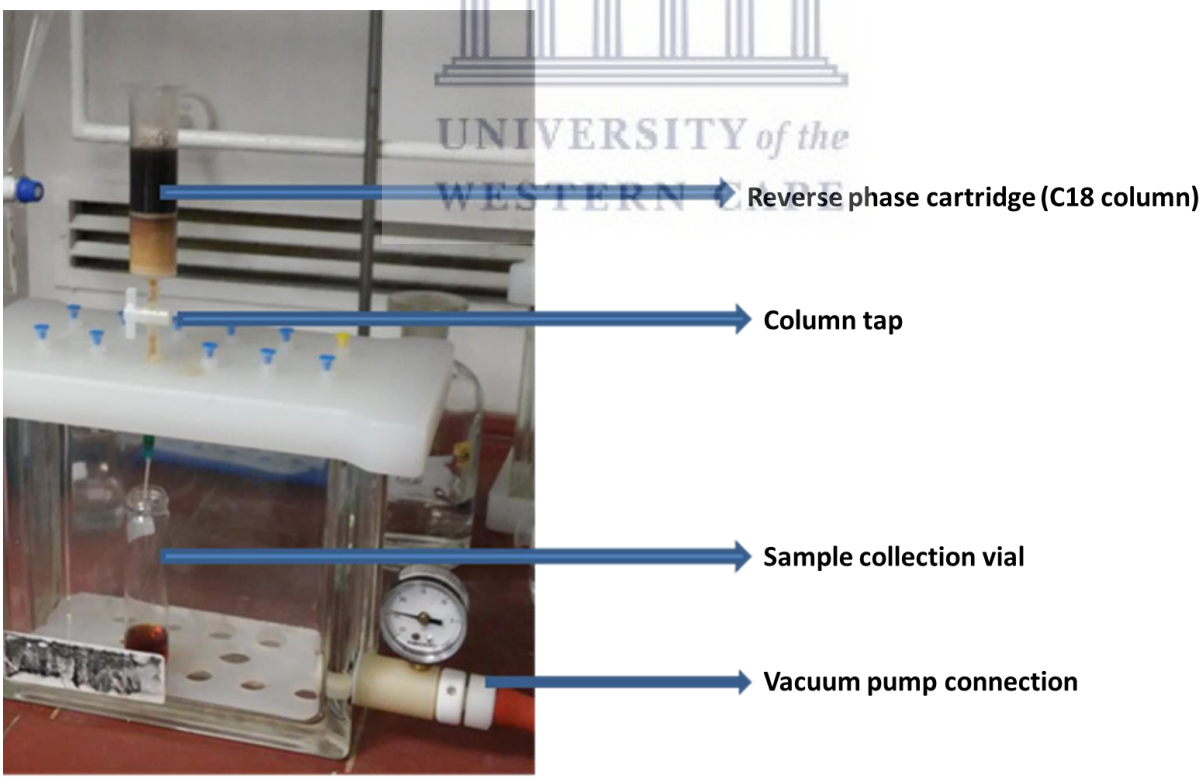


Figure 2. 3. A typical cartridge solid phase extraction system

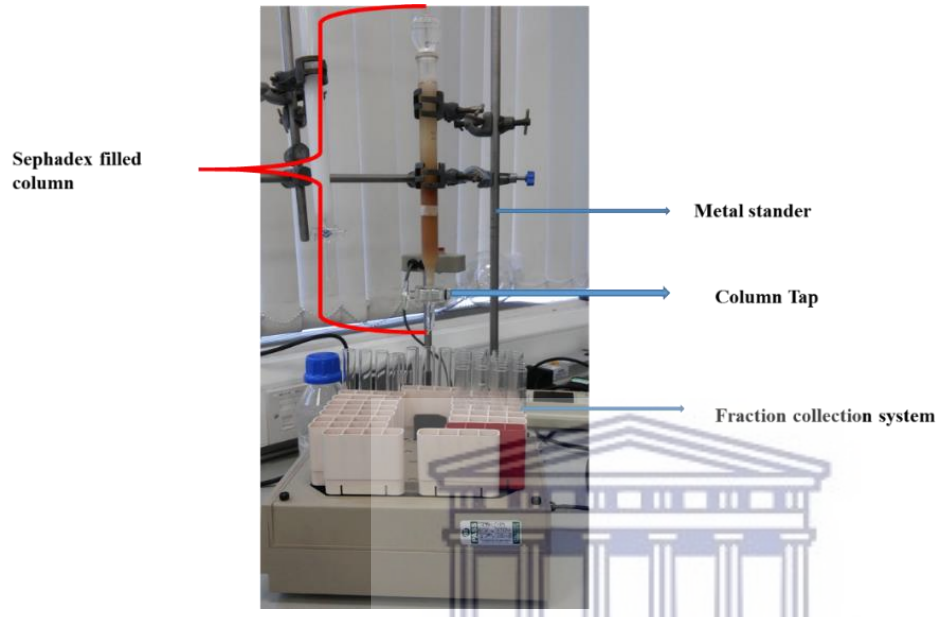
The loaded compounds were eluted from the column using 4 CV of RO water and collected as a wash step, to remove highly polar and unwanted media components, such as salts and sugars. The undissolved sample debris from the sample vial was re-dissolved in 25% MeOH and loaded into the column. Thereafter, 4 CV of 25 % MeOH was used for column elution and the fraction was collected as fraction 1 (e.g., coded: PVET-SPE-25). The remaining sample debris was dissolved in 50 % MeOH and loaded into the column. The column was eluted with 4 CV of 50 % MeOH and collected as fraction 2, (e.g., coded: PVET-SPE-50). The above process was repeated for 100 % MeOH and 100 % MeOH with 0.005 % trifluoroacetic acid (TFA), coded PVET-SPE-100 and PVET-SPE-100TFA, respectively.

The addition of TFA in the MeOH assists in the elution of strongly bound organic compounds. The fractions were partially dried using a rotary evaporator at 40 ^{o}C before they were transferred to pre-weighed vials for overnight drying under nitrogen gas on a 40 ^{o}C heating block. Once completely dried, the samples were weighed, and weights recorded. The fractions were analysed using ^{1}H nuclear magnetic resonance (NMR) spectroscopy or bioassay, before further purification. NMR was used to evaluate interesting chemical functional groups, splitting patterns and chemical shifts and bioassays were used to evaluate the anti-microbial activity of the extracts or fractions.

2.5.4.2. Size exclusion chromatography

After chemical analysis of the generated fractions with SPE, the samples were further purified by size exclusion chromatography (Figure 2.4) using a 50:50 ratio of MeOH / dichloromethane (DCM) on a Sephadex LH-20 (20-100 µm) mesh (Sigma-Aldrich, USA). The sample (PVET-SPE-100 %) generated from SPE in section 2.6.3.1 was dissolved in 100 % MeOH and then loaded onto the Sephadex mesh column. A total of 17 fractions were collected, followed by 10 min eluting at 100 % MeOH. Fractions 1-7, 8-10, 11-13 and 14-17 were pooled together to generate a total of 5 fractions, including the 100 % MeOH final elution step. Fractions with similar colours were pooled together as advised (personal communication). The fractions were dried overnight under nitrogen

gas at 40 °C, then subjected to NMR before further purification with high-performance liquid chromatography (HPLC).



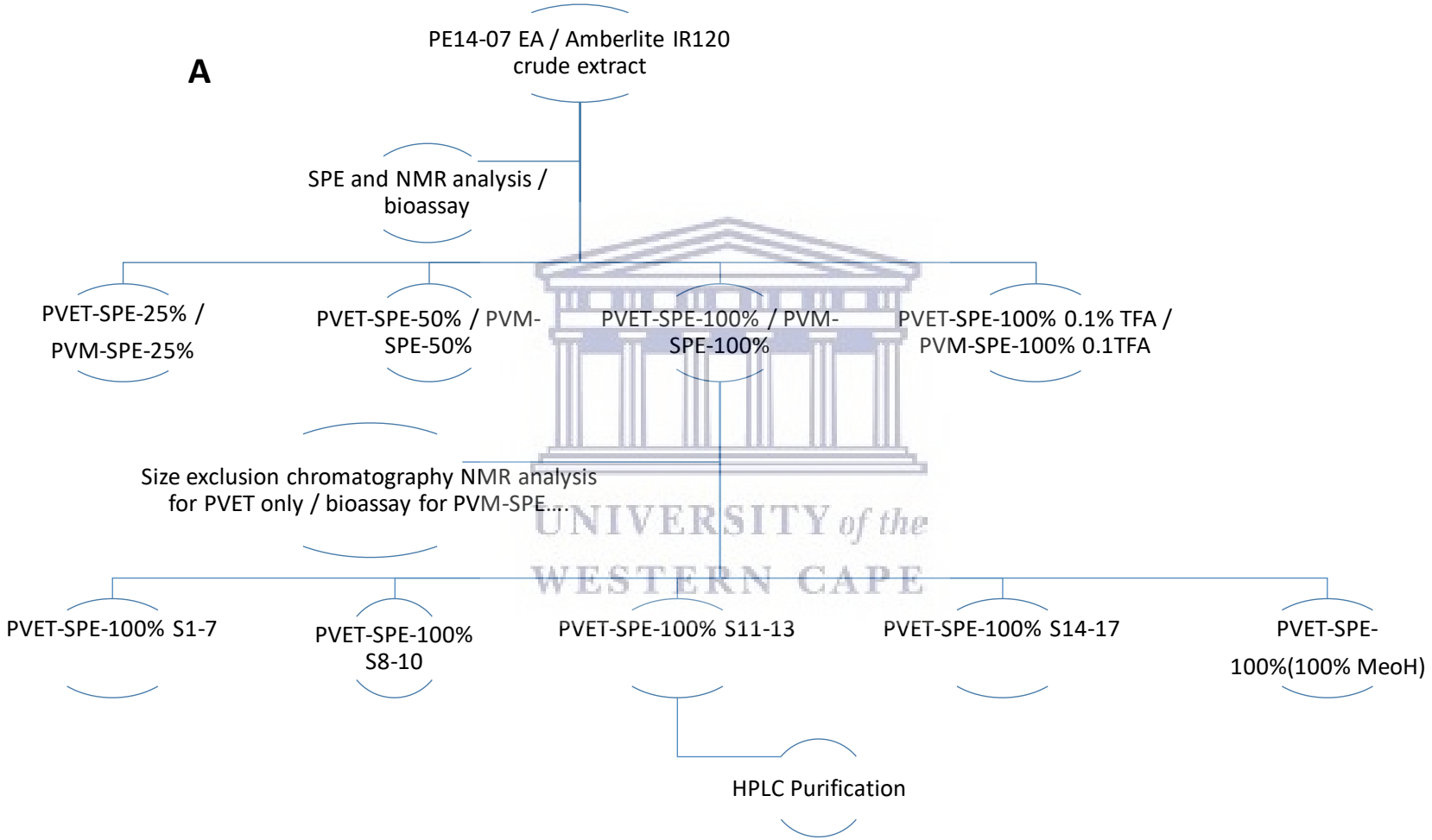
.

Figure 2. 4. Size exclusion set-up used during purification

2.6.3.3 Liquid-liquid partitioning

To concentrate compounds of similar polarity in the fractions, crude extracts were subjected to liquid-liquid partitioning. Samples were partitioned with solvents of increasing polarity, namely hexane, DCM, 1-butanol (BUT), MeOH, and RO water. The crude sample was re-suspended in RO water and mixed with DCM in a 1:1 ratio and the solvents were separated using a separating funnel. The crude sample was extracted three times with DCM. The RO water fraction was extracted two times with 1-butanol using equal volumes and dried with a rotary evaporator at 40 °C. The DCM fraction was dried using the rotary evaporator and dissolved using 70 ml of 90 % MeOH. The dissolved fraction was extracted with equal volumes of hexane three times and the 90 % MeOH fraction adjusted to 50 % MeOH using RO water. The 50 % MeOH fraction was re-extracted with DCM and dried with a rotary evaporator at 40 °C. A total of four fractions were generated (MeOH / water, BUT, DCM and hexane). The samples were analysed for chemically

interesting features using NMR or bioassays to select fractions for HPLC purification. Figure 2.5 shows a graphic representation of the fractionation workflow followed during the study and fractions generated, starting from the crude extracts of PE14-07 and PE08-149B bacteria.



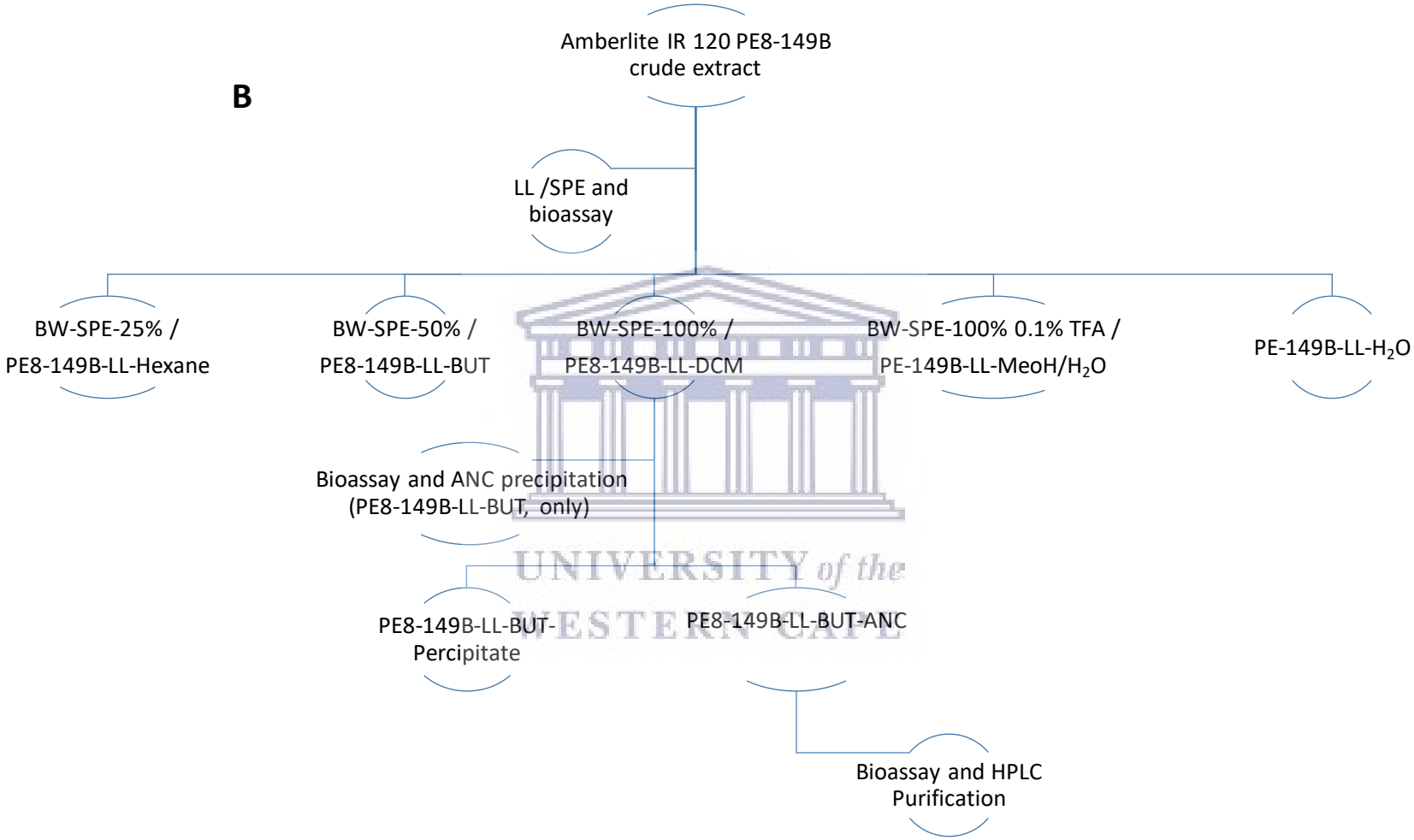


Figure 2. 5. Graphical illustration of the extraction process. A; PE14-07 and B; PE8-149B

2.6.3.4 Purification of natural products

After NMR and bioassay experiments, fractions were subjected to HPLC fractionation using an ACE C18 (Advanced Chromatography Technologies, Scotland) semi-prep column (250 mm × 4.6 mm**)** or Phenomenex Kinetex semi-prep C18 5 µm (250 mm × 10 mm) column, with a sample injection volume of between 60 – 200 µl. HPLC was performed using a gradient, consisting of water / MeOH (95 % / 5 %) and MeOH (100 %). The 95 % ultra-pure water solution was used to prevent the column from collapsing in 100 % water. Gradient conditions used in the study are outlined in Table 2.1. All solvents used were HPLC-grade. HPLC generated fractions were evaluated for anti-microbial activity and analysed with NMR. The active fractions were collected from multiple injections to generate enough material for compound characterisation.

Table 2. 1. HPLC conditions used for purification of secondary metabolites in this study

Sample ID (fraction)	Flow rate (ml/minute)	Column	Solvents	Gradient
PVET-SPE-100 S11-14	2	ACE C18	A: 95:5 (water: MeOH) B: 100 % MeOH	Solvent B: 60 % for 0 to 20 min, followed by 100 % B for 10 min and then 0 % B for 5 min for a total run time of 35 min
PVET-SPE-100	1.5	Phenomenex Kinetex semi-prep C18	A: 95:5 (water: MeOH) B: 100 % MeOH	Solvent B: 60 % - 95 % in 15 min, 95-100 5 min and a 10 min hold at 100 % for a total run time of 30 min
LL-BUT-ANC-149B	1	Phenomenex Kinetex semi-prep C18	A: 95:5 (water: MeOH) 0.01 % TFA B: MeOH 0.01 %TFA	Solvent B: 50 %-95 % for 0-15 min, followed 95 % -100 % in 5 min and then a 10 min hold for a total run time of 30 min
Tropodithietic acid (TDA)	1	Phenomenex Kinetex semi-prep C18	A: 0.1 % TFA water B: 0.1 % acetonitrile	Solvent B: 20 % to 100 % in 25 min, followed by a 5 min hold at 100 % for a total run time of 30 min

The generated fractions were dried by nitrogen gas on a 40 ^{o}C heating block or in an incubator and thereafter analysed by NMR spectroscopy for the following structural elucidation experiments. The HPLC systems used for the study were an Agilent 1260 infinity (Agilent Technologies, USA) and Dionex Ultimate 3000 (Thermo Scientific, United States of America), both equipped with diode array detectors (DAD) for UV / Vis detection.

2.6.3.5 Analysis, identification and characterisation of the HPLC purified samples

Nuclear magnetic resonance (NMR) spectroscopy is an analytical technique which makes use of the resonance properties of nuclei such as ^{13}C, ^{1}H, ^{31}P and ^{15}N when subjected to strong magnetic fields to elucidate the structure of organic compounds. The HPLC purified samples were analysed using NMR spectroscopy (Bruker Ascend^{TM} 400MHz and 600MHz, USA) with both ^{1}H and ^{13}C chemical shifts (δ ppm) on the spectrum referenced using the NMR solvent peak (MeOH-d_4: δ_H 3.31, δ_C 49.15 or DMSO d_6: δ_H 2.5 and δ_C 39.52). Basic 1D ^{1}H and carbon NMR experiments were conducted to obtain information about the chemical environment of these elements in the organic compounds. More advanced 2D NMR experiments such as correlation spectroscopy (COSY), heteronuclear single-quantum correlation spectroscopy (HSQC) and heteronuclear multiple bond correlation spectroscopy (HMBC) were used to characterise pure compound(s) and aid in structural elucidation. These experiments yield information on proton-proton (COSY) or carbon to proton (HSQC and HMBC) correlation for a compound and allow for establishing connections of the identified groups into a finished organic structure. NMR experiments were conducted at the University of the Western Cape, chemistry department NMR unit and the University of Aberdeen, Marine Biodiscovery centre.

High-resolution mass spectrometry (HRMS) data were acquired for use in the generation of the molecular formula of the purified compounds to assist in structural elucidation. The HRMS detector is coupled to an ultra-HPLC (UHPLC) for chromatographic separation before mass analysis. Samples were analysed using a ThermoFisher Scientific Orbitrap HRMS system (USA)

at the University of Aberdeen, Scotland and Waters Synapt G2 UHRMS (United Kingdom) at the CAF, Stellenbosch University, South Africa. Both systems make use of the soft electrospray ionisation (ESI) technique to produce the ions, meaning that the technique generates fewer fragments. The ionisation was performed either in positive or negative mode, respectively generating the quasimolecular ion as $[M+H^+]$ together with other adducts (e.g. $[M+Na^+]$) for positive mode and $[M-H^-]$ for negative mode.

Xcalibur software (Thermo Scientific, USA) was used to generate molecular formulas. The molecular formula is generated using the mass to charge ratio (*m/z*) of the ions e.g. $[M+H]^+$ which is equal to the exact mass of the protonated compound. The obtained molecular formulas were corrected by the addition of hydrogen (or sodium, if applicable) or removal of the hydrogen for positive or negative mode, respectively. Usually, the molecule should have a double bond equivalent (DBE) with 0.5 as a decimal. The allowed mass error of the generated molecular formula was 5ppm. The mass error refers to the acceptable difference between theoretical *m/z* and experimentally observed *m/z*. And DBE refers to the number of hydrogen molecules required to convert all Π-bond to single bonds and all rings within a compound into an acyclic structure. The generated molecular formula and MS data were also used for dereplication purposes by checking for potential matching compounds using ChemSpider, PubChem, MarinLit and SciFinder.

The solvents used as mobile phase were ultra-pure water and acetonitrile as solvent A and solvent B, respectively, with 0.1 % formic acid. The following gradient conditions used for solvent B were as follows; 0 % for 30 s, followed by a gradient to 100 % from 30 s to 8 min and then holding it for 2 min (8-10 min) before dropping it down to 0 % and holding for 2 min (10-12 min) at a 0.32 ml / min flow rate.

2.5.5. Tropodithietic acid (TDA) extraction and identifications

PE14-07 was grown in 50 ml marine broth and GYM. The experiment was done to evaluate the ability of GYM to trigger the production of TDA. The marine broth was used as it is shown to

trigger TDA production in *Pseudovibrio* genus (Harrington et al., 2014). The cultures were incubated for 5 days in the dark without shaking. TDA was extracted using a method by Prosby et al., 2011. The cultures were extracted with 50 ml of 1 % formic acid in EtOAc twice. The extract was dried at 40 ºC. The dried TDA rich extract was subjected to an anti-microbial test and high-performance liquid chromatography (HPLC)-ultraviolet (UV-Vis) analysis for characterisation of TDA (method in Table 2.1). TDA presence was also investigated with the use of HRLCMS. The separation was performed with a 2.1 mm ID, 50 mm, 1.8 µm Agilent Eclipse Plus C18 RP-column at 40 °C with water and acetonitrile with formic acid as mobile phase. A gradient solvent system was used with the following conditions: 0.8 ml/min flow rate. The gradient starts at 85 % water and decrease to 40 % within 1.8 min and then to 0 % in 0.2 min and keep at 0 % for 0.8 min and return to 85 % in 0.2 min.

3. Chapter 3: Microbial strain prioritisation

Bacteria from marine environments have proven to be a great source of new bioactive compounds, which hold therapeutic value as potential drug candidates (Malve, 2016). These compounds are essential in addressing the current anti-microbial resistance crisis and the lack of therapeutic compounds for certain diseases and also a provision of alternative drugs with improved efficacy or reduced side effects from current medication (Debbab et al., 2010; Paul and Sinha, 2016). Marine invertebrates have contributed to the current drugs which are available in the market such as ziconotide, cephalosporin, vidarabine and trabectedin (Malve, 2016) while several marine bacterial compounds are in clinical trials, such as brentuximab vedotin, soblidotin, tasidotin and salinosporamide A (Malve, 2016; Pereira, 2019). Despite the successful bioprospecting of marine bacteria-derived drugs, microorganisms from the marine environment remain an under-explored niche as a source of bioactive SMs. Of the characterized SMs from the marine environment, 25 % are of bacterial origin, with the number expected to rise significantly (Li et al., 2015).

To further explore marine bacteria as a source of bioactive compounds, marine invertebrate-associated (MIA) bacteria previously isolated from ten marine invertebrates namely, *Distaplia skoogi* (Ascidians), *Spongia sp.*001RSASPN (Porifera), *Pseudodistoma africanum* (Ascidians), *Eudistoma* sp. 011RSASPN (Ascidians), *Halianthella annularis* (Cnidaria), *Sabellastarte longa* / *Pseudobranchiomma longa* (Kinberg, 1866) (Annelida), *Leptophytum foveatum* (Rhodophyta), *Hypnea spicifera* (Rhodophyta) and *Hamacantha (Vomerula) esperioides* (Porifera). These marine invertebrates were collected along the coast of South Africa (Chapter 2). Porifera are a well-known source of bioactive compounds, with this phylum contributing up to 32 % (4851 out of ~15000) of the known marine SMs (Mehbub *et al*., 2014). A wide variety of bacteria isolated from these marine invertebrates have yielded a new range of chemical entities worldwide (Blunt et al., 2014). The marine ascidians have yielded more than 1200 compounds belonging to different SM classes such as alkaloids, NRPs and PKs. The isolated compounds have shown pharmaceutical importance

in cancer and as antivirals (Dou and Dong, 2019). As stated in Chapter One, most anti-microbial studies conducted on samples from the South African marine environment mainly focused on the invertebrates themselves, and not the microorganisms living in symbiosis with the invertebrates, hence the motivation for investigating the knowledge gap.

Distaplia skoogi has previously been identified as a source of SM. Three indole compounds have been isolated namely, 3,6-dibromoindole, 6-bromo-3-chloroindole and 6-bromo-2-oxindole. The isolated compounds showed moderate cytotoxicity against MA-MB-231 breast cancer cells (Bromley et al., 2013). It is worth noting that the above study did not test for anti-microbial activity. However, as reported in previous studies indole compounds are known to possess anti-microbial activity (Rajan and Kannabiran, 2014), and therefore *Distaplia skoogi* needs to be investigated as a source of anti-microbial compounds.

The *Spongia* genus has been studied as a source of bioactive SMs in several studies (Anand et al., 2006; Máximo et al., 2016; Noyer et al., 2011; Rueda et al., 1998), including the associated microbial diversity (Karimi et al., 2017). Crude extracts from various *Spongia* sp. showed bioactivity against a number of bacterial species such as *S. aureus*, *Micrococcus* sp., *Pseudomonas aeruginosa*, *Bacillus spaericus*, *B. subtilis* and *C. albicans* as well as against Ehrlich carcinoma and epidermoid carcinoma cell lines (Máximo et al., 2016; Anand et al., 2006). *Hypnea* genus organic extracts have shown anti-microbial, anti-oxidant, and other biological activities (Zerrifi et al., 2018), while *Hypnea spicifera* ethanolic extract has shown bioactivity against *S. aureus* and *E. faecalis* (Stirk et al., 2003). It is worth noting that these examples are not from specimens sampled from the South African marine environment. Studies have shown that samples of the same invertebrate species / genus collected from different locations have the potential to possess new chemistry (Noyer et al., 2011; Nuñez et al., 2008) and unique microbial diversity due to different factors such as temperature, salinity, light and pH of the inhabited environment. Therefore, MIA bacteria isolated from the South African marine environment hold the potential to be a source of new chemistry.

There is limited information available on the *Pseudodistoma africanum* ascidian with the only available literature on its discovery (Millar, 1954) indicating a potential source of new bioactive compounds and bacteria. Additionally, *Pseudodistoma africanum* is endemic to the South African marine environment (Ascidiacea World Database, n.d.). Similarly, the only published studies available for *Hamacantha (Vomerula) esperioides* focused on taxonomic classification, diversity and spatial distribution of the sponge (Maduray, 2013). Similarly, the literature search for microbial endosymbionts or as a source of bioactive compounds was fruitless as only database classification data appeared for *Pseudobranchiomma longa* (Kinberg, 1866) and *Leptophytum foveatum*.

In this study, MIA bacteria isolated from the above discussed marine invertebrates located in Algoa Bay (Port Elizabeth, South Africa) and Kalk Bay (Cape Town, South Africa) were screened for their ability to produce bioactive SMs, utilising small scale One Strain Many Active Compounds (OSMAC) culturing to prioritise bacteria for further investigation. Due to the re-discovery of known SMs from marine invertebrate-associated bacteria and fungi (Schmidt-Dannert, 2015), some of the isolated marine bacteria were also evaluated as a source of new compounds through the interrogation of their sequenced genomes using bioinformatic tools (i.e. genome mining). The purpose of sequencing was to assess the presence of new BGCs to enable strain prioritisation and assist in making linkages to the characterised compounds in Chapter 4. The hypothesis for this is that South African MIA bacteria possess the potential to be a source of new bioactive active compounds against bacteria (e.g., anti-microbial resistant strains) as the bacteria in this environment are understudied. The bioactive SMs were also characterised using NMR, and LC-MS (Chapter 4).

3.1. Anti-microbial and anti-inflammatory screening

Twenty-three MIA bacteria were screened for their anti-microbial potential using the agar overlay, well diffusion and high throughput screening (HTS) bioassays. The anti-microbial results are tabulated in Table 3.1 and Figure 3.1 (agar overlay and well diffusion). The anti-microbial study

was qualitative. Therefore, extracts or bacteria were regarded to have anti-microbial activity when there was a clear zone of clearance around the well or colony as illustrated in Figure 3.1. Evaluation of anti-microbial screening from marine bacteria was conducted with 8 microbial test strains covering Gram-negative (*E. coli* 1699 and *P. putida* ATCC27853), Gram-positive (*B. cereus* ATCC10702, *M. aurum* A+ and *S. epidermidis* ATCC14990) bacteria and fungi (*A. fumigatus* MRC and *C. albicans* NIOH). Additionally, the anti-inflammatory activity of the isolates was evaluated.

Agar overlay and well-diffusion assays were conducted in our laboratory, whilst HTS assays were conducted by external collaborators (Medina Foundation; Granada, Spain and MarBio; University of Tromso, Norway). Results for these bioassays are highlighted red in Table 3.1. These are reported as qualitative results. For agar overlay and well-diffusion assays a clear distinctive zone around the well or colony was regarded as anti-microbial when compared to the negative control. For the anti-inflammation results provided by MarBio, fractions with > 50 % inhibition of TNF-α were defined as active. Prior to performing HTS assay samples were pre-fractioned with solid phase extraction, to minimize false positives, as it has been well documented that complex microbial extracts interfere with the HTS assays resulting in a high number of false positives (Cutignano et al., 2015; Higgs et al., 2001). Having multiple assays (e.g., anti-inflammatory) is essential during primary screening, as secondary metabolites (SMs) are known to possess diverse biological activities.

Table 3. 1. Bioactivity screening of marine invertebrate-associated bacteria using well-diffusion, agar overlay and HTS (results shown in red) assays.

Organism ID	*E. coli* 1699	*P. putida*	*S. epidermis*	*B. cereus*	*M. aurum*	MRSA	*C. albicans*	*A. fumigatus*	Anti-inflammatory
KB07-58		ZBA*-				ZBB			
KB08-21.1			GYM*-	GYM*					TSB
KB08-22.1			GYM*-	GYM*-					GYM
KB08-48	TSA*-		GYM*-				ZBB, TSB	ZBB, TSB	ZBB
KB11-25			GYM*-	GYM*-				ZBB, GYM	ACM
KB11-44					GYM		ZBB	ZBB	
KB18-15						ACM			
PE06-105				GYM, TSA*		ACM		TSB, ZBB	
PE06-117	ZBA*-								GYM
PE06-34	ZBB, TSB, GYM*-	ZBB		TSB		ACM			
PE06-56	GYM, ACM*	GYM/ACM*	GYM, ACM*	ACM*	GYM				
PE08-149A		GYM					TSB	TSB	
PE08-149B	GYM*, ZBB	GYM*	GYM	GYM*	GYM				
PE08-33(3).1									ACM
PE12-107	GYM*-							ACM	ACM
PE12-116	GYM*-	GYM	GYM		GYM				
PE12-126								ZBB	
PE13-161			GYM				ZBB	ZBB	
PE13-172	ZBA*-			ZBA*-					ZBB, TSB
PE14-07	GYM, ZBB	GYM	GYM, ZBB	GYM*, ZBB*	GYM		ZBB, TSB	ZBB, TSB	GYM
PE14-104	GYM, ZBB	GYM	GYM, ZBB				ZBB, TSB	ZBB	
PE14-115							ZBB	ZBB	
PE14-63	GYM		GYM, ZBA*-	GYM			ZBB, TSB	ZBB, TSB	

Legend: Blank space indicates no observed activity; absence of * indicates activity observed only in liquid culturing, * indicates activity observed when cultured in liquid and solid media, and *- means activity only observed when cultured in solid media and not in liquid

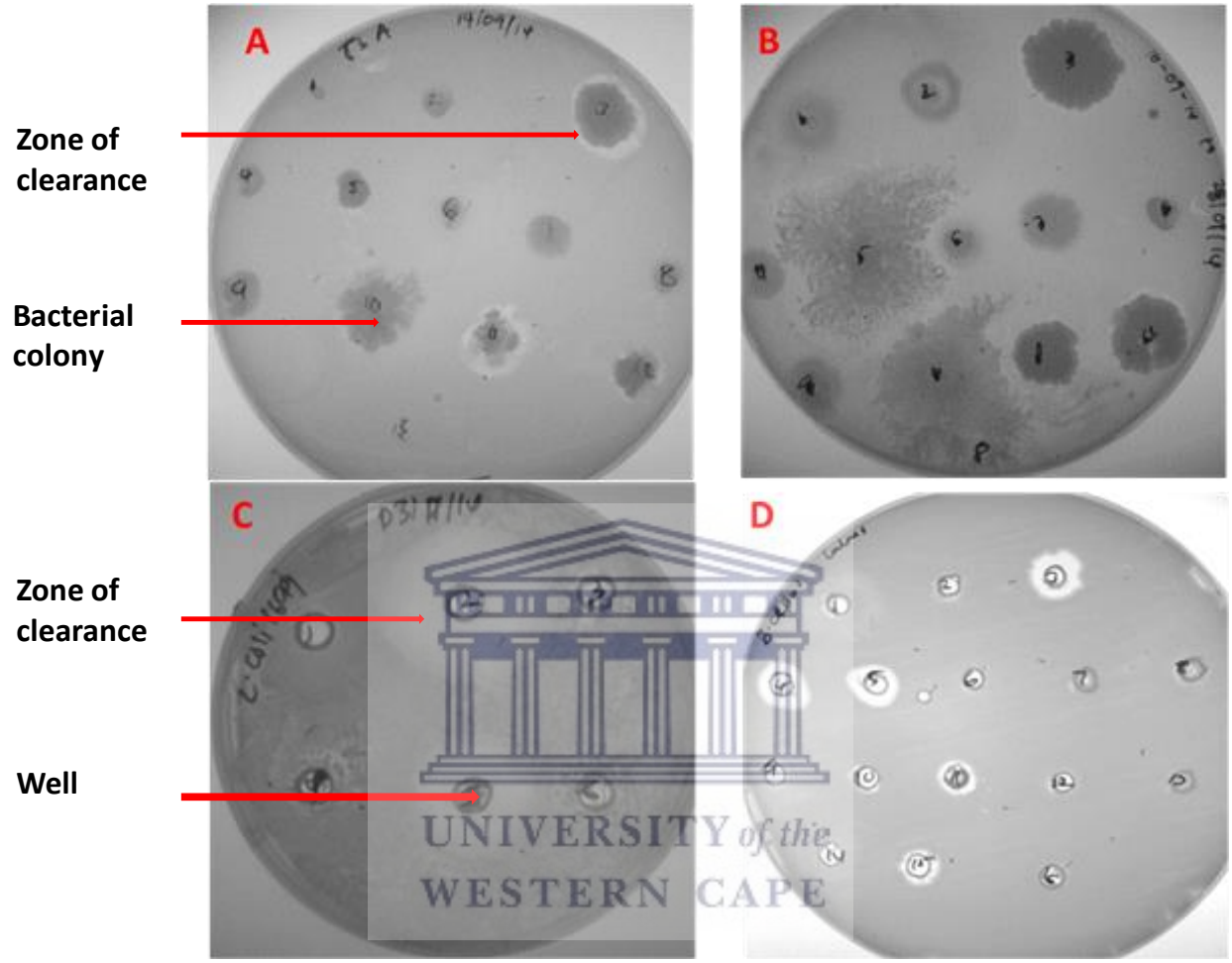


Figure 3. 1. Images illustrating an overlay assay (A; B) and well diffusion assay (C; D), for some of the marine invertebrate-associated bacteria tested during the study. Positive results are indicated by the clear zone around the well or spotted colony.

Briefly, microorganisms were cultured in four different media (GYM, ZBB, TSA and ACM) to increase the chance of inducing SM production as per the OSMAC principle (Chapter 1), with the results tabulated in Table 3.1. Three strains (PE12-126, PE08-33(3).1 and KB18-15) did not demonstrate anti-microbial activity for either liquid or solid fermentation experiments when tested with the overlay and well diffusion assays, suggesting a possibility of hydrophobic bioactive compounds with low diffusion properties. All isolates showed some biological activity in at least

one of the screening bioassays / methods, with 48 % of isolates displaying activity against *E. coli* 1699 and *S. epidermis*, respectively (Table 3.1).

Discovery of bioactive compounds for Gram-negative pathogens is trailing behind, compared to Gram-positive antibiotics (WHO, 2018). 48 % of the MIA bacteria showed anti-microbial activity towards Gram-negative bacteria, requiring further interrogations of these bacteria which may aid in the biodiscovery of new anti-Gram-negative compounds. The lower discovery rate of compounds active against Gram-negatives is mainly due to the multidrug / toxin efflux pumps system and lower low permeability porins on the cell membranes, as a result, the active compounds do not penetrate the cell or expelled before they can act (Silver, 2016; Zgurskaya et al., 2015).

The German / North American haemolytic-uraemic syndrome and bloody diarrhoea outbreak (May 2011), which resulted in 39 deaths from 810 infections, is a clear example of the need for new antibiotic therapies, as the virulent *E. coli* strain O104: H4 was found responsible for these infections was resistant to multiple antibiotics (Bielaszewska et al., 2011). In South Africa, a study on the ESKAPE pathogens showed an increase in antibiotic resistance over two years (2015 to 2016) as demonstrated by Perovic et al. (2018). In general, there is an increase in anti-microbial resistance by pathogenic microorganisms. The bioactivity against Gram-negative and antibiotic-resistant strains in this study suggest potential sources of compounds for dealing with the ESKAPE pathogen-related mortality worldwide (Schinke et al., 2017).

It is worth noting, that a significant number of the isolates showed activity against the two multi-drug resistant test strains (*E. coli* 1699 and MRSA). KB07-58, PE06-34 and KB18-15 showed activity against MRSA while PE06-34, PE06-56, PE06-117, PE08-149B, PE12-116, PE12-107 PE14-07, PE14-63, PE14-104, and KB08-48 inhibited the growth of multidrug-resistant *E. coli* 1699. The multidrug-resistant *E.coli* 1699 has been genetically engineered to be resistant to over 52 known antibiotics (list in Appendix II), while MRSA is resistant to β-lactam antibiotics (e.g. penicillin, methicillin and oxacillin), daptomycin and linezolid (Kaur and Chate, 2015).

Bioactivity against multidrug-resistant strains is a key feature for researchers to look at in developing new drugs with a unique mode of action (Hentschel et al., 2001), which will, in turn, be valuable in addressing antibiotic resistance. The results, therefore, present the much-needed hope towards addressing the lack of antibiotics against Gram-negative bacteria. In order to evaluate the relevance of the bioactivity against the multidrug resistance indicator strains, pre-fractionation or minimum inhibitory concentration of the crude extracts will have to be performed, to eliminate false negatives and positives.

The activity against *M. aurum* serves as a possible indicator for anti-tuberculosis activity as both species belong to the *Mycobacterium* genus. This assumption is because microorganisms of the same genus share some metabolic processes and utilise the same enzymes. Twenty-two per cent of the marine isolates were responsible for bioactivity against *M. aurum*. TB is a major burden to the South African population. It was the second leading cause of death among young people in 2010 (Casey et al., 2015). In 2013 South Africa was one of the top six countries adversely affected by TB worldwide (Knight et al., 2015). However, the anti- *M. aurum* bioactivity will need to be evaluated against *M. tuberculosis* to assess the full potential of MIA bacteria as a source of anti-tuberculosis compounds.

Twelve of the isolates displayed anti- *A. fumigatus* activity while eight displayed activities against both *A. fumigatus* and *C. albicans*. The anti-fungal bioactivity profiles revealed culturing-triggered SM production as different culturing conditions resulted in different bioactivity profiles. Despite most fungal infections being opportunistic, for a country like South Africa, it is critical to have anti-fungal compounds due to immunocompromised individuals as in the case of Human immunodeficiency virus (HIV) infections, and South Africa has the highest HIV infections worldwide (Bärnighausen et al., 2008; Kahn et al., 2006; Vandormael et al., 2018). Immunocompromised individuals are prone to secondary infection decreasing their chances of survival (Filice and Pomeroy, 1991). The immune system gets compromised when individuals are diagnosed with solid organ tumours, HIV, hematologic malignancy, or who have other conditions that require immunosuppressive therapies such as a solid organ or stem cell transplants (Memoli

et al., 2014). Fungal secondary infections are increasing due to increased transplants being performed, according to USA data between 1980 and 1997 fungal infections claimed 6,534 lives (Low and Rotstein, 2011). Fungal infections affect between 2 – 49 % of adult patients and mortality ranges between 20 – 70 % (Soysal, 2015). Fungal infections of concern include invasive aspergillosis, candidiasis, cryptococcosis and zygomycosis (Pappas et al., 2010; Puebla, 2012). These pathogens are also developing resistance to antibiotics such as fluconazole, echinocandin and Posaconazole (Pappas et al., 2010), which prompt a focus on anti-fungal compound search.

Anti-inflammatory activity was observed for 39 % of the MIA isolates from fractions extracted with Sepabeads® SP-207 (Hydrophobic) resin. In the human body, inflammation is part of the complex biological response to remove injury or harmful stimuli such as pathogens, damaged cells, or irritation. This response leads to many physical symptoms such as pain, fever, and swelling, as a result of many associated changes such as vasodilation, increased vascular permeability, and plasma extravasation (Srilekha et al., 2017), hence, the need for anti-inflammatory agents. Marine bacteria have yielded some anti-inflammatory compounds such as honaucins A–C, lobophorins A, N-substituted brominated monoterpene phenazine, lavanducyanin, cyclomarin A and salinamides A (S. Rangnekar and Khan, 2015). Our study suggests that the MIA bacteria could be a source of anti-inflammatory agents.

Marine bacteria in this study showed broad, narrow and species-specific bioactivities as depicted in Table 3.1. For example, PE14-07 showed activity against Gram-negative (*E. coli* 1699 and *P. putida*), Gram-positive (*B. cereus* and *S. epidermis*) and fungal (*A. fumigatus*) strains, while KB11-25 demonstrated anti-microbial activity against Gram-positive (*B. cereus* and *S. epidermis*) bacteria only. KB18-15 only proved to be effective against multidrug-resistant *Staphylococcus aureus* (MRSA)-MB5393. A total of thirteen MIA bacteria showed multiple anti-bacterial activities. A similar pattern was also observed by Anand and co-workers (2006) when working with bacteria isolated from marine invertebrates. The multiple bioactivity properties observed from this study indicates that South African marine bacteria are a great source of antibiotics for pathogenic microorganisms and anti-inflammatory.

Secondary metabolite production by microorganisms is highly influenced by the different stimuli (e.g. temperature, media, other organisms and pH) and therefore, altering these parameters may activate or enhance the production of SMs (Chai et al., 2012; Reen et al., 2015). In this study, growth media and the state of the media (solid or liquid) were some of the parameters investigated during the anti-microbial evaluation. These parameters and variations were tested as there are no universal media for SM production for various microbial strains (Frisvad, 2012). Changes in anti-microbial profiles were observed for several isolates (see Table 3.1 and Figure 3.2). Examples include PE14-07 which displayed activity against *B. cereus* when fermented in solid media (ZBA and GYM), whereas when in a liquid medium, it showed anti-microbial activity against *B. cereus*, *E. coli* 1699, *P. putida, S. epidermis* and *M. aurum* (GYM and ZBB). KB07-58 on the other hand, only displayed activity against *P. putida* on solid media (ZBA) and no activity in a liquid medium. Most of the anti-bacterial activities were observed on GYM media while ZBB and TSB showed more anti-fungal and anti-inflammatory bioactivity, suggesting the presence of multiple compounds with diverse biological targets (Figure 3.2). ACM had the most MRSA activity. Results obtained thus far were used to select the optimal media for further studies.

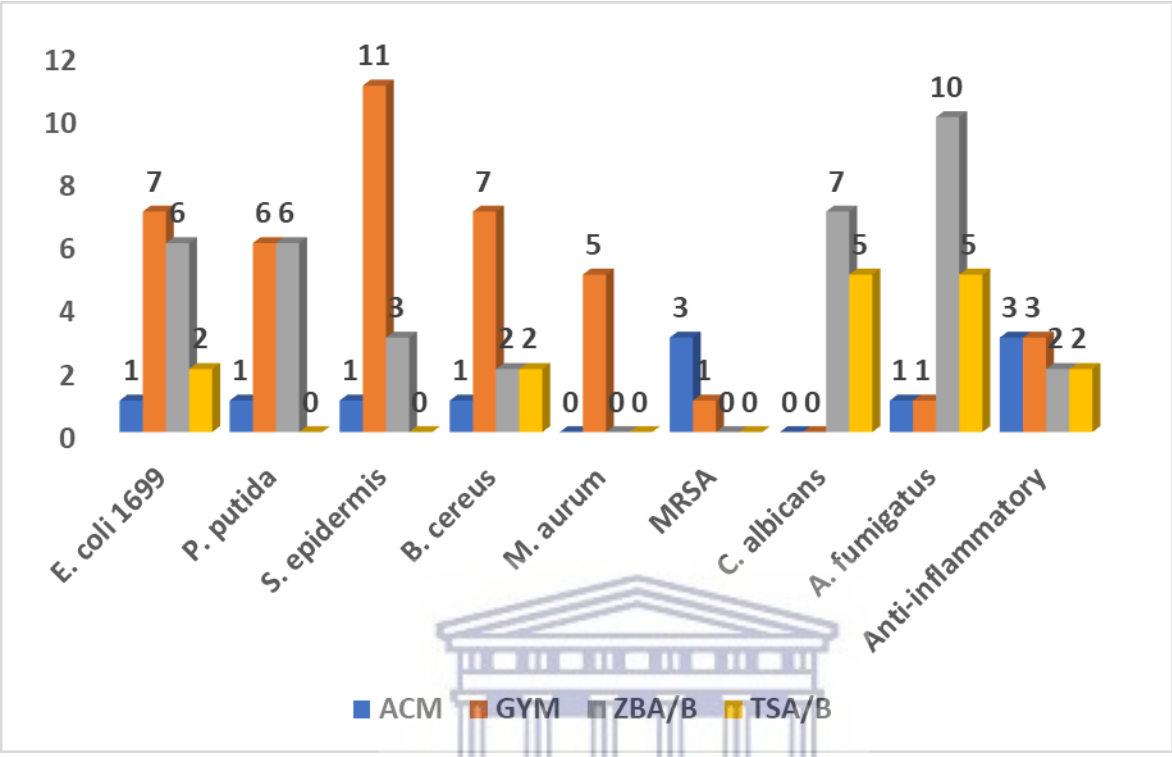


Figure 3. 2. Bar graph illustrating the number of bioactivities per media

The present study findings showed that several marine bacteria were induced into bioactive SM production when cultured in liquid media. This might be due to bacteria producing molecules that assist in nutrient scavenging, or response to stress from non-natural habitat, laboratory conditions. Due to the lack of studies on the effect of media state in the production of bioactive SM for non-filamentous bacteria, the preliminary data suggests that non-filamentous bacterial SM production is favoured in liquid medium with shaking fermentation, as more bacteria were able to show anti-microbial activity when grown on liquid media compared to solid media (Figure 3.3).

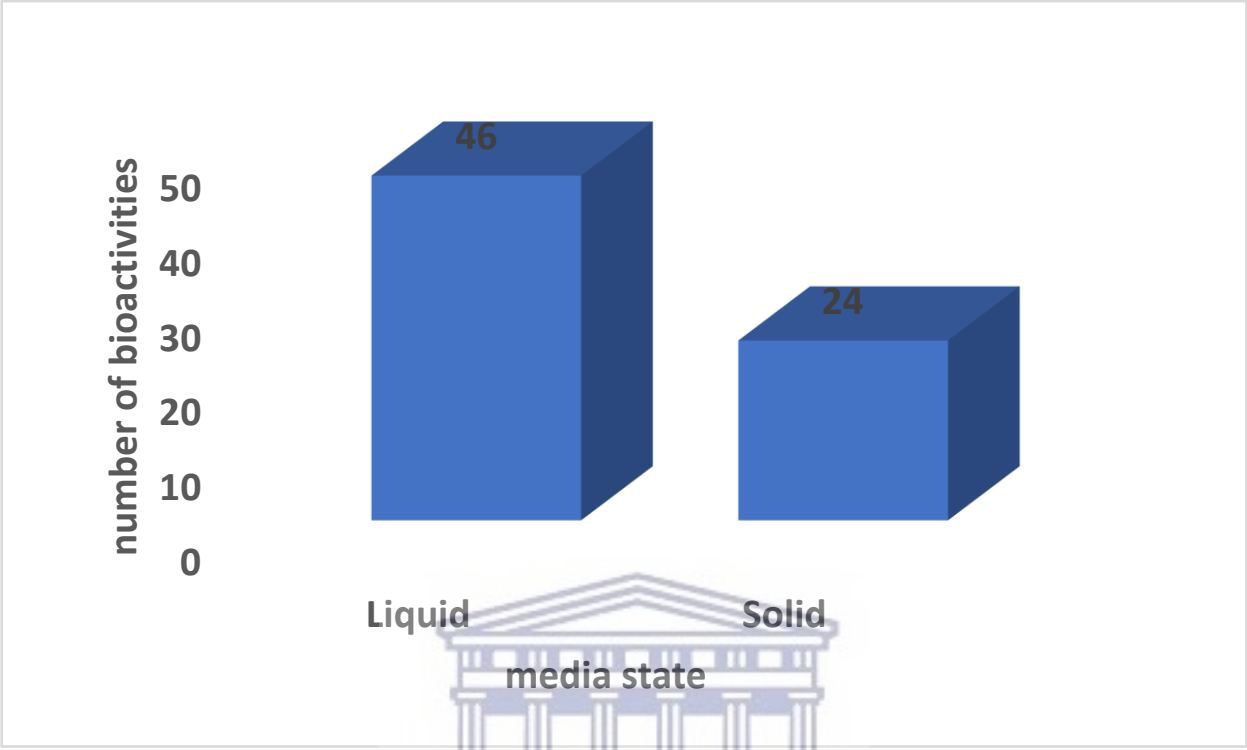


Figure 3. 3. Bar graph illustrating total bioactivities for solid and liquid media

The effect of the state of the media was observed in the production of pyrrocidines A and B by *Cylindrocarpon* sp. LL-Cyan426 were produced only when the culture was grown on solid media vs liquid media (shaken or unshaken) (Bigelis et al., 2006). Upon culturing *Penicillium commune* QSD-17 on solid rice media, the production of two new compounds was achieved which was not observed in liquid media namely; isophomenone and 3-deacetylcitreohybridonol (Gao et al., 2012). Media change effects were also observed in the production of capistruin, whereby medium M9 suppressed the production and medium M20 induced the production of capistruin (Knappe et al., 2008). *P. expansum* 40815 showed production of transcription of polyketide synthase, aromatic prenyltransferase, and terpenoid synthase when grown in liquid while in solid it showed only high transcription of meroterpenoid biosynthesising enzymes (Kim et al., 2016). There is a knowledge gap in the studies to evaluate the effect of media state in bacteria, as literature search revealed that the majority of studies are conducted on fungi.

However, these studies highlight the need for media optimisation and that there is no "one solution" to increasing product titre and triggering the expression of BGCs in microorganisms.

Media components also play a critical role in the production of secondary metabolites with simple changes such as the carbon or nitrogen source resulting in the induction of SM producing pathways (Vandermolen et al., 2013). This was demonstrated by our marine bacteria, as all media (Appendix I) used contained the same primary nitrogen source (yeast extract). The main difference was in their primary carbon sources, whereby GYM and TSA / B contain glucose whilst ZBB / A and ACM contained peptone and sodium pyruvate, respectively. GYM, ZBB and TSB (Appendix I) yielded the most anti-microbial activity hits compared to those cultured in ACM, which yielded the least anti-microbial activity. A possible reason for the poor performance of ACM media could be due to the activated charcoal, which absorbs the produced SMs making them unavailable to act on the test strains or due to the poor growth of the strains on the media as visually observed. Picot and Grenouillet documented that activated charcoal easily absorbs aromatic and heterocyclic compounds (Picot and Grenouillet, 1994).

KB08-48 displayed anti-*E. coli* 1699 (TSB) and anti-*S. epidermis* (GYM), while PE13-116 and KB07-58 isolates showed anti-microbial activity when grown on GYM and ZBB media, respectively. None of the test strains displayed identical anti-microbial profiles across all media tested, which was an indication of a stimulus-triggered expression by certain media components. In addition, no strain displayed activity under all test conditions, of which a similar phenomenon was observed by Matobole and co-workers (2017) when screening marine bacterial endosymbionts isolated from marine invertebrates (*Isodictya compressa* and *Higginsia bidentifera*) on solid media after using the OSMAC approach. The above observations also indicate the production of different compounds by some of the isolates.

As stated above several marine bacteria showed anti-microbial activity against the test strains when grown in GYM media. This could be because the media supplement such as yeast and malt extracts supplied the bacteria with precursors required for the synthesis of SM or contained BGC

expression-triggering molecules. In response to different medium components, the metabolism of microorganisms is altered resulting in the induction or inhibition of SM production (Crüsemann et al., 2016). Discovering the compounds responsible for inducing bioactive SM production is essential in fermentation optimisation. Through determining the SM inducing elements, minimal fermentation / defined media could be established which will minimise the current purification difficulties faced by the SM science community, due to the use of complex media such as GYM (Khan et al., 2013).

The difference in the screening assay outcomes for the targets shows the significance of multiple assays for the evaluation of potential bioactive SM producers, as one assay would have been prejudiced against certain strains. The prejudice can be due to a lack of compatibility by the active compound, for instance, in the well-diffusion assay the compound needs to diffuse into the agar for the bioactivity to be observed. Highly diffusing compounds may result in larger clear zones compared to poorly diffusing compounds, and no clear zones for non-diffusing (membrane-bound) compounds.

A number of studies have focused on microbial diversity of marine invertebrates found in Algoa Bay (Matcher et al., 2017; Walmsley et al., 2012; Waterworth et al., 2017), with limited published data covering MIA bacteria and their potential as sources of bioactive SMs, as only three articles are available (Cwala et al., 2011; Dorrington et al., 2016; Matobole et al., 2017), and no studies on the Kalk bay region to our knowledge. Our study indicated that South African marine invertebrates are a good source of MIA with the potential to be utilised as a source of valuable bioactive compounds. It is worth noting that the study focused on other MIA bacteria other than *Streptomyces*. The study is also acting as a motivation to study other marine invertebrates from the South African coastal environment as a source for SMs producing microorganisms.

3.2. Molecular identification of marine invertebrate-associated bacteria

The marine invertebrate associated (MIA) bacteria were identified by sequencing the 16S rRNA genes and the closest relatives were determined by BLASTn analysis (Table 3.2). The isolated bacteria belong to three different classes of microorganisms, namely Bacilli, Alpha- and Gamma-proteobacteria. All classes identified have previously been isolated from marine invertebrates, such as marine sponges and tunicates (Dupont et al., 2014; Kennedy et al., 2008; Kjeldsen et al., 2010; Margassery et al., 2012). Marine invertebrates are known to live in symbiosis with bacteria (Zuppa et al., 2014). Their well-known symbiotic relationship with microorganisms could be the reason these invertebrates are a great source of bioactive compounds, as research has shown that symbiotic microorganisms are the true producers of many bioactive SM previously isolated from marine invertebrates (Mehbub et al., 2014).

Table 3. 2. Identification of endosymbiotic strains screened for production of bioactive compounds using 16S rRNA analysis.

Strain ID	Closest BLASTn match	Identity Percentage (%)	coverage (%)	Closest relative Accession no.	Source
1. PE06-56					
	Bacillus mycoides strain D109_CV6R	99	100	MK883205.1	*Distaplia skoogi Michaelsen*, 1924
2. PE12-116					
	Vibrio splendidus LGP32	98	99	FM954972.2	*Eudistoma* sp. 011RSASPN
3. PE08-149B					
	Bacillus mycoides strain TS27	99	100	MN710449.1	*Spongia (Spongia)* sp.001RSASPN
4. PE14-07					
	Pseudovibrio ascidiaceicola strain F423	95	100	NR_041040.1	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886
5. PE14-104					
	Vibrio pomeroyi strain CAIM 578	100	100	NR_025547.1	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886
6. PE12-126					
	Bacillus sp. strain ZSQ1	89	99	MF403055.1	*Eudistoma* sp. 011RSASPN

7. PE14-63	*Vibrio tasmaniensis* strain Mj67	99	100	GQ455010.1	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886
8. PE06-34	*Bacillus anthracis* strain L9	100	100	KX832711.1	*Distaplia skoogi Michaelsen*, 1924
9. KB07-58	*Bacillus megaterium strain Rhizo_2*	95	93	MG263616.1	*Halianthella annularis*
10. KB18-15	*Bacillus thuringiensis* strain K5-4	99	100	KJ161422.1	*Sabellastarte longa*
11. PE14-115	*Vibrio sp. F15(2010)*	95	100	GQ861541.1	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886
12. KB08-21.1	*Bacillus wiedmannii* strain B9	99	100	MH196886.1	*Leptophytum foveatum*
13. KB08-22.1	*Bacillus wiedmannii strain* B9	100	100	MH196886.1	*Leptophytum foveatum*
14. KB08-48	*Bacillus anthracis strain JS41*	93	100	HQ857773.1	*Leptophytum foveatum*

15. KB11-25	*Bacillus thuringiensis strain* HER1410	100	100	KT887580.1	*Hypnea spicifera*
16. PE06-117	*Bacillus anthracis strain* AJGENE16011	98	100	KX443418.1	*Distaplia skoogi Michaelsen*, 1924
17. PE08-33(3).1	*Bacillus altitudinis* strain NL1	91	100	KM054685.1	*Spongia (Spongia)* sp.001RSASPN
18. PE12-107	*Pseudovibrio* sp. AU643	92	100	LN878629.1	*Eudistoma* sp. 011RSASPN
19. PE13-172	*Pseudovibrio ascidiaceicola* strain POLY-S9	99	100	KF155263.1	*Pseudodistoma africanum* Millar, 1954
20. PE13-161	*Psychrobacter faecalis* strain AP4Ob	92	100	KR051250.1	*Pseudodistoma africanum* Millar, 1954
21. KB11-44	*Bacillus toyonensis*, isolate Fr 15.1 E	91	100	LN995802.1	*Hypnea spicifera*
22. PE08-149A	*Erwinia* sp. CC10D 1	99	100	KM187110.1	*Spongia (Spongia)* sp.001RSASPN

23. PE06-105					
	Lysinibacillus fusiformis strain RB-21	99	96	CP010820.1	*Distaplia skoogi Michaelsen*, 1924

All isolated genera have shown the ability to produce anti-microbial compounds in some studies, which is in support of our findings (Devi et al., 2010; Margassery et al., 2012; Skariyachan et al., 2014). The majority of the strains identified in this study belong to the *Bacillus*, *Vibrio*, and *Pseudovibrio* genera which have previously been isolated from various marine invertebrates and marine environments (Hardoim and Costa, 2014; Thomas et al., 2010) such as *Polymastriapenicillus*, *Hyrtios altum*, *Haliclona simulans* and *Polycitor proliferus* (Alex et al., 2013; Alex and Antunes, 2015; Thomas et al., 2010). Numerous *Bacillus* species isolated from the marine invertebrate *Haliclona simulans* demonstrated anti-microbial activity against *Listeria monocytogenes*, *Clostridium perfringens* and *Bacillus cereus* (Phelan et al., 2012). *Pseudovibrio* P1Ma4, *Vibrio* P1MaNal1 and *Citricoccus* P1S7 isolated from the marine invertebrate *Phorbas tenacior* revealed anti-microbial, anti-plasmodial and anti-oxidant activity (Dupont et al., 2014).

The bacteria belonging to the genus *Bacillus* have yielded commercially available pharmaceutical antibiotics such as pumilin (Bhate, 1955), polymyxin (Sauberan and Bradley, 2018), gramicidin (Ostroumova et al., 2015) and surfactin (Chen et al., 2015). However, *Bacillus* remains a potential source of new compounds as genome mining demonstrated. Both *Pseudovibrio* and *Vibrio* genera are well-known producers of anti-microbial compounds with few compounds characterised compared to *Bacillus. Pseudovibrio* species, in general, have not been well explored as a source of SM (Flemer et al., 2012; Graça et al., 2013), with only three isolated compounds isolated showing bioactivity thus far (Crowley et al., 2014).

Interestingly, a strain belonging to the genus *Erwinia* (PE08-149A), a well-known plant pathogen, (including marine plants) was isolated (Amin et al., 2010). *Erwinia* has been found to live in marine environments together with brown alga known as *Padina pavonica* and the isolated strain showed no anti-microbial activity (Ismail et al., 2016). However, a member of this genus, *Erwinia carotovora* ssp. *carotovora* strain has been shown to produce the beta-lactam antibiotic carbapenem, suggesting potential within the genus (Barnard et al., 2007; Horinouchi et al., 2010).

Carbapenems are an important class of antibiotics as the currently available compounds are used as the last line of defence or against suspected resistance and have broad-spectrum activity (Papp-Wallace et al., 2011). Characterisation of PE08-149A to species level is essential in establishing whether it belongs to the *Erwinia carotovora* ssp. *carotovora*. or not, as this will shed some knowledge on whether other *Erwinia* sp. can produce bioactive compounds.

In our study, PE08-149A showed activity against *P. putida*, *A. fumigatus* and *C. albicans* when grown on TSB. Therefore, whole genome sequencing and bioassay-guided isolation studies on PE14-149A are essential in establishing whether it may produce carbapenem congeners or other SMs. To date, this genus has never been found in association with any marine invertebrate to our knowledge, making our study the first to report on the anti-microbial activity (anti-yeast, -fungal and -bacterial) of *Erwinia* sp. (PE08-149A) in association with a marine sponge.

PE14-104, -63 and -115 strains were identified as belonging to the genus *Vibrio*, and all displayed different anti-microbial profiles suggesting the presence of diverse BGCs within the genus. Interestingly, PE14-104 and -115 both showed activity against *A. fumigatus* and *C. albicans* when cultured in ZBB, indicating a possibility of the same compound being responsible (Table 3.1). However, bioactivity guided LC-MS profiling could be used to establish the presence of the same compounds and/or to identify differences in the chemical profiles of the active fractions. KB08-21.1 and KB08-22.1 displayed anti-microbial activity against the same test strains when cultured in GYM, and both showed 99.78% sequence identity at the 16S rRNA gene level to *B. wiedmannii* strain B9, further suggesting that the isolates represent the same species. However, the isolates showed anti-inflammatory activity on different media; TSB (KB08-21.1) and GYM (KB08-22.1), and therefore these differences need further validation. Three strains, PE06-34, PE06-117 and KB08-48 were all identified as *B. anthracis*; however, they all displayed different anti-microbial profiles (Table 3.1) which could suggest that they do not represent identical genotypes. However, genome comparisons would be necessary to assess this. It is worth noting that not all *B. anthracis* strains can cause anthrax. The gene encoding for anthrax is found on a plasmid (pXO1) which is not present in all isolates / strains of *B. anthracis* (Koehler, 2009).

However, when dealing with such precaution was taken. This does highlight the risk associated with the isolation of microorganisms as there is a possibility of introducing pathogens into the laboratory.

The *Pseudovibrio* strain PE14-07 demonstrated an anti-microbial profile similar to the *Pseudovibrio sp*. observed by Kennedy and co-workers (2009). The only difference in our study was that the strain did not show anti-microbial activity against MRSA and was not tested against *B. subtilus*. The absence of activity against MRSA could be due to different media components and culturing conditions which were utilised in our study. Another possibility could be the lack of pathways responsible for the observed bioactivity in our isolates compared with those from the Kennedy study. Unfortunately, the isolates from that study have not been subjected to whole genome sequencing, precluding a genomic comparison. The other *Pseudovibrio* sp. (PE13-172 and PE12-107) isolated in this study displayed distinctive anti-microbial profiles (Table 3.1), suggesting a possibility of the same species but different strains. These observations indicate the possibility of using bioassays to dereplicate isolates with identical 16S rRNA genes.

Regardless of whether the exact role of the endosymbiont is known, studies have shown that they are extremely important to their hosts. Their importance has been demonstrated, whereby the fate of the host is determined by the dominance of a certain bacterial endosymbiont (Garate et al., 2017). For example, *Pseudovibrio* sp. was discovered to be in low abundance in diseased *Rhopaloeides odorabile* and *Callyspongia aff. biru* marine sponges compared to healthy representatives (Versluis et al., 2018), suggesting a mutualistic relationship. A marine invertebrate (genus *Tedania*) associated *Pseudovibrio* sp. inhibited the growth of a *Bacillus* sp. capable of degrading invertebrate material (Versluis et al., 2018). Additionally, endosymbionts are known to influence the fitness of their host metabolically or immunologically (Garate et al., 2017; Versluis et al., 2018). The fitness of the host is enhanced by the bacteria with the production of SMs (toxins, amino acids and antibiotics) to assist in fighting pathogens, and supplying vitamins/cofactors for other vital metabolic processes (Jeganathan et al., 2013; Romano, 2018). The presence of *Pseudovibrio* sp. in our samples could indicate that the sponges were healthy at the time of

sampling. This statement is based on available data indicating *Pseudovibrio* is beneficial to the host (Romano, 2018). However, we cannot say for certain as no metagenomic bacterial analysis was performed on the sponge, nor has a baseline microbial community been established for "healthy sponges" of these genera and species.

Based on the broad anti-microbial activity profiles demonstrated by PE14-63, PE08-149B and PE14-07 (Table 3.1), these strains were selected for further genomic characterisation as potential sources of new SMs. The genome mining was done to check for the presence of new and known biosynthesis gene clusters (BGCs) within the three genomes.

3.3. Genome mining of secondary metabolites

The decline in the effectiveness of the classical SMs discovery methods, mainly due to re-discovery of known compounds has led to the development of other strategies such as genome mining. This strategy was developed to save time, money and to solve re-discovery of compounds as a major bottleneck (Cuadrat et al., 2017). Genome mining forms part of the –omics approach to fast-track the discovery of new compounds by identifying BGCs likely to be responsible for the production of the majority of known SM classes (Adamek et al., 2017).

Advancements in genomics and bioinformatics have demonstrated that the SM biosynthesis potential of microbes is usually much greater than what is observed with any biological assay and chemical analysis (Huang et al., 2016). The undiscovered SMs are there because the majority of the BGCs identified through genome sequencing and annotation are silent (un-transcribed) under laboratory fermentation conditions, and are therefore referred to as cryptic pathways (Xu et al., 2016). This leads to underestimating the potential of microbes to produce SM even for well-characterised SM producing microorganisms (e.g. actinomycetes) when using classical approaches (Matsuda et al., 2015).

Genome mining potential was showcased with the discovery of a new NRP BGC from *Streptomyces coelicolor* M145, in addition to a new tris-hydroxamate tetrapeptide iron chelator, coelichelin**,** being isolated with the assistance of precursor predictions (Challis et al., 2005). Genome mining of the *Streptomyces clavuligerus* genome resulted in the discovery of two new sesquiterpene synthases, indicating the possibility of two new terpenes (Hu et al., 2011). In this study, genome mining was used to evaluate the potential production of SMs by PE14-07, PE14-63 and PE08-149B. Using web-based bioinformatics tools such as Antibiotics & Secondary Metabolite Analysis SHell (antiSMASH) 4.0 / 5.0, Bagel 3, and Prediction Informatics for Secondary Metabolomes (PRISM), genomes from these strains were mined and analysed.

3.3.1. Whole genome sequencing

Bacterial strains PE14-07, PE08-149B and PE14-63 were selected for whole genome sequencing (WGS) based on properties mentioned in Section 3.2. The draft genome of PE08-149B was sequenced in a separate study and its genome sequence is publicly available on the NCBI database (accession no. LRPH00000000) (van Zyl et al., 2016). The *de novo* assembly of the sequence data resulted in a draft genome size of 5.8 Mbp, 5.7 Mbp and 5.91 Mbp for PE14-07, PE14-63 and PE08-149B, respectively (Table 3.3). The PE08-149B genome showed above 96 % average nucleotide identity (ANI) to *B. mycoides* WSB10204, VD078, KBAB4 and BPN40 indicating that they are the same species, corresponding to the 16S rRNA identification performed initially (Table 3.2). The analysis of the PE08-149B genome sequence revealed that it had a GC content of 35.1 %.

Table 3. 3. Genome analyses of PE14-07, PE14-63 and PE08-149B strains.

Strain ID	Number of reads	Average read length	Number of Contigs	Largest contigs (bp)	N50	Genome size (bp)	GC %	Coverage
PE14-63	2237468	187.4	177	1010546	422541	5657393	44	74.1x
PE14-07	1713057	231.71	78	447698	156767	5836631	51.3	68x
PE08-149B	4625328	193.7	101	97,016	-	5911306	35.1	118x

Legend: - Data not available

PE14-07 showed 96.85 % ANI to *Pseudovibrio* sp. AU243 when using the JSpeciesWS tool indicating a clear similarity between the two genomes, leading to the conclusion that the two are the same species. However, since both species are not taxonomically identified to species level, this could indicate the possibility of a new species. The GC (guanidine-cytosine) content of PE14-07 was found to be 51.3 %, which is identical to *Pseudovibrio* sp. POLY-S9 (6.6 Mb). *Pseudovibrio* sp. JE062 (5.7 Mb) and *Pseudovibrio* sp. FO-BEG1 (5.9 Mb) both have a GC content of 52.4 % which is the highest *Pseudovibrio* GC content thus far based on the NCBI available genomes (Alex and Antunes, 2015; Bondarev et al., 2013). The lowest GC content recorded for *Pseudovibrio* is 45.2 % from strain AD26. PE14-07 had the same cumulative genome size as a *Pseudovibrio* strain FO-BEG1 isolated from a marine sponge *penicillus* in the Atlantic Coast, Portugal (Alex and Antunes, 2015).

The PE14-63 genome revealed a 96.24 % and 96.03 % ANI to *Vibrio tasmaniensis* UCD-FRSSP16_25 and *Vibrio_tasmaniensis*_ZS-17, respectively, while it had below 90 % ANI to other *Vibrio tasmaniensis* species when analysed using JSpeciesWS. The lower similarity of PE14-63 to the other *Vibrio tasmaniensis* species could suggest the strain belong to a different group (species/ subspecies) together with UCD-FRSSP16_25 and ZS-17; however, this claim requires verification. PE14-63 had a GC content of 44 %, which is similar to other *Vibrio tasmaniensis* genomes available on NCBI databases.

3.3.2. PE14-07 secondary metabolite biosynthesis gene clusters

A comprehensive genomic analysis of BGCs distribution within the *Pseudovibrio* genus was conducted by Versluis et al., 2018. The study analysed 31 *Pseudovibrio* genomes and found that all genomes contained at least one bacteriocin cluster and a terpene BGC. Other BGCs found in most of the strains genomes were NRPs, siderophores and T1PK-T3PKs, while butyrolactone and homoserine lactone (Hserlactone) BGCs were present in two and six *Pseudovibrio* strains, respectively. Some BGC were unique to certain strains. For example, *Pseudovibrio* sp. POLY-S9,

Pseudovibrio sp. 8H04 and *Pseudovibrio axinellae* were the only strains with a T3PKS, T3PKS-hserlactone and Trans PKS-NRPS hybrid, respectively. The study highlighted the importance of sequencing each strain as some BGCs are only present in certain strains within a species. Aligning with our suggestion of not solely relying on 16s rRNA sequence for strain prioritisation, the PE14-07 genome was analysed for BGCs with antiSMASH to identify its potential to synthesise unique SMs. The identified BGC were compared with BGC from other studies.

Based on the antiSMASH output, a total of seven SM gene clusters were identified from the genome of PE14-07 (Table 3.4, Appendix VI). The discovered BGCs included a Type I-III PK hybrid which is interesting as type III PK are commonly found in terrestrial microbes (Shelest et al., 2015). The type I-III PK hybrid pathway genes showed 55 % and 50 % gene similarity to type I-III PK hybrid pathways in *Pseudovibrio sp.* Ad14 AD1 and *Pseudovibrio ascidiaceicola* strain DSM 16392, respectively. The low sequence similarity indicates the possible production of different and new PK hybrids by the three strains. The T1PK-T3PK BGC contains a gene encoding for RNA ligase activity indicating the possibility of a bacteriophage being the source of the BGC. The entire pathway showed 90.35 % similarity to *Pseudovibrio* FO-BEG1 based on NCBI BLASTn analysis, further indicating a potential new hybrid PK production by PE14-07 as no hybrid PK has been isolated from FO-BEG1.

Table 3. 4. Secondary metabolite pathways identified from PE14-07 based on antiSMASH, PRISM and Bagel analysis.

	antiSMASH			PRISM	Bagel
Secondary metabolite	**Gene cluster sizes (kb)**	**Highest Nucleotide similarity (%)**	**Similarity to characterised cluster***	**Secondary metabolite**	**Secondary metabolite**
Siderophore	22.4	No significant result	None	✓	-
NRP	39.4	95.1 (*Pseudovibrio* sp. Ad23 clone 3)	Rimosamide (14 %)	✓	-
PK	-	-	-	✓ ✓	-
PK Hybrid (Type I and III)	37.2	90.4 (*Pseudovibrio* sp. FO-BEG1)	None	-	-
Bacteriocin	14.2	86.2 (*Pseudovibrio* sp. FO-BEG1)	None	-	✓
Terpene	20.9	85.6 (*Pseudovibrio* sp. FO-BEG1)	None	-	-

betalactone	19.7	86.0 (*Pseudovibrio* sp. FO-BEG1)	Fengycin (13 %)	-	-
Hserlactone	15	86.3 (*Pseudovibrio* sp. FO-BEG1)	None	✓	-

***Similarity to the characterised clusters was determined using antiSMASH 5.0 internal settings, NB; the % on similarity to characterised cluster column refers to the number of genes. The highest nucleotide similarity represents NCBI Blastn results for the whole cluster.**

None of the gene clusters identified from PE14-07 demonstrated significant similarity to previously characterised BGCs, as analysed using antiSMASH (Table 3.4). Additionally, the cluster blast function on antiSMASH was utilised to analyse the similarity of the BGCs to any cluster detected from the NCBI GenBank, i.e., no experimental characterisation of the cluster (antiSMASH database). This step is further used as a dereplication method to assess the novelty and uniqueness of the BGC. For example, for the betalactone pathway, only 18 % of the genes showed similarity to a BGC from *Oceanicella actignis* strain DSM 24423, while 50 % of T1PK-T3PK genes showed similarity to those from a BGC from *Pseudovibrio* sp. Ad14 AD1. BLASTn results are also shown in Table 4.4 for all identified BGCs. These findings suggest that the PE14-07 strain is a potential source of new SMs from known SM classes.

Similarly, 14 % of the NRP BGC genes showed similarity to the genes responsible for the production of rimosamide in *Streptomyces rimosus* subsp. *rimosus* ATCC 10970 based on antiSMASH analysis. Both BGCs share two genes encoding a hypothetical protein and taurine dioxygenase. The hypothetical protein in contig 32-10 encodes for a GSCFA domain-containing protein. Rimosamide has no recorded direct bioactivity; however, it has been shown to enhance the anti-*B. cereus* activity of blasticidin S (McClure et al., 2016). However, given the very low percentage similarity and the lack of the core biosynthetic genes for rimosamide synthesis, it is unlikely that this pathway encodes the production of rimosamide, and rather is most likely to produce a new peptide.

The identified NRP was further characterised using NaPDos, NCBI BLASTn and NRPSpredictor2. NaPDos revealed that the NRP cluster contains three types of C-domains namely, LCL, dual and heterocyclisation. NaPDos is a bioinformatics tool used for the detection of SM domains, specifically KS-domain (PKs) and C-domain (NRPs) by analysing protein or (meta) genomic sequences. In addition to the identification of the domains, NaPDos classifies the domain(s) into seven different subtypes based on the reaction they catalyse (Ziemert et al., 2012): i) the LCL domain catalyses the formation of a peptide bond between two L-amino acids, whereas

the ii) DCL domain catalyses peptide bond formation between an L-amino acid and a D-amino acid in a growing chain, iii) the heterocyclisation domain is involved in the peptide bond formation and cyclisation of cysteine, threonine or serine residues, while iv) the epimerization domain catalyses the conversion of L-amino to D- amino acids within the chain. v) the modified amino acid (modAA) domain modifies the incorporated amino acid and vi) the dual domain performs both condensation and epimerisation of the incorporated amino acid. Lastly, vii) hybrid domains catalyse the condensation of an amino acid to connect to the polyketide, resulting in a hybrid of PK/NRP (Fischbach and Walsh, 2006; Rausch et al., 2007).

The presence of a heterocyclisation domain in the PE14-07 cluster suggests the production of a macrocyclic peptide. This assumption was further supported by PRISM (Results Appendix VI) structural prediction feature which showed cyclic peptides as possible final structures (Figure 3.4). During development, PRISM's ability to predict structures is validated by calculating the Tanimoto coefficient between the predicted structure and structure of known BGCs with known structures. The Tanimoto coefficient is a calculation of similarity between two chemical structures and a Tanimoto coefficient median of 0.67 is indicative that the predicted structures are similar to the true structure (Skinnider et al., 2017). However, due to the limited biosynthesis and enzymology information on SMs and the fact that the program relies on known protein homology to predict structure, it results in the program failing to predict novel enzymes and thus chemical structures. Additionally, in some instances the point of tailoring is not possible to predicate; for example halogenase can catalyse the insertion of chlorine in multiple sites within the compound (Skinnider et al., 2020, 2017). For these reasons, it is not possible yet to predict structures with 100 % confidence from BGCs.

Three of the identified C-domains showed some identity to the syringomycin BGC C-domain from *Pseudomonas syringae* pv. *syringae* except for one which was identical to the C-domain of bleomycin BGC from *Streptomyces verticillus*. The KS-domain found in the first module is related to the KS-domain found in the BGC of epothilone A produced by *Sorangium cellulosum*, (45% identity) which is classified as a hybrid KS subtype. However, since the identity percentages are below 85 %, it continues to indicate that the C-domains in the NRP BGC could catalyse the formation of a different compound.

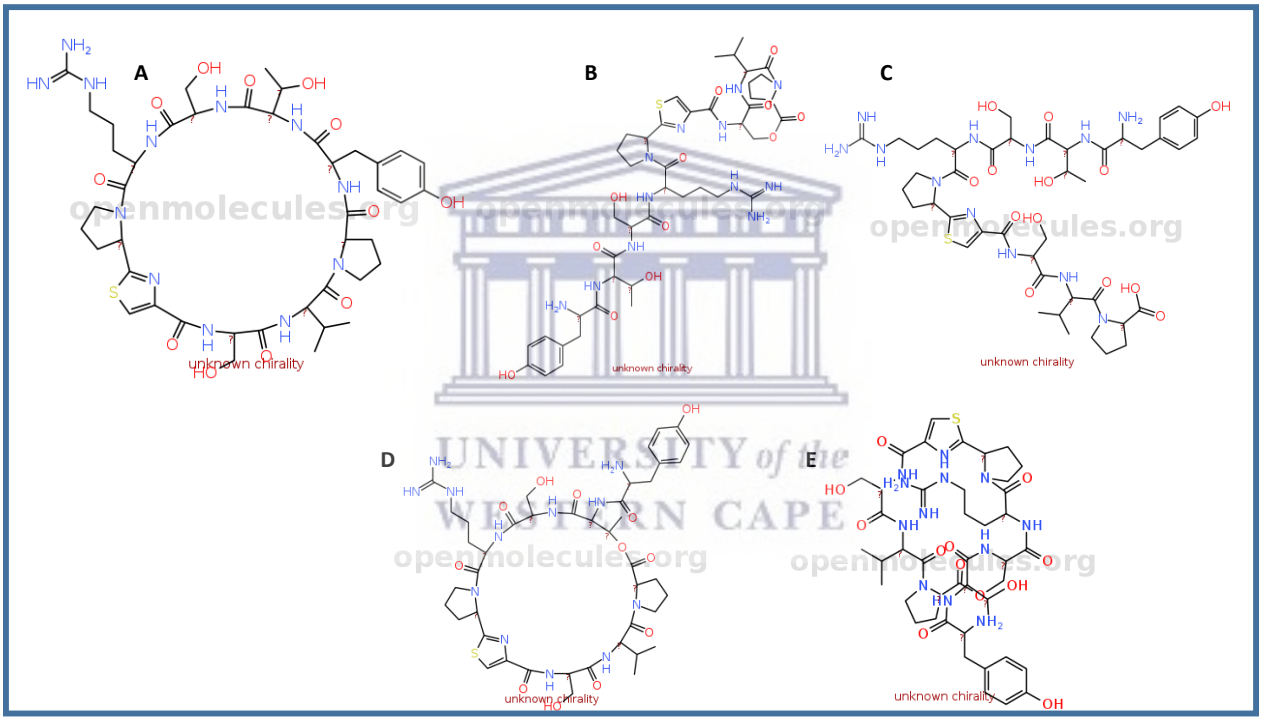


Figure 3. 4. PRISM predicted structures for PE14-07 NRP gene cluster (generated using open molecules)

Further analysis of the four NRPS genes revealed seven A-domains indicating that seven amino acids are likely incorporated in the synthesis of a complete NRP molecule (Fischbach and Walsh, 2006; Trauger and Walsh, 2000). NRPSpredictor2 analysis revealed the NRP amino acid structural backbone as Pro-Gln-Arg-Pro-Cys-Thr-Gln. Interestingly, a study showed macrocyclic peptide production by a *Pseudovibrio* strain, the peptide was specifically active against *B. subtilis* (Vizcaino et al., 2010). It is tempting to speculate that PE14-07 produces a similar bioactive

peptide; however, further bioinformatic analysis of the pathway and the corresponding domains (C- KS- and A-domain) suggests that the NRP cluster is new, based on the pathways to which it has similarity. Therefore, given all the analyses presented for this pathway, it would not be possible to predict the final structure and bioassay-guided identification and / or heterologous expression are necessary to study the compound produced by this pathway.

A BGC encoding a NRP-independent siderophore was identified on the PE14-07 genome. Siderophores are iron chelating components, iron is important for microorganisms in general as it is used as cofactors for enzymes. There are low levels of dissolved iron in the marine environment. The limited availability of dissolved iron (<0.2nM) creates competition amongst marine organisms, hence the abundance of siderophore BGCs in marine microbes which gives the producers an added advantage over others (Boiteau et al., 2016). The PE14-07 siderophore BGC cluster was induced by GYM medium, and this was confirmed with the use of CAS media assay. Further characterisation of the cluster is presented in Chapter 4.

The identified siderophore BGC is about 14.7 kb (Figure 3.5) in length. The presence of genes with similarity to lysine/ornithine N-monooxygenase (*sidC*), *IucA-IucC* (*sidD* and *sidF*) and *IucD* (*sidE*) suggests a hydroxamate siderophore (Hirschmann et al., 2017). All biosynthetic genes showed similarity to ochrobactin biosynthesis genes except for *sidG* which showed similarity to a gene on the NRP-PK hybrid cluster from *Streptomyces* sp. responsible for the production of the compound JRBIR-06, based on Minimum Information about a Biosynthetic Gene cluster (MIBiG) database. Ochrobactin siderophore has been previously isolated from marine bacteria such as *Ochrobactum* sp. SP18 and *Vibrio* species (sp. DS40M5 and sp. S4BW). A literature search for anti-microbial activity of ochrobactin was fruitless (Hirschmann et al., 2017). JRBIR-06 showed anti-cancer activity and belongs to a group of SMs known as antimycin-Type. These groups of compounds are known to possess anti-cancer, anti-inflammatory and anti-microbial activity (Li et al., 2017).

Figure 3. 5. An overview of the PE14-07 siderophore gene cluster as predicted by antiSMASH. Legend grey other genes, light red- core biosynthesis genes (BGs), dark red – additional BGs, blue- transport-related genes

In addition to the core biosynthesis gene, other genes are present which encode for proteins involved in activities such as transporting of the siderophore in and out of the cells and modification of the core structure. Downstream of the cluster lies a gene predicted to encode an extracellular ligand-binding receptor *(sidJ*). The antiSMASH translated protein sequence shows a 99.75 % similarity to ABC transporter substrate-binding protein. Upstream of the gene cluster lies a transmembrane transport gene (*sidB*) encoding for a protein belonging to the major facilitator superfamily (MFS). A protein from the same family has been proven to facilitate enterobactin excretion in *E. coli* (Han et al., 2013). A putative integrase gene is located upstream of the cluster suggesting that integration / horizontal gene transfer (HGT) was the mechanism by which the BGC was acquired by PE14-07. Integrase was previously speculated to be responsible for the HGT of 3-dichloropropene catabolic pathways genes (Poelarends et al., 2000). Siderophores are not only important as potential anti-microbial SMs, but they can also even serve as iron overload relievers and drugs for other diseases (Chen et al., 2019).

The betalactone BGC cluster was identified on the PE14-07 genome with 13 % of the genes showing similarity to a characterised fengycin NRP cluster from *Bacillus velezensis* FZB42. It is worth noting that cluster blast of the betalactone BGC cluster showed no similarity to clusters from *Pseudovibrio* strains based on antiSMASH, despite the NCBI BLASTn analysis of the whole cluster showing 86 % similarity to *Pseudovibrio* FO-BEG1. Additionally, antiSMASH showed that PE14-07 and FO-BEG1 contain different BGCs. The cluster contained a sequence encoding for a gene transfer agent, showing 46 % similarity to *Salinarimonadaceae* bacterium HL-109. This suggests that the gene cluster could have been obtained via horizontal gene transfer.

Genome analysis of PE14-07 with PRISM (Table 3.4) resulted in the identification of two PK BGCs, one NRP and an acyl-homoserine lactone (Hserlactone). Both PRISM and antiSMASH predicted a BGC which encode for SM mainly functioning as a quorum sensing (QS) compound (Hserlactone). Genome mining of QS compounds is very important for the discovery of SMs, as QS compounds are known to form part of other SM production and pathogenicity as they are involved in gene expression regulation (Barnard et al., 2007; Churchill et al., 2011; Liu and Li, 2011). Understanding the role of this compound is imperative in the future discovery of SMs within the *Pseudovibrio* genus and other bacteria. They can be added to the fermentation media to increase product yield or induce expression as part of an OSMAC approach. Conversely, if the compound inhibits expression / synthesis of SMs, its expression can be suppressed using genetic tools enabling the expression of the target SM.

For example, *Erwinia* sp. uses Hserlactone based systems to regulate carbapenem expression. Hserlactone (e.g. 3-Oxo-C6-HSL) binds and activates the response regulator gene *car*R, which triggers the expression of the carbapenem BGC (Barnard et al., 2007). The Hserlactone system was also shown to regulate the production of phenazine-1-carboxamide in *Pseudomonas chlororaphis* HT66 through mutation experiments. Additionally, overexpression of genes responsible for Hserlactone (3-OH-C6- Hserlactone, 3-OH-C8-Hserlactone and 3-OH-C10-Hserlactone) compounds production resulted in increased production of phenazine-1-carboxamide by 2.3-fold compared to the HT66 wild strain (Peng et al., 2018). The Hserlactone system has been shown to inhibit the expression of enterocyte effacement genes in enterohemorrhagic *E. coli* (Yang et al., 2018).

Genome analysis using Bagel versions 3 and 4 (Table 3.4) resulted in the identification of a bacteriocin BGC belonging to the class of sactipeptides, which supports the prediction by antiSMASH. Sactipeptides are RiPP synthesised with the coupling of the sulphur in a cysteine molecule to the α-carbon of the next amino acid (thioether bond), giving rise to structural diversity and bioactive compounds (Himes et al., 2016). The identified sactipeptides main biosynthesis gene showed only 45 % similarity to a hypothetical protein from *Emiliania huxleyi* CCMP1516,

indicating the possibility of a new sactipeptide bacteriocin. Bagel also identified modification genes within the BGC such as nitrite reductase, uroporphyrinogen-III C-methyltransferase and probable cysteine desulfurase enzymes which are responsible for modification reactions.

Tropodithietic acid (TDA) production is common in the marine *Rhodobacteraceae* family, which the *Pseudovibrio* genus belongs to. TDA has been shown to have anti-cancer and anti-microbial activity. Homologues to the six (*tdaABCDEF*) genes responsible for TDA biosynthesis were found in the PE14-07 genome when using TDA genes from *Pseudovibrio* sp. p12 in a comparative search (Table 3.5, Appendix VII). The TDA BGC is located within a 130 kb contig identified as contig 3. The genes showed an average of 82.25 % of identity to *Pseudovibrio* sp. P12 TDA genes.

Table 3. 5. TDA biosynthesis genes identified from PE14-07 using *Pseudovibrio* sp. P12 TDA genes.

Gene	Function/ name	Nucleotide identity to *Pseudovibrio* sp. 12 (%)
tdaA	Transcriptional regulator, LysR family protein	78.69
tdaB	Glutathione S-transferase domain protein	79.23
tdaC	Prephenate dehydratase	83.56
tdaD	Acyl-CoA thioester hydrolase	87.36
tdaE	Acyl-CoA dehydrogenase	85.23
tdaF	Phosphopantothenoylcysteine decarboxylase	79.23

BLASTn analysis of the contig showed 82.84 % to *Pseudovibrio* sp. FO-BEG1 which is a known TDA-producer (Porsby et al., 2011). The mining of TDA BGC shows the shortfall of SM mining tools such as antiSMASH as they focus on classes or groups and do not identify all SM genes / clusters which may encode the production of bioactive compounds. This indicates the need for the

further development of bioinformatics tools within SM discovery. The advancement in this area will also facilitate the linking of compounds to enzymes responsible for their biosynthesis. This further indicates that phenotypical screening is very important as it will show both known classes and unknown classes, encouraging the pairing of the approaches.

The presence of the bacteriocin and terpene BGCs within the PE07-14 genome aligns with previous studies (Versluis et al., 2018), which may indicate the importance of their products in the survival of the strain in the environment, as they may aid in fighting against competing microbes. Isolation and characterisation of products from the two BGCs could aid in understanding the exact purpose they serve within the *Pseudovibrio* genus. Terpenoids have been shown to benefit the host (invertebrate) by precluding predation and fouling organisms, while the bacteriocin could be important in the elimination of competition (Song et al., 2015; Versluis et al., 2018).

An interesting observation during the study was that 5 out of the 7 BGCs identified in the genome of PE14-07 are most closely related (>85 % nucleotide identity; Table 3.4) to sequences belonging to the *Pseudovibrio* FO-BEG1 strain. Despite the two strains sharing only 83.81 % ANI, indicating that they are different species, this suggests that the two strains shared a common ancestor but evolved differently due to different environmental pressures such as salinity and temperature. Furthermore, NRP-independent siderophores within the *Pseudovibrio* genus are randomly distributed with no geological or host specific pattern. The strains were isolated from seawater, sea sediment, bryozoans and sponges scattered worldwide. Interestingly, even species isolated from the same host have a different genomic composition (Versluis et al., 2018).

The unique gene clusters and genes observed among closely related bacterial strains or related BGCs respectively could be due to either the gaining and / or losing of tailoring enzymes, recombination, gene deletion, or horizontal gene transfer of the BGCs (Dittmann et al., 2015; Doroghazi and Metcalf, 2013). *Pseudovibrio* genomes are characterised by the presence of mobile elements which facilitates genomic variation within microbes (Alex and Antunes, 2015). This was observed with the presence of an integrase within the siderophore BGC from PE14-07.

3.3.3. PE08-149B secondary metabolite gene clusters

A comprehensive genome analysis of the genus *Bacillus* was conducted whereby a total of 1566 species were analysed for the presence of SM BGCs using antiSMASH (Grubbs et al., 2017). The study found that NRP BGCs are the most abundant class with PKs being the least abundant, while NRP-PKs hybrid pathways are common within this genus. Bacillibactin was found to be preserved in most of the bacilli strains, indicating its importance for the survival of the strain. To evaluate the potential chemical diversity of PE08-149B bioinformatics tools such as antiSMASH, Bagel 3 / 4 and PRISM were used in this study. The data was used to assess new BGCs and compare with other *Bacillus* species.

Genome mining of PE08-149B resulted in the identification of 11 SM BGCs belonging to different classes, namely terpene, NRP, bacteriocin (RiPP) and siderophore (Table 3.6, Appendix VI). The most abundant BGCs were bacteriocins (5), followed by NRPs (3) with a total of 230 kb dedicated to SMs biosynthesis making up 4 % of the genome. Cluster analysis showed the highest nucleotide similarity to clusters present in a range of *Bacillus* species (Table 3.6). One of the bacteriocin pathways (19 kb), predicted to encode a lasso peptide, showed 100 % similarity to the paeninodin BGC from *Bacillus anthracis* strain Ames. Lasso peptides were originally associated with the proteo- and actinobacteria phyla and were only recently identified from the Firmicutes phylum through genome mining (Maksimov et al., 2012; Zhu et al., 2016).

Paeninodin has previously been shown to not display any anti-microbial activity (Zhu et al., 2016). The other putative bacteriocin BGCs showed no similarity to characterised pathways. Further analysis with Bagel 3 confirmed the lasso peptide BGC and only predicted 3 BGCs instead of the 4 proposed by antiSMASH. The additional predicted putative bacteriocin gene clusters belong to the Head_to_tail_cyclized_peptides (2 clusters), and a Linear azol(in)e-containing peptides (LAP). The difference could be attributed to the reasons stated in Section 3.3.2.

Table 3. 6. Secondary metabolite pathways identified from PE08-149B based on antiSMASH, PRISM and Bagel analysis.

	antiSMASH			PRISM	Bagel
Secondary metabolite	**Gene cluster sizes (kb)**	**Nucleotide similarity (%)**	**Similarity to known cluster**	**Secondary metabolites**	**Secondary metabolite**
Siderophore	14	98.18 (*Bacillus weihenstephanensis* KBAB4)	Petrobactin (100 %)	✓	-
Bacteriocin	14	96.72 (*Bacillus cereus* strain ZB201708)	Paeninodin (100 %)	✓ ✓ ✓	✓ ✓ ✓
	10	88.48 (*Bacillus tropicus* strain LM1212-W3 plasmid p1)	None		
	10	98.64 (*Bacillus mycoides* strain BPN401)	None		
	24	96.71 (*Bacillus mycoides* strain TH26)	None		
	19	98.16 (*Bacillus mycoides* strain TH26)	None		
NRP	29	99.59 (*Bacillus weihenstephanensis* strain WSBC 1020)	None	✓ ✓ ✓	-
	43	98.65 (*Bacillus weihenstephanensis* strain WSBC 10204)	Bacillibactin (100 %)		

	63	94.95% (*Bacillus wiedmannii* bv. *thuringiensis* strain FCC41 plasmid pFCC41-4-144K)	Lactocin S (10 %)		
Betalactone	25	93.11 (*Bacillus wiedmannii* bv. *thuringiensis* strain FCC41)	Fengycin (40 %)	-	-
Terpene	18	98.37 (*Bacillus weihenstephanensis* strain WSBC 10204)	None	-	-

The identified NRP-independent siderophore cluster (14 kb) was 100 % identical to that of the petrobactin BGC from *Bacillus anthracis*.

The genes of the 43 kb NRP identified by antiSMASH were similar to the biosynthesis genes of a bacillibactin cluster by 100 % from a *Bacillus velezensis* FZB42. Bacillibactin is a common NRP siderophore to many species within the *Bacillus* genus (Lee et al., 2011), and the high similarity of the pathway identified suggests that the PE08-149B is capable of producing the same siderophore (bacillibactin) (Figure 3.6).

The 60 kb NRP BGC exhibited 10 % genes similarity to the BGCs from *Lactobacillus sakei* and *Streptomyces* sp. Acta 2897, encoding for lactocin S and skyllamycin, respectively. Both lactocin S (Ross et al., 2010) and skyllamycin (Pohle et al., 2011) have anti-microbial activity against bacteria and fungi. However, due to low similarity, this BGC is unlikely to encode the synthesis of these compunds and rather expected to encode a new NRP.

Analysis of the 29 kb NRP identified two modules with two C-domains and A-domains. The C-domains were identified as heterocyclisation and DCL suggesting a cyclic peptide. The DCL domain showed 35 % identity to a C-domain from a lychenicin BGC while the heterocyclisation domain showed 39 % identity to a bacitracin A BGC C-domain; both from a *Bacillus licheniformis*. Since NaPDoS sequence analysis are below 85 % identity this suggests the production of a new or uncharacterised NRP. NRPSPredictor2 could only predict one substrate as cys, the failure could be due to the presence of an uncharacterised A-domain. The proposed structure based on PRISM analysis is shown in Figure 3.6.

The 63 kb NRPS-ladderane hybrid cluster contained five modules according to antiSMASH; however, the cluster contained 6 C-domains. NaPDos analysis of the C-domains revealed a presence of three LCL, two DCL and one modAA C-domains subtypes. The C-domains showed less than 40 % identity to microcystin, actinomycin and syringomycin BGCs C-domains. Based

on NRPSPredictor2 the amino acid backbone is Leu-X-Lys-Thr-X (X representing unknown amino acids). This indicates the presence of two new A-domains within the cluster. The proposed structure based on PRISM analysis is shown in Figure 3.6.

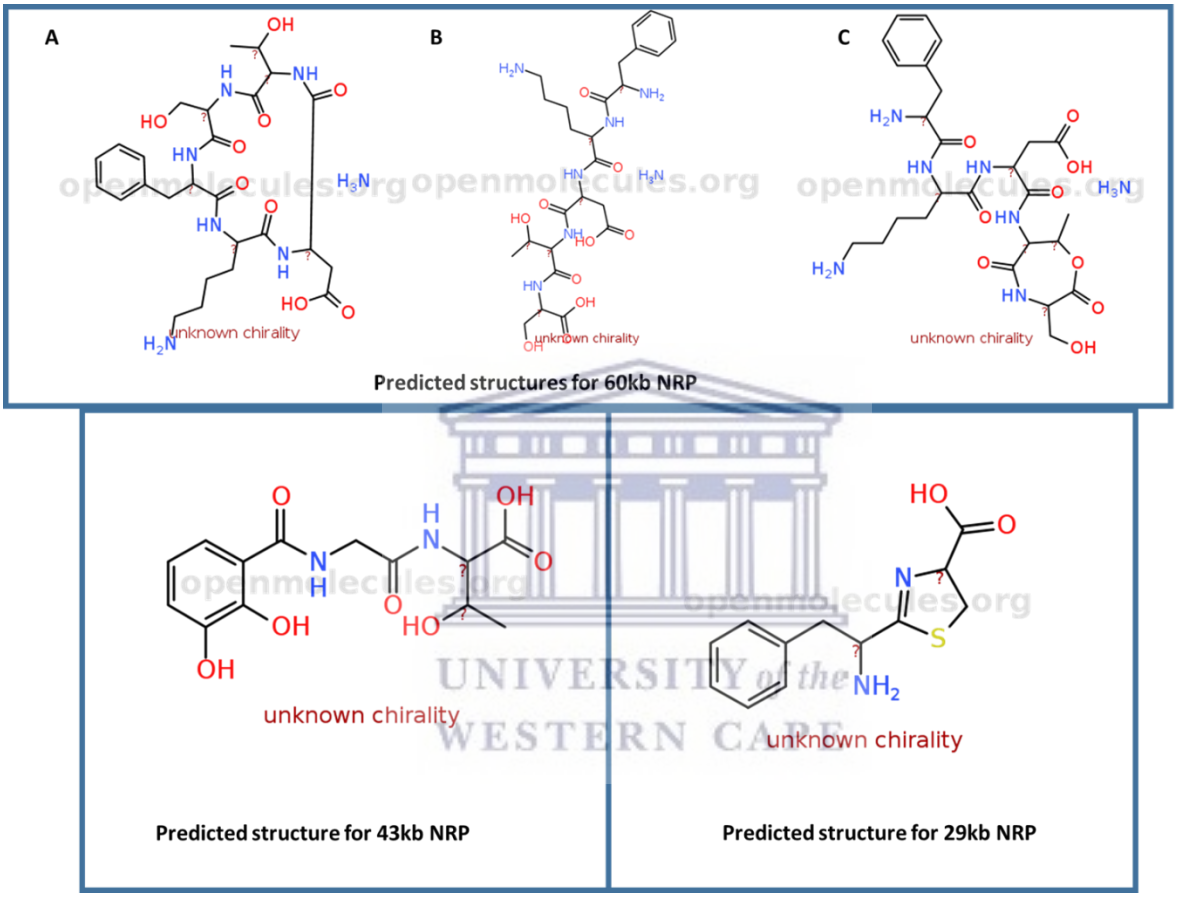


Figure 3. 6. PRISM predicted structures for PE08-149B NRP gene clusters (generated using open molecules)

PRISM analysis (Table 3.6, Appendix VI) of the PE08-149B genome resulted in a total of 7 clusters. Four of the identified clusters were classified as NRPs, while two were found to be bacteriocins, and one was an unknown thio-template. The unknown thio-template cluster was incomplete as it contains only two open reading frames, an A-domain and nitroreductase domain. The pathway was considered incomplete as it lacked the minimum domain requirements for complete synthesis of NRP, which include adenylation, peptidyl carrier protein, condensation and thioesterase domain. Bagel 3 analysis (Table 3.1, Appendix VI) of PE08-149B genome resulted in identification of 3 bacteriocin BGCs while antiSMASH resulted in a total of 5 clusters. PE08-149B contains a betalactone BGC with 40 % genes similarity to the fengycin BGC. Fengycin production is common amongst the genus *Bacillus* (Deleu et al., 2008); however, due to low similarity, the cluster likely produces a different compound.

3.3.4. PE14-63 secondary metabolite biosynthesis gene clusters

A total of six putative SM biosynthesis gene clusters were identified from the PE14-63 genome using antiSMASH (Table 3.7, Appendix VI). Amongst the identified clusters were two bacteriocin-encoding (RiPP) and one siderophore pathway. The siderophore pathway identified from PE14-63 shows 100 % gene similarity to a vibrioferrin biosynthesis pathway from *Vibrio parahaemolyticus*. Vibrioferrin was discovered from *Vibrio parahaemolyticus* in 1992 while the structure was only fully elucidated in 1994 (Yamamoto et al., 1994, 1992), and the BGC encoding for its biosynthesis was later discovered through metagenomic library screening (Fujita et al., 2011). Vibrioferrin has been characterised as a siderophore with low iron affinity compared to other siderophores and is susceptible to photolysis (Amin et al., 2009), with no study evaluating the anti-microbial activity of the compound as yet, to our knowledge.

Strains within the genus *Marinobacter* capable of vibrioferrin biosynthesis form a mutualist symbiosis with dinoflagellates (*Gymnodinium catenatum* and *Scrippsiella trochoidea*), whereby the host benefits from the iron release from vibrioferrin during photolysis while the bacterium

obtains carbon fixed by the host (Timmermans et al., 2017). This led us to speculate that the sponge (*Hamacantha* (*Vomerula*) *esperioides* Ridley and Dendy, 1886) host encountered in the present study might also be benefiting from the *Vibrio* sp. PE14-63 in the same manner. Siderophores are one of the most common groups of SMs within marine bacteria due to low levels of bioavailable iron in the marine environment (Årstøl and Hohmann-Marriott, 2019). Our study further affirms this as siderophore BGCs were present in all three strains (PE14-63, PE14-07 and PE08-149B).

Table 3. 7. Secondary metabolite pathways identified from PE14-63 based on antiSMASH, PRISM and Bagel analysis.

	antiSMASH			PRISM	Bagel
Secondary metabolite	**Gene cluster size(kb)**	**Nucleotide similarity (%)**	**Similarity to characterised cluster**	**Secondary metabolite**	**Secondary metabolite**
Siderophore	14.9	94.43 (*Vibrio splendidus* strain BST398)	Vibrioferrin (100 %)	✓	-
Bacteriocin	21.4	88.90 (*Vibrio splendidus* strain BST398)	None	-	-
	11	93.74 (*Vibrio alginolyticus* strain K08M4)	None		
PK				✓ ✓	-
Aryl polyene	42.0	84.29 (*Vibrio chagasii* strain ECSMB14107)	APE Vf (85 %)	-	-
betalactone	26,9	89.75 (*Vibrio splendidus* strain BST398)	None	-	-

PUFA, hglE-KS	55.8	91.94%*Vibrio cyclitrophicus* strain ECSMB14105	Eicoseicosapentaenoic acid (50 %)	-	-

The identified bacteriocin genes showed 100 % similarities to clusters from other *Vibrio* strains based on antiSMASH analysis. However, none of the BGCs have been characterised yet. Bacteriocins BGCs are common to *Vibrio* species such as *Vibrio tasmaniensis*, *Vibrio lentus*, *Vibrio atlanticus*, *Vibrio cyclitrophicus* and *Vibrio alginolyticus*. However, Bagel 3 and 4 analysis (Table 3.7, Appendix VI) could not corroborate the identification of bacteriocins in this genome sequence. To validate whether antiSMASH predicted the bacteriocins correctly, the predicted protein sequences of the BGCs were analysed using the NCBI BLASTp. Further sequence analysis provides some evidence to suggest that both putative bacteriocin BGCs were correctly predicted by antiSMASH as encoding bacteriocins since they contained conserved APP, DUF692 (Appendix VI) and PepP domains.

The APP domain, also known as "X-Pro aminopeptidase" catalyzes the release of the N-terminal amino acid, while the PepP domain is responsible for the transport and metabolism of amino acids. The peptidase is required for cleaving off the precursor (peptide chain) during the biosynthesis of bacteriocins (Arnison et al., 2013). The DUF protein family has been associated with circular bacteriocins (Collins et al., 2017; Gabrielsen et al., 2014). Based on the above information, we concluded that antiSMASH correctly identified BGCs as a bacteriocin. The fact that Bagel 3 and 4, are bioinformatics tools specifically aimed at identifying bacteriocins and did not identify any bacteriocin BGC within the PE14-63 genome demonstrates that these tools are not perfect. Encouraging the use of multiple tools to correctly assess the full biosynthetic potential of these microorganisms.

It is worth mentioning that misidentification of a bacteriocin by antiSMASH was observed during the genome mining of *Pseudoalteromonas ruthenica* by Machado and co-workers, whereby the BGC was identified by antiSMASH as a bacteriocin, but it encoded for a flagella operon. The genes within the cluster encoded two hypothetical proteins, namely a muramoyltetrapeptide carboxypeptidase and a 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase, involved in the biosynthesis of peptidoglycan and lysine (Machado et al., 2015). This

misidentification can be due to the different algorithms used by various programs detecting low similarity between these proteins and known bacteriocin encoding proteins. Thus, with the suite of current tools a level of manual curation is still called for to prevent under- or over-estimation of the genetic potential of an organism and avoid algorithm or database-related errors.

The identification of a ω-6 polyunsaturated fatty acid (PUFA) pathway suggests that PE14-63 could have the potential to produce compounds that can be used as a precursor for an anti-inflammatory agent. The detection of a PUFA pathway comes as no surprise as they are widely distributed in marine bacteria (Chenard and Lauro, 2017). PUFAs are ecological participators in assisting microorganisms in dealing with low temperature, high pressure and salinity in the environment (DeLong and Yayanos, 1986; Nichols, 2003).

The predicated pathway for the production of an aryl polyene (APE) was identified. APEs are yellow pigments, specialized polyunsaturated carboxylic acids, with anti-oxidant properties and function similarly to carotenoids (Kunwar and Priyadarsini, 2011; Kurutas, 2016). APE BGCs have been identified as the most widespread family of bacterial BGCs (Cimermancic et al., 2014; Johnston et al., 2021). APEs are present in most major bacterial genera; however; despite being produced by diverse bacterial genera they share a remarkably similar chemical scaffold (Cimermancic et al., 2014). APEs have been shown to protect bacteria from reactive oxygen species. In a study, Poplawsky and co-workers showed that APE (xanthomonadin) protected *Xanthomonas campestris* pv. *campestris* against photooxidative damage whereby a mutant APE *X. campestris* pv. *campestris* did not survive under normal light (Poplawsky et al., 2000). The PE14-63 colony was cream in colour during our study suggesting that the APE BGC is cryptic or being expressed in very low quantities to be visible on an agar plate or broth.

The identified betalactone BGC genes showed 36 % similarity to genes from *Pseudoalteromonas rubra* strain OCN096 NRP BGC and showed no similarity to any characterised gene cluster. However, the entire cluster BLASTn analysis revealed that the cluster is 89.75 % similar to a *Vibrio splendidus* sequence. Based on the low gene similarity PE14-63

produces a different compound to *Pseudoalteromonas rubra* strain OCN096. Therefore, studies are required to establish if PE14-63 betalactone BGC produces a bioactive compound against any of our test strains and to characterise the structure.

PRISM analysis (Table 3.7, Appendix VI) of PE14-63 detected two PK clusters, with the second cluster containing multiple thiolation domains (also known as acyl carrier proteins), which antiSMASH failed to predict. To validate the authenticity of the domains in clusters I and II we conducted an NCBI BLASTp analysis which showed similarity to keto-ACP-synthase III (100 %) and 3 hydroxyl-ACP-dehydratase (97 %), respectively. This result again indicates the importance of using multiple bioinformatics tools and NCBI BLAST analysis for the identification of possible pathways. Analysis of bacterial genomes on multiple platforms will minimise the over or underestimation of the potential of bacteria to produce SM. For example, in our study antiSMASH didn't predict any PK cluster while PRISM predicted two clusters therefore if only antiSMASH was we would have concluded that there are no PK clusters within the PE14-63 genome. Additionally, these could have resulted in PE14-63 being overlooked in future studies interested in studying PKs clusters.

PRISM also detected a NRP-independent siderophore (Appendix VI), synthesised by a NRPS-independent siderophore (NIS) synthase. The enzymes are classified based on the precursors used for the amide bond formation (Årstøl and Hohmann-Marriott, 2019). The PE14-63 synthase is classified as a NIS type B by PRISM suggesting that it catalyses the formation of a peptide bond between a hydroxamate-amine with an α-ketoglutaric acid.

PE14-63 showed the highest similarity to *V. tasmaniensis* at 16S rRNA sequence level. Therefore, a small-scale assessment of BGC distribution within the sequenced *V. tasmaniensis* genomes was conducted to compare classes of SM present within their genomes. Several *V. tasmaniensis* strains genomes from the NCBI database were analysed using antiSMASH and compared with PE14-63 results to determine the BGCs distribution within the species (Table 3.8). The analysis revealed that aryl polyene, bacteriocin and betalactone BGCs are present within all analysed genomes. The

presence of these classes could indicate their importance in the survival of the organism in the environment as with the bacillibactin cluster within the genus *Bacillus*. The NRP clusters showed the lowest degree of similarity (data not shown), and further analyses would have to be conducted to correlate their diversity and distribution among this species. The detection of NRP clusters with low sequence similarity within the genomes of several of the *V. tasmaniensis* species indicates the potential to discover novel peptides from this species.

Interestingly, some strains contain an ectoine pathway while others do not. Ectoine is essential in saline environments as it assists in dealing with osmotic and cold pressure (Kuhlmann et al., 2011; Reshetnikov et al., 2011). Due to its ability to stabilise organic molecules, it is very important in the pharmaceutical and cosmetic industry. PUFA has been linked to assisting marine microorganisms with environmental stress such as hydrostatic pressure and low temperature (Kato and Nogi, 2001; Nichols, 2003).

Table 3. 8. Comparison of SM biosynthesis gene clusters amongst genomes of *V. tasmaniensis* strains downloaded from NCBI using antiSMASH.

Secondary metabolite	PE14-63	1A01	BST398	10N.222.45. A8	ORI231	LGP32	ZS-17	UCD-FRSSP16_2	LMG 20012	DSM_19640	1F-267
Siderophore	✓	x	✓	x	✓	✓	x	x	✓	✓	x
Bacteriocin	✓✓	✓✓	✓	✓	✓✓	✓	✓✓	✓✓	✓	✓✓	✓
Ectoine	x	x	x	✓	x	✓	x	x	✓	x	✓
Arylpolyene	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NRP	x	✓	x	✓	x	x	✓	✓	✓	x	x
PUFA-hglE-KS (hybrid)	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓
betalactone	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NRPS- betalactone (hybrid)	x	x	x	✓	x	x	x	x	x	x	x

Legend: number of ticks indicate the total number of identified BGC within the genome.

4. Chapter 4: Isolation and characterisation of SM

4.1.Extraction, isolation and characterisation of natural products.

The marine isolates *Pseudovibrio* sp. PE14-07, *Bacillus* sp. PE08-149B and *Vibrio* sp. PE14-63 were selected for further studies to characterise and identify the diverse anti-microbial activity observed when fermented on GYM media (Table 4.1). This specific growth condition (GYM media) was chosen since it elicited anti-microbial activity for all three isolates. The initial fractionation was performed using amberlite IR120, amberlite IRA 910 and ethyl acetate (EtOAc) after a 14 day fermentation period, to cover multiple chemical properties of SMs. The utilisation of different purification systems is essential as SMs are known to have very diverse chemical properties (e.g., non-polar, polar, charged, etc.).

Both amberlite IR120 and amberlite IRA 910 resins bind compounds using ionic exchange principles. The ionic exchange happens between two phases which depends on the concentrations and affinity of the ions to the immobilised ions (Alexandratos, 2009). The use of ion exchange as SMs separation method dates back to 1945 whereby it was showcased by the separation of thiamine from riboflavin (Herr, 1945). Amberlite IR20 has been previously used to recover nicotine (de Lucas et al., 1998) and amberlite IRA 910 was used to recover SMs classified as “organic acids” from water (Desauziers et al., 1998). Ionic resins were used to purify violacein and indolmycin from the bacterial broth of marine *Pseudoalteromonas luteoviolacea* bacteria (Dragull and Beck, 2012). Additionally, ionic exchange resin can be regenerated making them economical.

Ethyl acetate is commonly used in the separation of bioactive SMs from the fermentation broth and separates compounds based on polarity. Ethyl acetate has been shown to extract drug-like

molecules compared to other solvents such as MeOH, acetone and chloroform (Chairman et al., 2012; Skariyachan et al., 2014). Compared to other solvent extraction methods, ethyl acetate demonstrated suitable extraction properties (Kiran et al., 2018). Different classes of SMs have been isolated using the EtOAc solvent system such as lipids, peptides, quinones, lipopeptides and alkaloids (Skariyachan et al., 2014). In this study, the anti-microbial diversity of the extracted SMs was evaluated using well-diffusion bioassay with 30-80 % DMSO and reverse osmosis (RO) water was used to resuspend the fractions and as negative controls. DMSO was introduced as a solvent as some extracted compounds were not soluble in water.

Table 4.1 presents a summary of the anti-microbial results before and after fractionation with different methods. These results suggest the presence of different anti-microbial compounds as different anti-microbial profiles were observed under multiple purification strategies. When considering the genome analyses presented in Chapter 3 and taken together with the bioassay results obtained, it is not entirely surprising that multiple compounds conferring anti-bacterial activity were produced. For example, the presence of multiple compounds was observed when PE14-07 amberlite IR120 extract showed activity against the multidrug-resistant *E. coli* 1699, *S. epidermis, B. cereus, M. aurum, A. fumigatus* and *P. putida* while the amberlite IR910 extract showed activity against *S. epidermis, B. cereus* and *A. fumigatus*. This finding suggests that at least two different compounds were present in the fermentation broth as the resins (amberlite IR120 and amberlite IR910) display different anti-microbial profiles (Table 4.1 and Appendix VIII).

An interesting phenomenon was observed for PE14-63 after extraction with the cationic beads, whereby "new" activity was observed against *P. putida* and *M. aurum* as reported in Table 4.1. It could be that the active compounds were masked with the presence of large amounts of salts or an inhibiting compound from the media, an effect observed in a study by (Cutignano et al., 2015). In our study, the amberlite IR120 proved superior in the extraction of bioactive SMs, followed by EtOAc with amberlite IRA910 being the least effective. The cationic amberlite resins extracted different compounds with a broad anti-microbial spectrum from the marine strains.

Fractionation of the PE08-149B GYM fermentation broth using amberlite IR910 resin showed no anti-microbial activity when tested against all the test strains. However, the amberlite IRA120 and EtOAc crude extracts displayed anti-microbial activity against both Gram-negative and Gram-positive strains with no anti-fungal activity when using well diffusion assays. The absence of the anti-fungal activities in crude extract could be due to the decreased sensitivity of the well-diffusion assays compared with the HTS anti-fungal assays performed by our collaborators.

The expression of BGCs is highly regulated in microbes and can be influenced by several factors such as light, temperature, competition and growth stage (Scherlach and Hertweck, 2009). In our study, culturing was conducted on a shaker at ambient temperature. Since this was not constant it could have positively or negatively influenced the expression of BGCs in our strains. A room with controlled parameters is ideal for this kind of work as these microbes are sensitive and it will help in understanding / studying the effect of each parameter. Other parameters which could have influenced the difference are culturing vessels (McCartney bottles vs flasks), media volume (10 ml vs 250 vs 500 ml), fermentation batches and different reagent suppliers.

Table 4. 1. Anti-microbial results for crude extracts of PE14-63, PE14-07 and PE08-149B after extraction from GYM media fermentations.

Sample / extract ID	*E. coli* 1699	*P. putida*	*S. epidermis*	*B. cereus*	*M. aurum*	*A. fumigatus*	*C. albicans*
Crude extract PE14-07	Yes	Yes	Yes	Yes	Yes	Yes	No
PE14-07 amberlite IR910	No	No	Yes	Yes	No	Yes	No
PE14-07 amberlite IR120	Yes	Yes	Yes	Yes	Yes	Yes	No
PE14-07 EtOAc	Yes	Yes	Yes	Yes	Yes	No	No
Crude extract PE14-63	Yes	No	Yes	Yes	No	No	No
PE14-63 Amberlite IR910	No	No	No	Yes	No	No	No
PE14-63 amberlite IR120	Yes	Yes	Yes	Yes	Yes	No	No
PE14-63 EtOAc	No	Yes	Yes	No	No	No	No
Crude extract PE08-149B	Yes	Yes	Yes	Yes	Yes	No	No
PE08-149B amberlite IR910	No	No	No	No	No	No	No
PE08-149B amberlite IR120	Yes	Yes	Yes	Yes	Yes	No	No

PE08-149B EtOAc	Yes	Yes	Yes	Yes	Yes	No	No

Legend: Yes, means anti-microbial activity was observed against the test strain

(a clear distinctive zone around the well was regarded as anti-microbial when compared to the negative control)

No, means the absence of anti-microbial activity against the test strain

(No was interpreted as growth around the well)

After anti-microbial profiling, PE14-07 and PE08-149B were selected for chemical characterisation as they showed desired diverse bioactivity compared to the PE14-63 strain. Genome mining also showed the presence of more BGCs within the genomes of the two strains (PE14-07 and PE08-149B). Additionally, the BGCs identified in the PE14-07 genome showed low similarity to known BGCs, increasing the chance of identifying new SMs. Literature has also highlighted the lack of studies on identification of SMs in *Pseudovibrio* spp. despite the reported anti-microbial activities (Naughton et al., 2017). Despite PE08-149B BGCs showing high similarity to some characterised BGCs, majority of the clusters showed lower similarities to characterised or known clusters (Table 3.6).

Invertebrate associated *Pseudovibrio* sp. have been established as potential producers of SMs as early as 2008 by Kennedy *et al*. using PCR and bioassays (Graça et al., 2013; Kennedy et al., 2009). However, to our knowledge, to date only a few compounds such as TDA, thiotropocin, heptyl prodigiosin and pseudovibriocin have been isolated from *Pseudovibrio,* with bioactivities attributed to them (Crowley et al., 2014; Raina et al., 2016; Vizcaino, 2011). Other SMs isolated from *Pseudovibrio* include fistularin-3, phenazine, 11-hydroxyaerothionin, verongidoic acid, aerothionin, homopurpuroceratic acid, purealidin L and aplysinamisine II (Nicacio et al., 2017). All mentioned metabolites have shown bioactivity in other studies, for example, fistularin -3 and 11-hydroxyaerothionin have shown bioactivity against Feline leukaemia virus and multiple microbes (*Staphylococcus aureus*, *Bacillus subtilis*, *Candida albicans* and *Mycobacterium tuberculosis*), respectively. Studies by (Gunasekera and Cross, 1992; Kernan et al., 1990) were the first to prove bacterium biosynthesis of these metabolites. The metabolites (fistularin -3 and 11-hydroxyaerothionin) were previously isolated from the marine invertebrates suggesting *Pseudovibrio* as the true producer of these metabolites (Carroll et al., 2019; Naughton et al., 2017). Only a few SMs from *Pseudovibrio* have been characterised to date despite the availability of genomic evidence showing the presence of diverse BGCs, indicating a pool of more SMs to be described.

Bacillus is the most studied of the three genera as a source of bioactive SMs. Secondary metabolites from different classes have been isolated such as lipopeptides, NRPs, PKs, polypeptides and RiPPs (Boottanun et al., 2017; Mondol et al., 2013). Compounds such as fengycin, bacillomycin, surfactin, difficidin and bacilysin have been isolated from the genus *Bacillus* (Boottanun et al., 2017). antiSMASH analysis demonstrated the presence of a fengycin-like BGC within PE08-149B genome with 40 % similarity suggesting production of different fengycin forms or completely different metabolites. antiSMASH also showed four bacteriocins, one terpene and one NRP BGC with no similarity to known clusters (Chapter 3, Section 3.3.3), indicating a potential source of new compounds.

4.1.1. Chemical- and bioassay-guided fractionation

Crude extracts were fractionated using bioassay-guided purification. The bioassay-guided fractionation made use of well-diffusion (test strains; *E. coli* 1699, *S. epidermis*, *B. cereus*, *P. putida*, *M. aurum*, *A. fumigatus* and *C. albicans*), SPE and HPLC methods. Chemically guided fractionation was accomplished with the use of SPE and NMR (400 / 600 MHz) to characterise potentially new SMs without focusing on anti-microbial activity. NMR was used to identify which fraction to focus on, based on which functional groups and splitting patterns were observed. HRLC-MS was used for the identification of metabolites. The purified metabolite was then tested for bioactivity using the above-mentioned test strains. PE14-07 was also analysed for the presence of TDA using LC-MS, as TDA biosynthesis genes were detected within the genome of PE14-07.

4.1.2. Isolation and purification of PE14-07 secondary metabolites

Solid-phase fractionation of the PE14-07 EtOAc extract yielded four fractions (PVET-SPE-25, -50, -100 and -100 TFA), and their anti-microbial activity was determined using a well-diffusion assay. The PVET-SPE-100 fraction retained the bioactive compounds as it gave activity against *E. coli* 1699, *P. putida*, *S. epidermis*, and *B. cereus*. The other fractions displayed mild to no activity. The anti- *A. fumigatus* activity was lost after fractionation, which could be due to two or

more compounds acting synergistically or the breakdown of the anti-fungal active compound in the column. The loss of activity after pre-fractionation was also observed in a study by Hoppers and co-workers, whereby anti-microbial activity against *E. coli* and *P. aeruginosa* disappeared after fractionation of the MeOH extract generated from a marine sponge *Haliclona* sp. (Hoppers et al., 2015). Due to the outcome of the bioassay, the PVET-SPE-100 fraction was subjected to further purification using RP-HPLC.

The EtOAc extract was also subjected to fractionation prioritisation using NMR. Out of the four SPE generated fractions, the PVET-SPE-100 extract spectrum showed peaks in the aromatic region (Figure 4.1) and was selected for further fractionation by size exclusion chromatography. The size exclusion chromatography resulted in PVET-SPE-100-S11-13 fraction with a high concentration of the aromatic compounds. The fraction was then purified with RP-HPLC (University of Aberdeen) to yield a pure fraction of compound 1 (described in Section 4.3). The prioritisation of the fraction with ^{1}H NMR peaks in the aromatic region was made since many SMs (organic compounds) contain aromatic rings and this particular region of the spectrum often suffers much less from signal overlap than the aliphatic region (Figure 4.2).

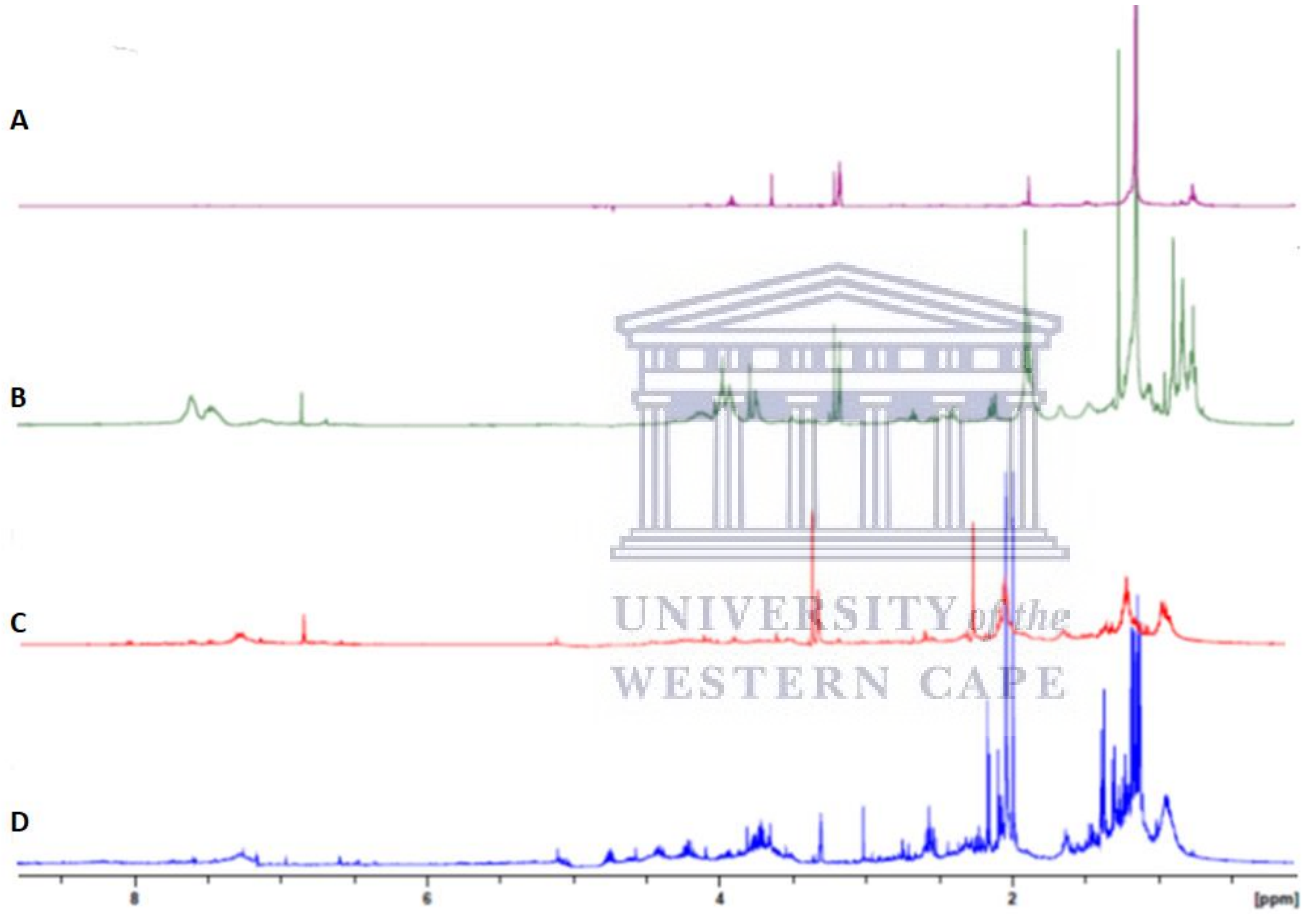


Figure 4. 1. ^{1}H NMR spectra of PE14-07 PVET-SPE fractions in MeOH-d_4 at 400 MHz. Labels A: PVET-SPE-100 TFA, B: PVET-SPE-100, C: PVET-SPE-50 and D: PVET-SPE-25

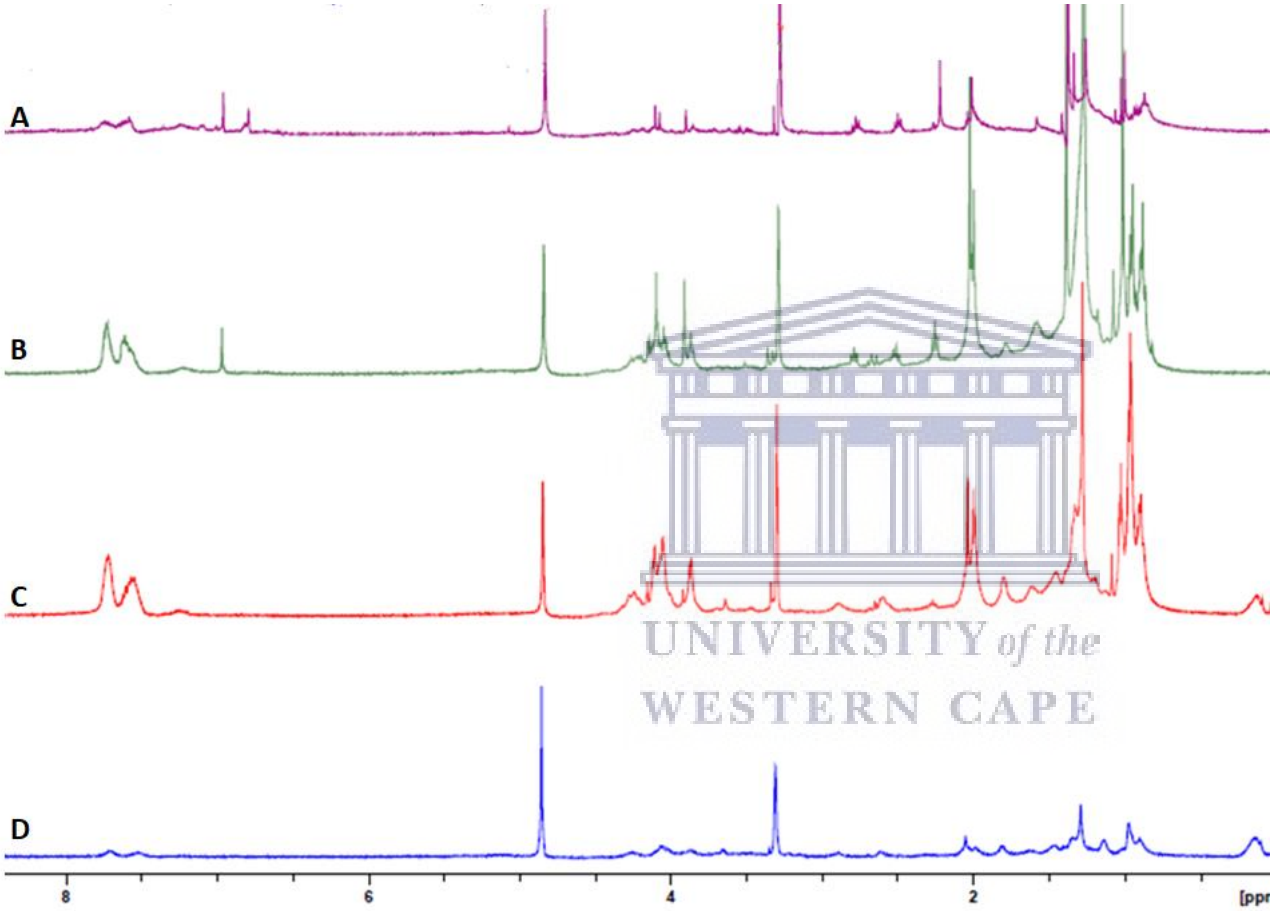


Figure 4. 2. ^{1}H NMR spectra of PE14-07 PVET-SPE-100 fractions after size exclusion chromatography MeOH-d_4 at 400 MHz. Labels A: PVET-SPE-100 S14-17, B: PVET-SPE-100 S11-13, PVET-SPE-100 S10-8 and PVET-SPE-100 S1-7

4.1.3. PE14-07 identified compounds.

Compound 1 (Bis(3-acetyloxy)-bis-(2,2-dimethyl propyl benzoate / Algoa PV1):

The PVET-SPE-100 S11-13 HPLC purification resulted in the isolation of Algoa PV1 (Figure 4.3). The structure of the compound was elucidated with the use of 1D and 2D NMR data (Table 4.2), while the molecular formula was determined using HRMS data. Compound **1** was isolated as a yellowish solid material, with a molecular formula of $C_{22}H_{30}O_8$ as determined by HRMS of the protonated ion (*m/z* 423.20). The compound has a double bond equivalent (DBE) of 8, indicating unsaturation, and has strong UV-VIS absorbance at 230 nm and 279 nm. The NMR Data spectra are reported in Appendix III (for this compound and others reported in the following sections), while HRLC-MS spectra are provided in Appendix IV.

Based on the NMR analysis, the compound was symmetrical (see the following discussion) with NMR indicating the presence of three methyl groups (thus a total of six methyl groups). The compound was established to feature a 1,2-disubstituted aromatic ring, based on a J constant of 5.76 Hz observed for the two aromatic proton signals. Moreover, the spectrum displayed two methyl singlets at δ 1.06 and 2.07, indicating that they were attached to a quaternary carbon. Two singlets belonging to methylene groups were also observed at 4.13 and 3.95. The total number of protons were half the total number of protons required by the molecular formula which led to the conclusion that the compound is symmetrical. The only COSY correlations which were observed were from the benzene ring resulting in limited structural information gathered, rendering COSY usefulness limited in this case. Due to this limitation, HMBC provided vital information in elucidating the chemical structure of compound 1 (Figure 4.4).

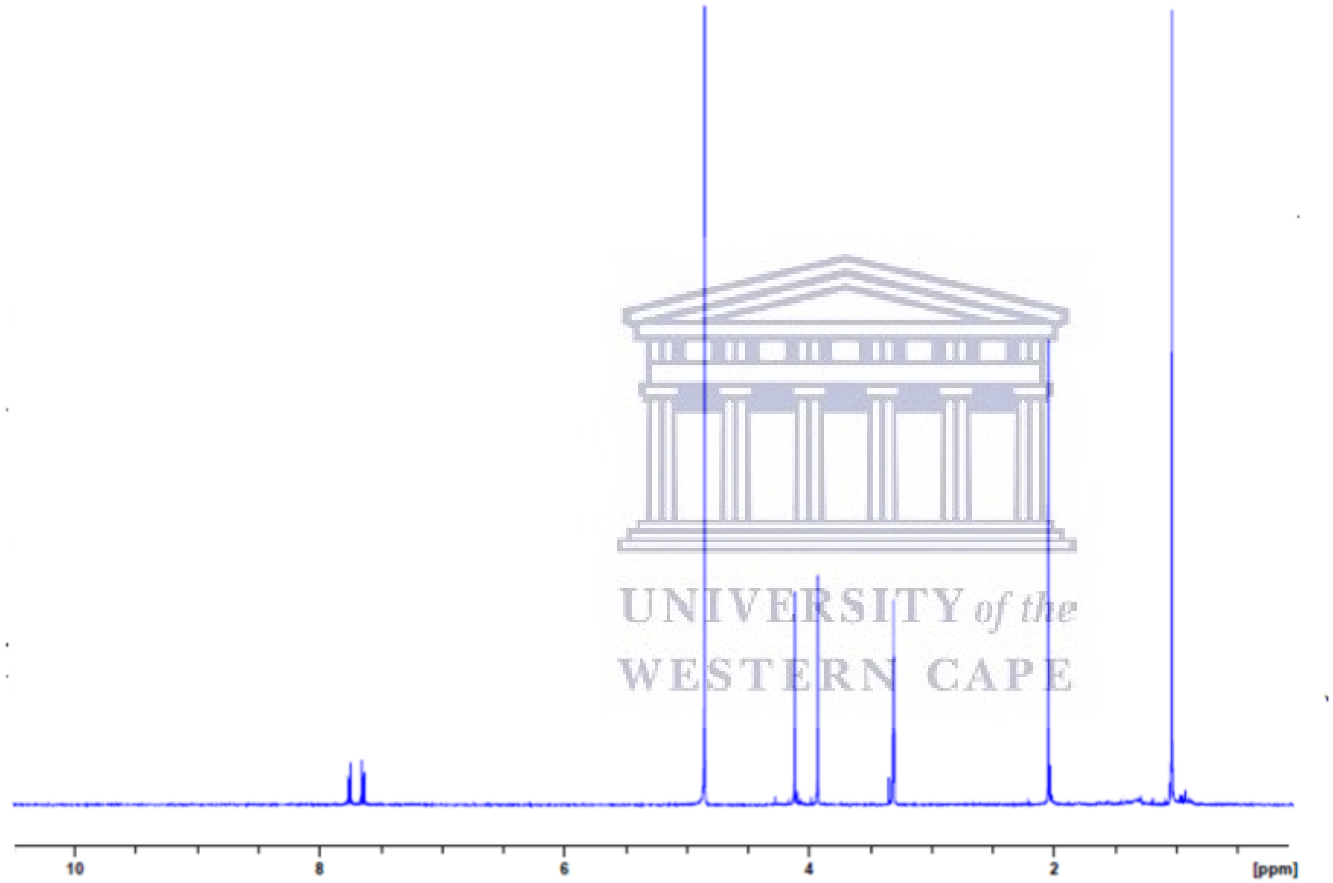


Figure 4. 3. ^{1}H NMR spectrum for compound 1 in MeOH-d_4 at 400 MHz.

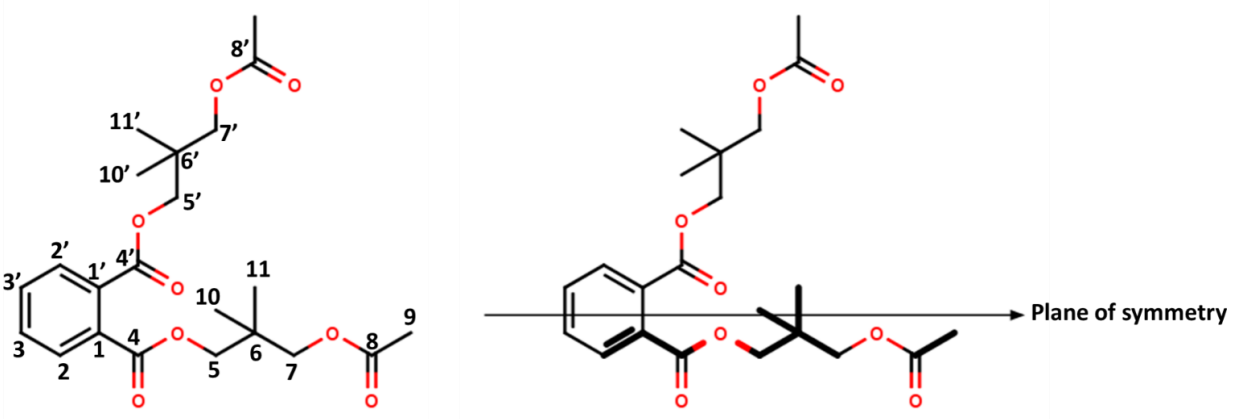


Figure 4. 4. Chemical structure of Algoa PV1 and HMBC correlation (Represented by black bold bonds).

The structure of the compound was searched on SciFinder, MarinLit, and PubChem but these searches did not produce a match, indicating the potential of a new compound. To affirm the new compound the structure needs to be searched in other databases such as Dictionary of Natural product and Antibase. The compound was also evaluated for anti-microbial activity against all test strains (*Escherichia coli* 1699, *Bacillus cereus* ATCC10702, *Pseudomonas putida* ATCC12633, *Mycobacterium aurum* A^{+}, and *Staphylococcus epidermidis* ATCC14990); however, no bioactivity was found against any of our test strains.

Table 4. 2. NMR Spectroscopic Data (400 MHz, MeOH-d_4) for Algoa PV1.

Position	δ ^{13}C, type	δ^{1}H (J in Hz) multiple
1 /1'	132, C	
2/2'	129, CH	8, 1H, dd
3/3'	131, CH	8, 1H, dd
4 / 4’	168, C	
5 / 5’	69, CH_2	4, 2H, s
6 / 6’	34, C	
7 / 17’	70, CH_2	4, 2H
8 / 8’	171.4, C	
9 / 9’	19, CH_3	2, 3H, s
10 / 10’	21, CH_3	1, 3H, s
11 / 11’	21, CH_3	1,3H, s

Legend: δ, indicate a chemical shift in reference to the NMR solvent.

Compound 2 (Algoa PV2):

The amberlite MeOH PE14-07 extract was further fractionated into four fractions using reverse phase-SPE, resulting in the PVM-SPE-100 fraction being active against *B. cereus* (Figure 4.5A). This fraction also displayed iron chelating activity (Figure 4.5D) when assessed by the CAS media well-diffusion assay. HPLC bioassay-guided purification of the PVM-SPE-100 fraction resulted in three neighbouring peaks which were yellowish and contained a compound that conferred anti-*B. cereus* activity. Since a siderophore gene cluster was predicted by the antiSMASH analysis of the PE14-07 genome, it is tempting to speculate that compound 2 is the product of this BGC. Siderophores are iron chelating compounds with diverse bioactivities such as anti-microbial, anti-parasitic and anti-cancer (Årstøl and Hohmann-Marriott, 2019; Mansson et al., 2011). In order to confirm that the BGC cluster is responsible for the anti-*B. cereus* and siderophore activity gene knock-out studies or assay guided purification of the active compound is required.

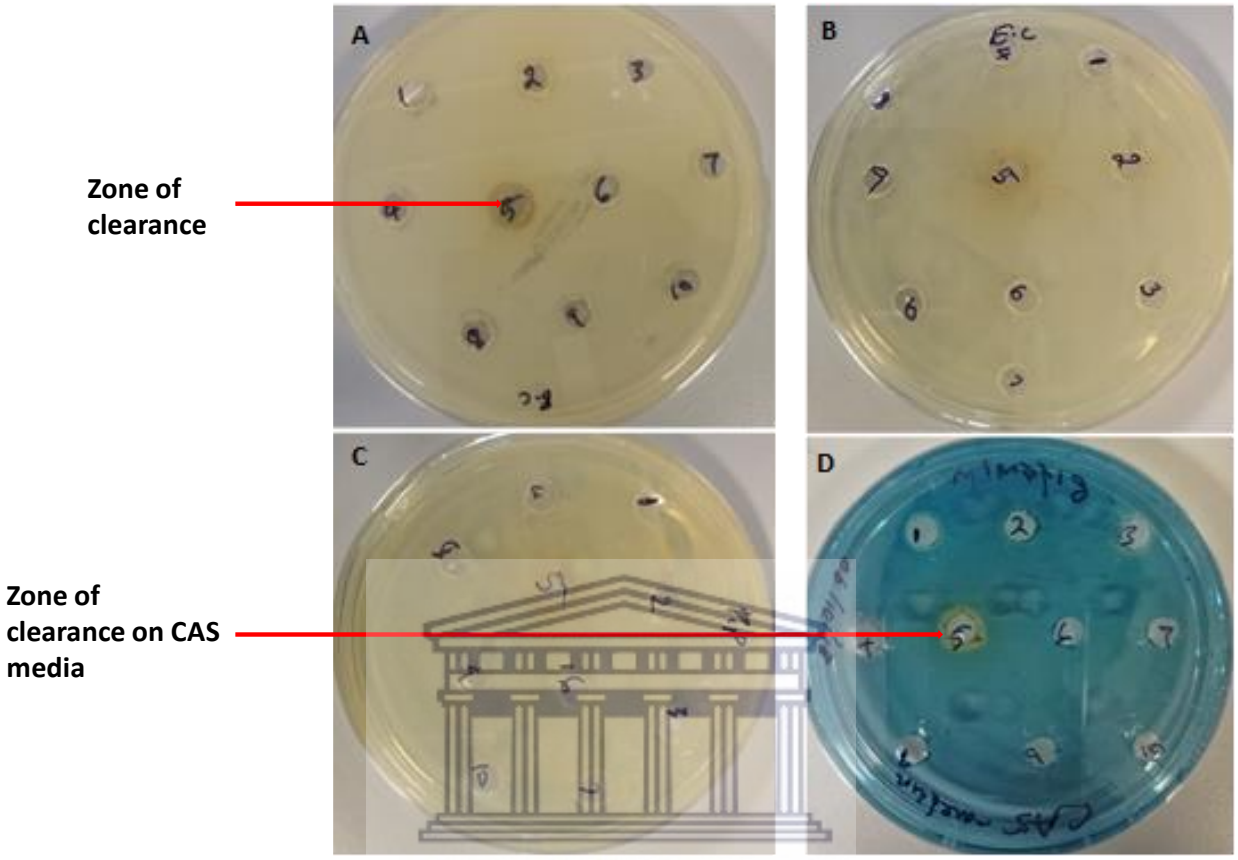


Figure 4. 5. Anti-microbial assay of PE14-07 SPE fractions and siderophore assay with anti-bacillus A: *B. cereus*, B: *E. coli* 1699, C: *P. putida* and D: CAS agar siderophore assay.

The PVM-SPE-100 fraction was subjected to HPLC (Figure 4.6) whereby fractions were collected, and three fractions showed both bioactivity and CAS activity. The HRMS data analysis of fractions revealed a common pseudomolecular ion ($[M+H]^+$) at *m*/*z* 421.2332 with a mass accuracy / error of 0.5 ppm in all three fractions, suggesting a compound with a molecular formula of $C_{22}H_{32}N_2O_6$ and a DBE of 5 (Figure 4.7). The compound was speculated to be responsible for the observed bioactivity since it was common amongst three neighbouring bioactive HPLC fractions. The SciFinder database search resulted in three compounds matching the formula; however, none of them had references to enable a comparison to be made. PubChem search resulted in over 1110 providing possible future reference compounds. However, the molecular formula search on the MarinLit database did not identify a match, indicating a potential new marine compound.

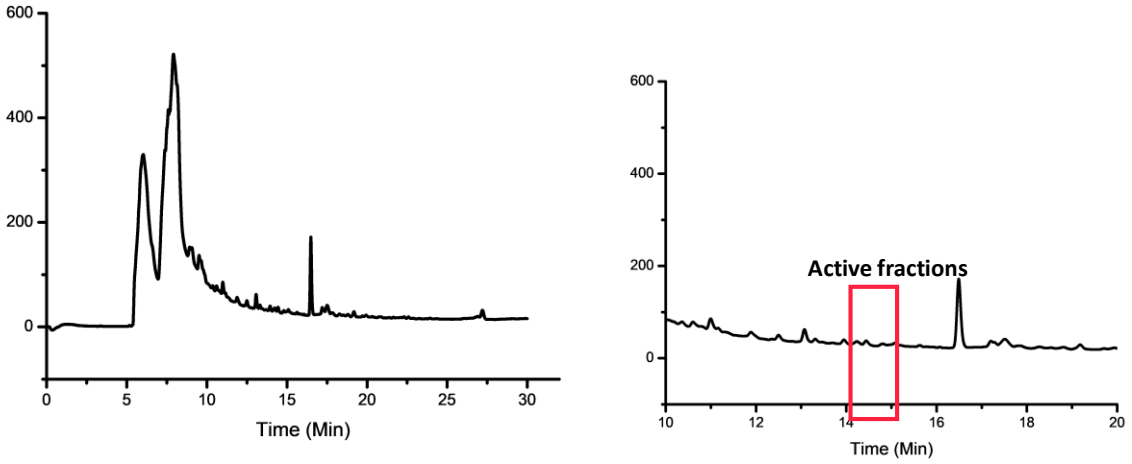


Figure 4. 6. HPLC spectrum of PVEM-SPE-100 highlighting peaks active against *B. cereus*.

Figure 4. 7. HRMS spectrum of the common mass (Algoa PV2) on positive ionisation

Due to insufficient amounts purified, we were unable to structurally identify Algoa PV2 using NMR experiments. The insufficient production of SMs by microorganisms is another hindering factor in SM discovery. More studies are required in understanding factors governing the production of SMs in microorganisms in general (van der Meij et al., 2017).

Compound 3 (Algoa PV3):

Tropodithietic acid (TDA) is a commonly produced compound by bacteria belonging to the *Rhodobacteraceae* family which the *Pseudovibrio* genus belongs to (Harrington et al., 2014). After detecting genes responsible for TDA biosynthesis within the PE14-07 genome, we evaluated its production in GYM. After fermentation of PE14-07 in GYM media, a TDA specific extraction was conducted (Porsby et al., 2011). TDA presence was evaluated using HPLC with a UV-VIS detector to characterise the compound. The HPLC chromatogram and associated UV spectrum for the major peak obtained (Figure 4.8) suggested the presence of TDA based on its signature wavelength absorbance at 300 and 356 nm. The TDA rich crude fraction was further subjected to HRLC-MS to analyse for the presence of TDA. The result indicated the presence of a molecular ion with *m/z* of 212.9680 further confirming the presence of TDA (Figure 4.9).

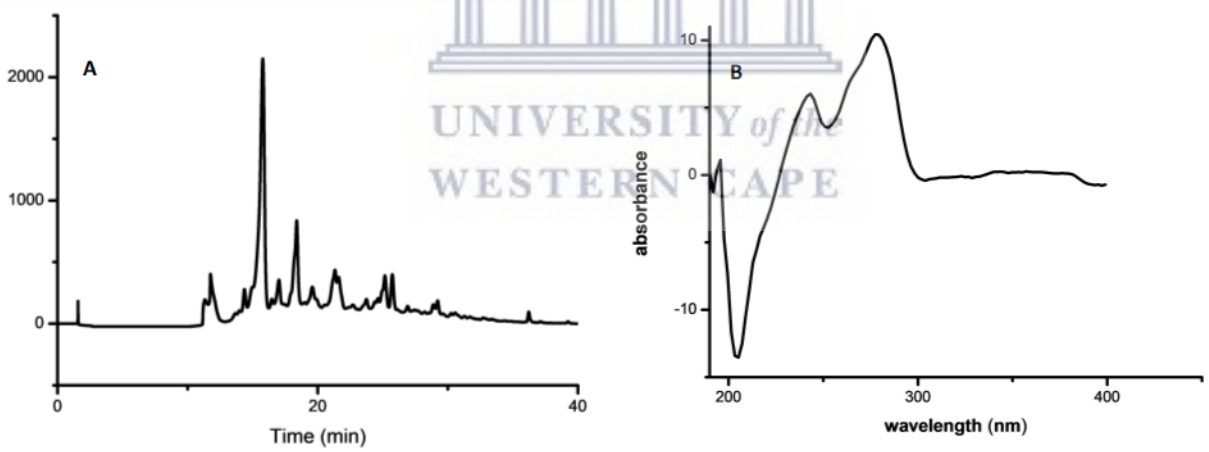


Figure 4. 8. HPLC chromatogram (A) and (B) UV spectra of the PE14-07 TDA-rich extract

The TDA fraction showed bioactivity against *E. coli* 1699, *S. epidermis*, *B. cereus*, *P. putida* and *M. aurum* from both marine broth and GYM medium. The bioactivities are similar to the one reported in literature except for *M. aurum* as there is no data, making our study the first to show anti-*Mycobacterium* activity. An interesting phenomenon was also observed in the production of TDA in PE14-07. Usually, the production of TDA is linked to the simultaneous production of a

brownish pigment in MB (Bruhn et al., 2005; Gram et al., 2010). A pigment was observed in our study in the extract prepared from MB fermentation (Figure 4.10). However, the pigment was not observed when TDA was produced by PE14-07 in GYM media. Harrington and co-workers have shown that TDA production is regulated differently as some species showed the production of TDA when grown on marine broth (MB) and not starch-yeast extract-peptone–seawater (SYP-SW) (Harrington et al., 2014). Due to different TDA regulatory mechanisms within the *Rhodobacteraceae* family (Berger et al., 2011), studies are required to establish whether the antiSMASH-identified Hserlactone BGC is active in the regulation of TDA synthesis in PE14-07. The pathway is responsible for the production of homoserine lactone compounds (e.g. 3OHC(10)-HSL) which have been shown to regulate TDA production in some bacteria (Berger et al., 2011; Harrington et al., 2014).

Assuming that the pigment produced by PE014-07 in MB is the same as the pigment previously identified to be linked with TDA production in other species, our study suggests that the genes responsible for TDA and pigment production are located on separate operons or are regulated differently (Figure 4.10). Perhaps there are genes within the TDA operon, or outside, that are downregulated in GYM and are responsible for pigment synthesis. The detection of TDA without the pigment in PE14-07 could therefore be useful in identifying the genes responsible for the biosynthesis of the pigment, for example by studying the expression of the TDA genes in GYM vs MB media.

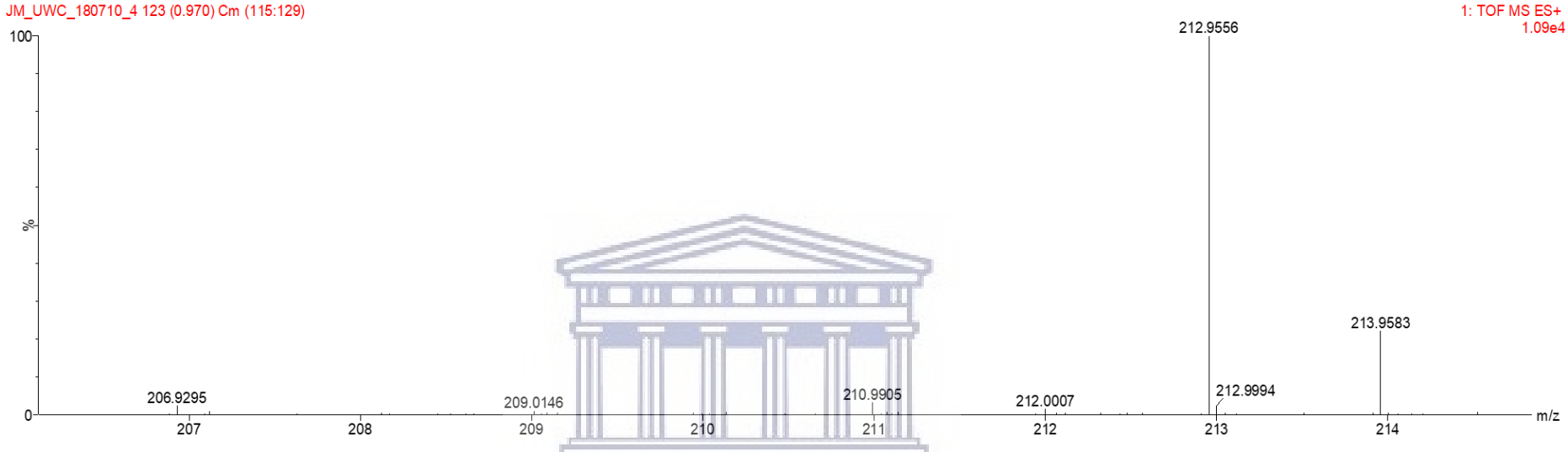


Figure 4. 9. HRMS spectrum of Algoa PV3

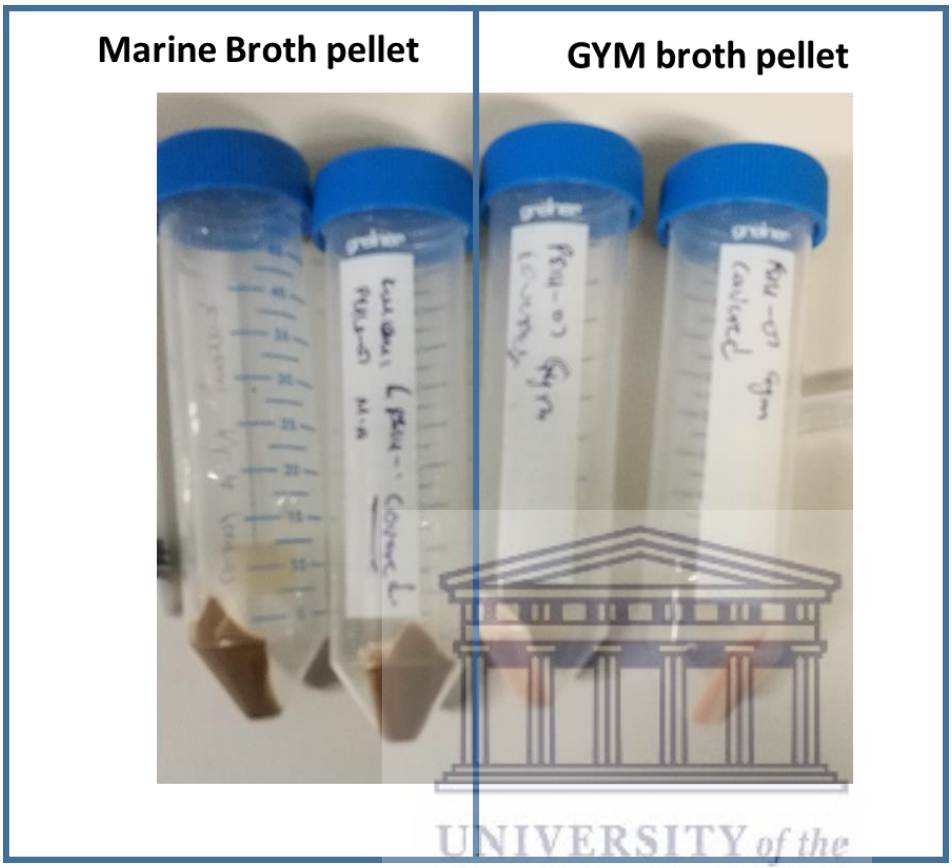


Figure 4. 10. Pelleted cells of PE14-07, illustrating brown pigment on marine broth and no pigment on GYM broth.

The presence of the *tda*A gene encoding a transcriptional regulator of the LysR family indicates that the protein is vital in the regulation of TDA synthesis despite it being suggested not to be the sole regulator of the pathway (Harrington et al., 2014). For example, *Pseudovibrio* sp. W64 and W74 produced more pigment and TDA when shaking (Harrington et al., 2014) while Belas and co-workers showed that TDA production increased in static culturing compared to when shaking (Sule and Belas, 2013).

Other studies have shown TDA is regulated by auto-induction and quorum sensing (compound utilised 3OHC (10)-HSL) in the *Roseobacter* clade. However, *Pseudovibrio* W74 indicated the presence of another regulation mechanism as it didn't conform to the known mechanisms (Geng

and Belas, 2010; Harrington et al., 2014). This shows diverse SMs regulation with different species and strains further highlighting that there is no "one solution fits all" scenario. In our study, TDA was produced without shaking. Future studies for PE14-07 should include studying the factors governing the expression of TDA such as light, shaking, fermentation vessel, quorum sensing and media components.

4.1.3.1.Hurdles encountered during purification.

The major challenge encountered during the purification of the metabolites was the production of enough material for chemical characterisation and further studies. This challenge was observed in fractions such as PVET-SPE-100 and other extracts where bioactivity was observed; however, the generated material was insufficient for NMR experiments. There is an urgent need for the optimisation of the *Pseudovibrio* sp. PE14-07 fermentation to ensure maximum SM yields.

4.1.4. Isolation and purification of PE08-149B secondary metabolites

Solid-phase fractionation of the PE8-149B extract resulted in the loss of activity with only the BW-SPE-25 fraction conferring weak bioactivity, therefore the SPE fractions were not subjected to further purification. The compounds could be strongly binding to the stationary phase creating difficulty in elution. As a result, liquid-liquid partitioning was used as a method for purification of SMs produced by PE08-149B. Liquid-liquid partitioning of the MeOH Amberlite IRA120 extracted SMs and bioassay shows a diverse polarity of the bioactive compounds as the active components were retained in the 1-butanol and DCM extracts, which were further subjected to HPLC for purification. The 1-butanol fraction yielded compounds active against *P. putida* while the DCM fraction retained activity against the other test strains (*E. coli* 1699, *B. cereus, S. epidermis* and *M. aurum*, Table 4.1).

4.1.5. PE08-149B identified compounds.

Compound 4 (Butyl-methyl-dicarbonate / Algoa BW1) was isolated as a brownish-yellowish solid using RP-HPLC (Figure 4.11). HRMS data (*m/z* 175.0607) obtained in negative ionisation mode suggested $C_7H_{12}O_5$ as the molecular formula, with the compound having a DBE of 2. The ^{1}H NMR spectrum showed no peaks downfield (above δ 4.2), ruling out the possibility of the compound containing phenyl (aromatic) and aldehyde functional groups (Table 4.4). The ^{1}H NMR spectrum did; however, display the presence of two methyl groups at δ 2.5 and 0.9 as a singlet and a triplet obtained in negative ionisation mode, respectively. The 2.5 methyl group is further downfield and is a singlet suggesting its position adjacent to oxygen. Three methylene groups were observed as triplet and multiplets, one of the methylene groups was further downfield suggesting that it is next to an oxygen.

COSY correlation showed the presence of a single substituted butane chain (Figure 4.12). The ^{13}C spectrum showed the presence of two ester carbonyl groups, and HMBC experiments revealed the attachment of the remaining methyl group to the molecule (Figure 5.12). The molecule showed anti-microbial activity against *P. putida* only. The compound structure was searched in SciFinder, MarinLit, Dictionary of NP and PubChem, but did not return any hits, suggesting that the compound may be new.

Table 4. 3. NMR Spectroscopic Data (400 MHz, MeOH-d6) for Algoa BW1.

Number	δ ^{13}C, type	δ^{1}H (J in Hz), multiple
1	14.0, CH_3	0.9, 3H, 7.4 Hz, s
2	19.0, CH_2	1.4. 2H, 7.5 Hz, m
3	29.2, CH_2	1.6, 2H, m
4	29.9, CH_2	4.0, 2H, 7.4 Hz, m
5	174.0, C	-
6	172.7, C	-
7	30, CH	2.6, 3H, 6.6 Hz, s

Legend: δ, indicate a chemical shift in reference to the NMR solvent.

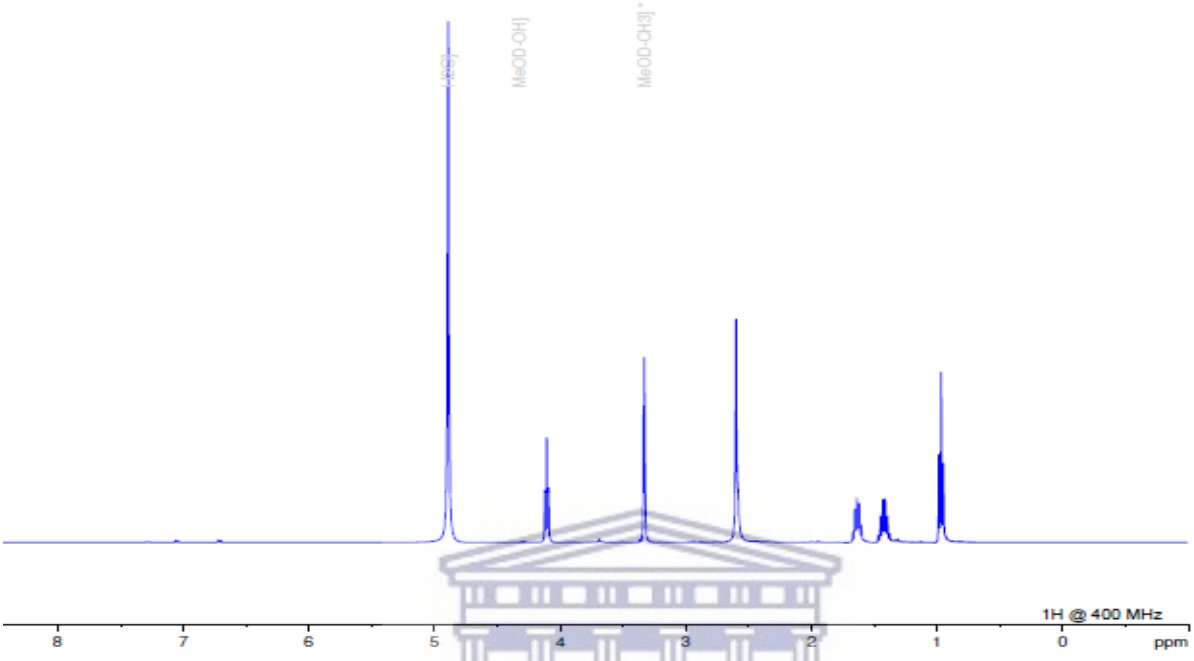
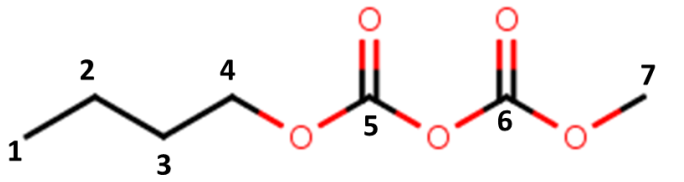
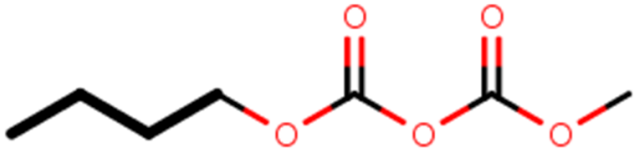
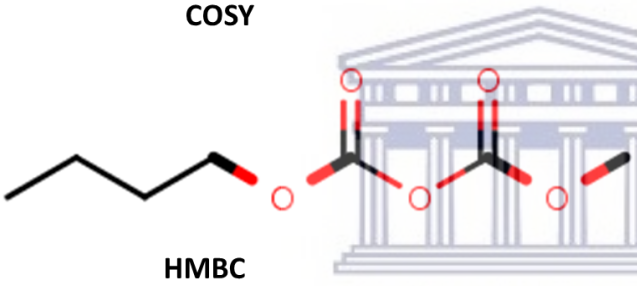


Figure 4. 11. ^{1}H NMR spectrum for compound 4 in MeOH-d_4 at 400 MHz.





COSY



HMBC

Figure 4. 12. COSY and HMBC correlation of Algoa BW1

Compound 5 (Algoa BW2) was isolated as a brownish-yellowish solid using RP-HPLC. The HRMS data m / z 277.1547 suggested that the molecular formula is $C_{15}H_{20}N_2O_3$. The compound has a DBE of 6. Unfortunately, due to the complexity of the structure and insufficient material for further experiments, only substructures were elucidated from the NMR data (Figure 4.12 and 4.13). The database search using monoisotopic mass and chemical formula resulted in compounds with NMR spectra that are different from the one generated in this study. Some difficulty was experienced with the presence of two CH_2 singlets between 2 and 2.5 ppm integrating for 6 and 2. Literature and database searches resulted in no compound with the same molecular formula linked to *Bacillus*, suggesting a new *Bacillus* bioactive SM, but it should be mentioned that it is also possible that the NMR fraction contained two compounds.

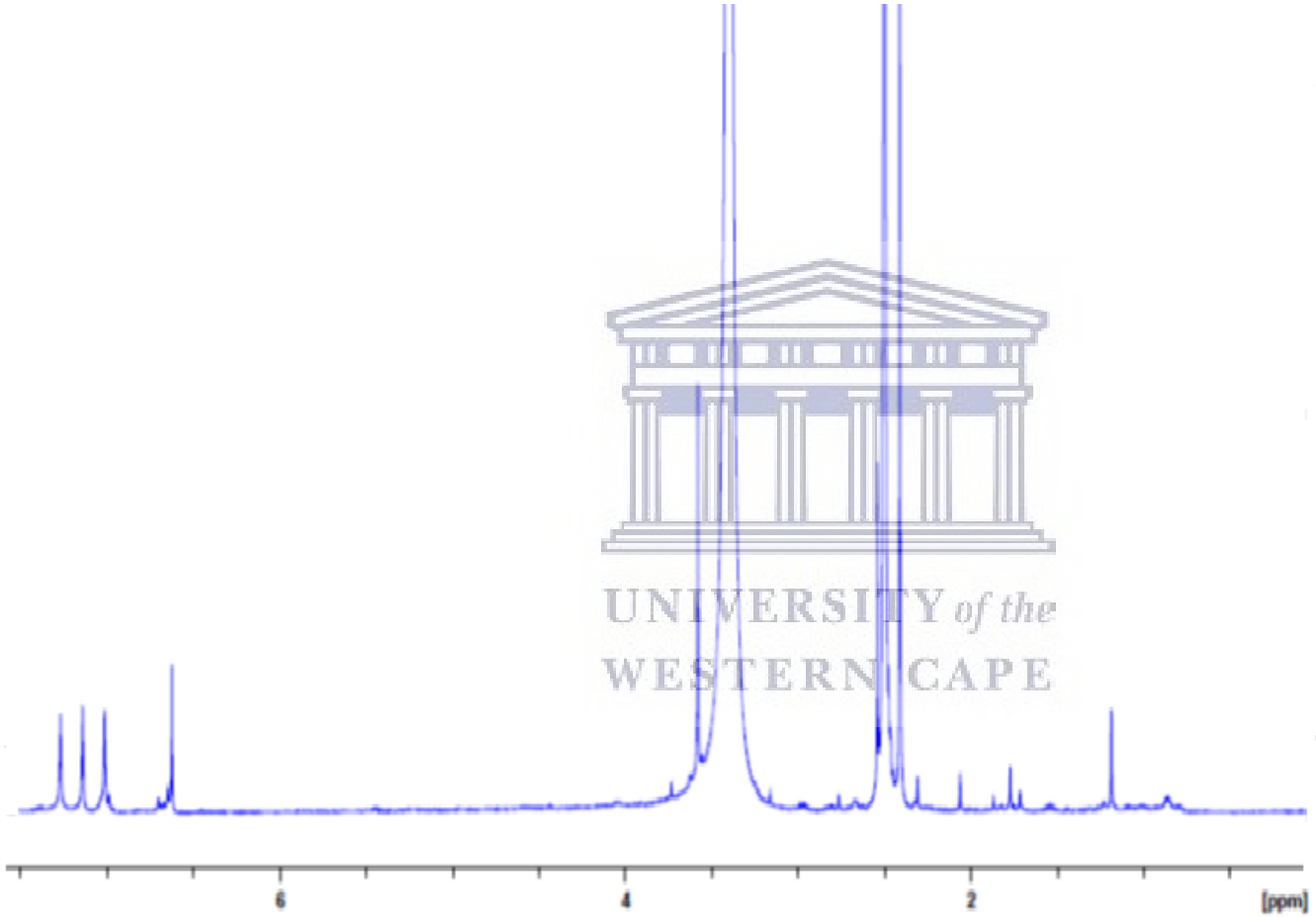


Figure 4. 13. ^{1}H NMR spectrum for compound 54 in DMSO-d_6 at 400 MHz.

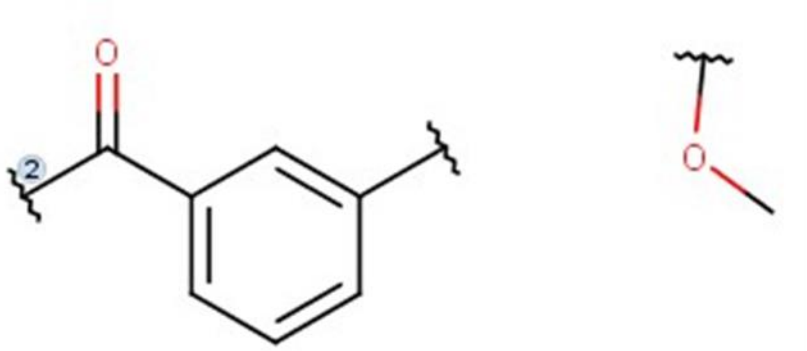


Figure 4. 14. Proposed substructures for Algoa BW2.

4.1.5.1.Hurdles encountered during purification

Regrettably, the DCM fraction contained only a low concentration of the bioactive compounds to be further characterised. The loss of activity presents a challenge for the purification of bioactive compounds from these strains, and alternative purification strategies will be required for future studies. It is worth noting that other fractions were not further purified, mostly due to loss of activity during the purification process.

5. Chapter 5: General discussion and conclusion

The discovery of antibiotics is one of the greatest scientific achievements of the twentieth century (Friedman et al., 2016). However, current global antibiotic resistance constantly spreading within hospitals, communities and environment threaten to undo all the progress made in the quality of health (Debbab et al., 2010). For example, drug-resistant pathogens are currently costing the health system huge amounts of money with high mortality rates and extended hospital stays from patients affected with previously easily treated diseases (Friedman et al., 2016). Additionally, according to the WHO, diseases such as malaria, diarrhoeal diseases, tuberculosis and HIV/AIDS are still claiming lives more significantly in low-income countries (WHO, 2020b).

Pathogenic microorganisms such as *Streptococcus pneumoniae*, *S. aureus*, *K. pneumoniae* and *Enterobacteriaceae* are proving to be the dominating pathogenic microorganisms in this dilemma (Friedman et al., 2016; Roca et al., 2015). In addition to drug resistance, emerging of new diseases (e.g., Ebola and Covid-19) and adverse side effects of some current medications has led to a lack of treatment options, which in turn has prompted scientists to reignite the drug and discovery research for new compounds.

Secondary metabolites are a good source of drug candidates. However, over-exploitation of the easily accessible sources of bioactive compounds such as plants and terrestrial microorganisms has led to the increasing rediscovery of known compounds. For these reasons, scientists aim to investigate so far underexplored sources such as marine environments and understudied microorganisms (Sabdono and Radjasa, 2011). The marine environment remains an under-explored niche as a source of bioactive SMs due to its size and the requirement for sophisticated sampling technologies (Li et al., 2015).

Marine invertebrates (sponges, soft-corals, molluscs, cnidarians, bryozoans and tunicates) are well-known sources of bioactive compounds with therapeutic properties (Sanchez et al., 2012).

Over 20000 marine-derived natural products have been discovered since 1950 and the majority were isolated from marine invertebrates (Leal et al., 2012). They yielded drugs or advanced drug candidates that are currently in the market or clinical trials such as trabectedin (Yondelis®), plitidepsin (Aplidin®), and 3-(2,4-dimethoxybenzylidene)-anabaseine (GTS-21) for the treatment of metastatic liposarcoma (Cordeiro et al., 2020), multiple myeloma (Leisch et al., 2019) and Alzheimer's (Rosas-Ballina et al., 2009), respectively.

However, marine derived SMs are not fully exploited due to the limited supply of the marine invertebrates material as huge amounts are required for sufficient early clinical development of the compounds (Steinert et al., 2014). For example, the amount of material available from the marine invertebrate *Lissodendoryx* sp. was insufficient for preclinical trials of halochondrins, anti-mitotic agents (Fiedler et al., 2005). Researchers have tried to build aquacultures to address the supply; however, the solution is not feasible as building and operating aquacultures is accompanied by extreme expenses and disease (Leal et al., 2014).

Fortunately, research has shown that in several instances, MIA bacteria are the true producers of many of the bioactive compounds isolated from marine invertebrates. In the literature, there are numerous examples of metabolites originally isolated from marines invertebrates, but which later have been shown to be either similar or even identical to compounds produced by MIA (Maslin et al., 2021; Piel, 2009, 2006, 2004). Microorganisms thus represent a sustainable supply of valuable biotechnological products compared to marine invertebrates as bacteria can be more easily manipulated through culturing and genetic techniques (Sabdono and Radjasa, 2011). Bacteria from marine environments have proven to be a great source of new bioactive compounds, which hold therapeutic value as potential drug candidates (Debbab et al., 2010). Marine bacteria are contributing to the drug development pipeline with a couple of drugs undergoing clinical trials such as soblidotin, tasidotin, and pulicatin A (Malve, 2016).

They are biosynthesised by complex enzymes such as NRPS and PKS. Based on their chemical properties and enzymes used to synthesise them, they are classified as PKs, NRPs, terpenes, alkaloids, peptides and lipopeptides. However, the majority of BGCs responsible for potential biotechnological / pharmaceutical important compounds are not expressed under laboratory conditions, posing a challenge in their production (Lindequist, 2016). OSMAC and heterologous expression are strategies that are used in triggering the expression of silent pathways. The OSMAC approach alters culturing parameters such as culturing vessel, state of the media, carbon sources, aeration rate and temperature to trigger the expression (Ruiz et al., 2010). Heterologous expression involves cloning of the cluster and expressing in a foreign host expecting the same product as in the native host (Galm and Shen, 2006; Julien and Shah, 2002).

In this study, 23 marine invertebrate-associated (MIA) bacteria isolated from the South African coast were exploited as a source of bioactive compounds and new chemical entities. Marine bacteria isolated from marine invertebrates were evaluated using the OSMAC approach and whole genome sequencing / genome mining. OSMAC parameters used during the study were media (four types) and media state (solid or liquid). Different anti-microbial profiles were observed under these different culturing conditions for the same strains indicating the expression of different BGCs using the OSMAC approach. The isolates were also evaluated for their ability to produce anti-inflammatory compounds.

The major limitation in studying the influence of metabolites produced by the bacteria was due to limited access to HRLC-MS and data processing facilities. This prevents multiple chemical analysis for the different metabolites being produced by the bacteria under different culturing conditions. However, based on the bioassay evaluation it is sufficient to conclude that different SMs were produced under different conditions as the anti-microbial profiles were different. This illustrated a wealth of SMs with potential to be antibiotics in the 23 MIA bacteria. The OSMAC approach has shown to be effective in activating expression of different BGCs, with growth and media optimisation being an area requiring focus for microorganisms to be suitable and sustainable producers of SMs.

Although the *Streptomyces* genus and other actinobacteria are known to include prolific producers of secondary metabolites (Subramani and Aalbersberg, 2012), they were not the focus of this study. The study focused on genera that are underexplored as a source of bioactive SMs to minimise the discovery of known and well characterised SMs. This was further motivated by genome mining studies which have illustrated those other phyla are also potential producers of SMS (Chapter 1, Section 1.6.5). The 16S rRNA sequence analysis revealed that the majority of the bacteria focused on in this study belong to the *Vibrio, Pseudovibrio* and *Bacillus* genera and they showed great potential as a source of bioactive SMs.

MIA bacteria showed bioactivity against two multidrug-resistant bacteria (*E. coli* 1699 and MRSA). Anti-microbial activity against multidrug-resistant bacterial strains is very important as it increases the chances of isolating new compounds with a new mode of action (Hentschel et al., 2001). The identification, purification and characterisation of the compounds responsible for bioactivity against *E. coli* 1699 and MRSA will assist in combating the negative effect caused by drug-resistant pathogenic microbes.

The anti-microbial result of the well-diffusion (chapter 3.1) and overlay assay represents the ability of the MIA isolate to produce extracellular bioactive SMs, and therefore, any intracellular bioactive SMs produced by the MIA isolates would have been excluded and may still represent a potential source of even more bioactivities and warrants further exploration. The intracellular compounds can be studied through techniques such as organic solvent extraction and ultrasound cell bursting before extraction and bioassay. Therefore, a study on both intracellular and extracellular compounds will represent a holistic assessment of the ability of the OSMAC approach used here to trigger expression of SMs within the isolates. Our findings are in correlation with the notion that MIA bacteria are a great source of compounds with potential pharmaceutical properties (Mohan et al., 2016).

The differences in anti-microbial profiles from isolates identified to be the same bacteria based on 16S rRNA analysis even though they were isolated from the same environment highlight the need

for molecular fingerprinting techniques, such as terminal restriction fragment length polymorphism (T-RFLP) in case of a targeted SM class or genome mining for the full genetic potential of the strains. Some studies have previously shown a shortfall of 16S rRNA analysis as SM discovery prioritisation tool. Comparative genome studies have shown that 16S rRNA sequence analysis is not enough to dereplicate as species with 100 % similar 16S rRNA sequences can harbour different SMs biosynthesis. A study on *Pseudovibrio* showed that isolates with the same phylogenetic relationship based on 16S rRNA harboured different SM (Esteves et al., 2017).

In this study, genome mining was chosen as a tool to further discover SM from MIA bacteria. The genome mining approach has gained popularity in drug discovery due to the drop in the cost of sequencing which resulted in an increase in available public genome data and increased availability of free informatic platforms (Aleti et al., 2015; Machado et al., 2015). The availability of sequences in the last 15 years or so was accompanied by tools that assisted in the understating of SM biosynthesis and mining which boosted the significance of the genome in drug discovery (Albarano et al., 2020). Genome mining is also important not only due to 16S rRNA being unreliable but because location and host have been shown to influence the type of SM BGCs in a bacterium. Therefore, genome mining will enable the discovery of new BGCs within the bacteria enabling exploration with techniques such as OSMAC, genetic manipulation and heterologous expression.

The genomes of the three MIA bacteria (*Bacillus* sp. PE08-149B, *Pseudovibrio* sp. PE14-07 and *Vibrio* sp. PE14-63;) with desired anti-microbial activities were sequenced and the genomes were analysed with antiSMASH resulting in the identification of both known and unknown BGCs. The unknown BGCs represent the potential for these bacteria to produce new bioactive compounds. The isolates contained BGCs which were unique to them; for example, PE14-07 contains a new beta lactone pathway.

Bacteriocins were the most abundant class of SMs from all strains (PE14-63, PE08-149B and PE14-07) (Figure 5.1). PE8-149B had the highest number of BGC followed by PE14-07, with

PE14-63 containing the lowest number. All genomes sequenced in this study were regarded as large genomes as they are above 5 Mbp in size. The abundance of SM BGCs within these bacteria is in agreement with the observation that large genomes contain more BGCs compared to smaller genomes (Cimermancic et al., 2014). The data gathered here clearly indicate the potential for production of new SMs as most of the predicted BGCs show low or no sequence similarity to characterised BGCs.

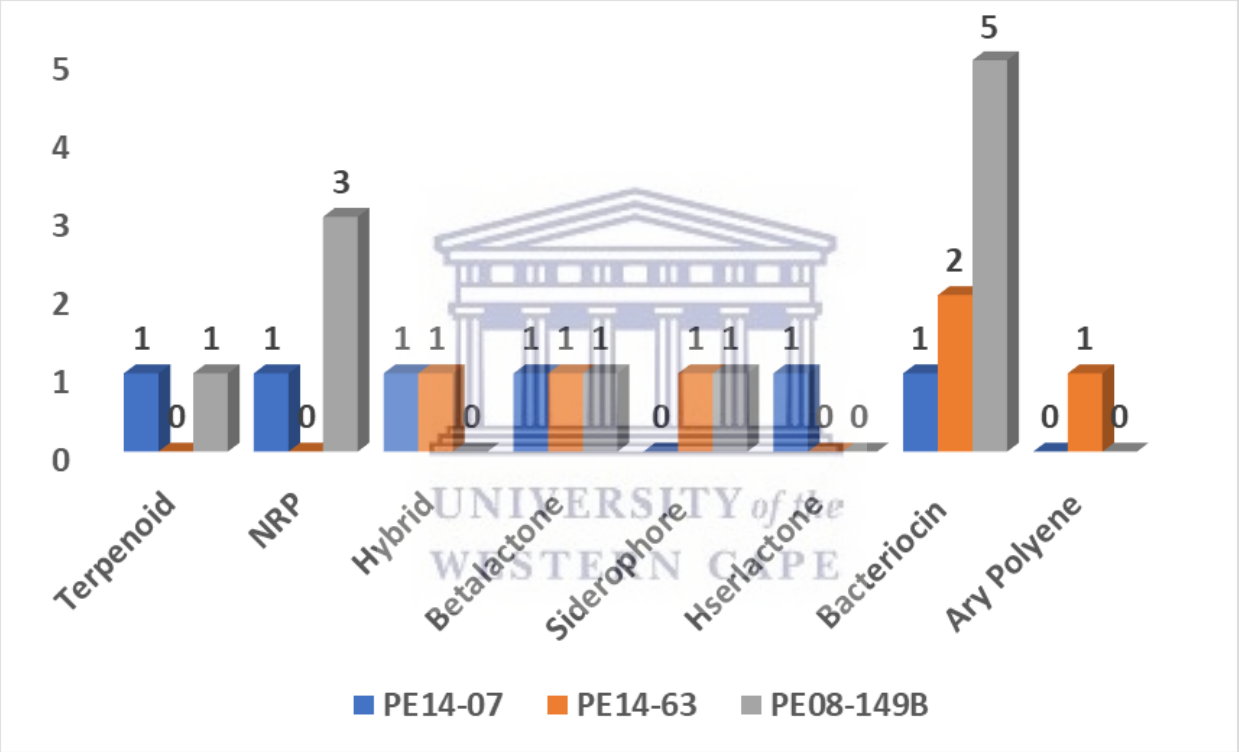


Figure 5. 1. Graphic representation of antiSMASH identified BGCs from PE14-07, PE14-63 and PE08-149B genomes.

The bioinformatics analysis results are in correlation with the suggestion in Section 3.1 and 4.1 (anti-microbial screening), in that multiple pathways are responsible for the diverse activity observed from these microorganisms. The identification of BGCs encoding for a NRP, bacteriocin and PK within a genome of one microorganism increases the chances of being a potential anti-microbial compound producer, as these classes of compounds are well known for anti-microbial properties (Ayuso-Sacido and Genilloud, 2005; Newman and Cragg, 2016; van Heel et al., 2013).

Genome mining does not only offer the ability to assess the genetic potential of an organism to produce SMs, but it also functions as a dereplication tool. This happens by comparing identified clusters to well characterised clusters. Dereplication utilising genome mining is a very important component of SM discovery. Genome mining involves the use of bioinformatics databases to analyse genome sequence of organisms and compares it to information in databases such as MIBiG. Dereplication was traditionally performed using chemistry tools such as LC-MS-MS, NMR, and UV-VIS coupled with database searches. Due to the limited sensitivity of some of the techniques, cost of the instruments, lack of expertise and no expression of the majority of the BGCs under laboratory conditions, their full usage is limited (Lin et al., 2008; Macintyre et al., 2014). Additionally, chemical dereplication is accompanied by lifetime cost as the majority of reliable databases require yearly licence renewals.

Alternatively, genome mining offers advantages such as the ability to interrogate the full SM biosynthesis potential of an organism, independent of the production of the SM in question. This enables the identification of silent gene clusters under laboratory conditions, a bottleneck that can then be addressed through OSMAC and genetic engineering approaches. Additionally, the genome mining cost is once-off and the sequence data can be reanalysed indefinitely on free platforms e.g., PRISM and antiSMASH, especially as these platforms are continuously updated (for example, the known cluster databases) and improved for more informative analyses to be conducted. These result in minimising chemical or bioassay analysis required. However, biological and chemical assays are still required to evaluate the biotechnological value of the SM experimentally as genome mining cannot indicate if a new compound will be bioactive with certainty. Hence, in this study, chemical, bioassay and genome-based approaches were used to complement each strategies' limitations. Enabling information-rich bacterial isolate prioritisation for SMs discovery instead of relying on one approach could lead to overlooking valuable strains.

The identification of all the BGCs (PE14-63, PE14-07 and PE08-149B) with low gene or sequence similarity to known and characterised BGCs from the MIBiG database is an indication that the isolates are yet to be fully exploited. The low genes / sequences / no similarity to BGCs in the

MIBiG database signifies potential new SMs from all three strains. The discovery of these BGCs enables the characterising of the pathways with the use of techniques such as cloning, OSMAC, genetic manipulation (promoter replacement) and expression studies (reverse transcription-polymerase chain reaction (RT-PCR) / transcriptomics) in efforts to study, isolate and identify biotechnological uses of compounds from the BGCs, opening doors for future studies.

However, one must bear in mind the risk of selecting a new BGC from genome mining results as current tools are unable to detect the potential biotechnological value of the product. The above implies that one might invest in BGCs with no anti-microbial activity in relation to drug discovery. Another major drawback is that only known classes will be identified leaving out any SMs that are synthesised by single or multiple enzymes that are distributed across the genome and not arranged as a single operon. This area still needs development in the bioinformatics field which can only happen as we understand more of the BGC biosynthesis, highlighting the need for not only applied research within the field of SMs discovery but also basic SM research.

The prediction of the structures with PRISM could assist in the identification and isolation of the SM as the predicted structure can be used to generate molecular weight, monoisotopic mass and chemical formula. The bioinformatical generated data could be used to mine HRLC-MS metabolic data obtain from analysis organic extract of the isolate to determine if a BGC of interest is expressed under the studied culturing condition. Additionally, monomer prediction (PRISM and antiSMASH) can be used to track and monitor metabolites during purification through an experiment such as isotope labelling.

NRP BGCs identified from PE14-07 and PE08-149B were characterised with bioinformatics tools (NaPDos and NRPSpredictor). Majority of the pathway domains (A- and C- domains) showed some low similarity to BGCs from *Streptomyces* and *Bacillus*, indicating a possibility of genetic exchange or this was because the majority of the characterised BGC are from *Streptomyces* and

Bacillus. Another reason could be mutations in the strain's genome. The *Pseudovibrio* genus is well characterised with regard to its ability to perform gene transfer as its genome has multiple mobile elements (Alex and Antunes, 2015; Versluis et al., 2018).

The ability of BGCs to be transferred or modified through the processes mentioned is part of the reason why the ability of isolates to produce new bioactivities should not be judged solely on their 16S rRNA phylogeny. In other words, “common” bacteria such as those belonging to the Proteobacteria and Firmicutes, which represent the phyla with the most cultured representatives, still warrant investigation as these may contain new gene clusters, especially if isolated from never-before sampled environments as seen in this study.

Future studies for the NRPs and other identified clusters include, but are not limited to, host optimisation (insertion of on and off plug sequences, promoters), heterologous expression, isolation of the compounds, chemical characterisation, minimum inhibition concentration assays and structural elucidation of the product. The above process can be useful in accelerating SM discovery and elimination of bottlenecks such as cryptic clusters and low product titre.

Chemically guided prioritisation of the extract using NMR resulted in the isolation of a new compound from *Pseudovibrio* PE14-07. The novelty of the compound was determined by databases and literature searches of compounds isolated from the *Pseudovibrio* genus and other SMs. Unfortunately, the compound showed no bioactivity against any of the test strains. A variety of assays, to assess, for example, anti-viral, anti-cancer and anti-diabetic activities are needed to evaluate the full potential of any isolated SM. However, the discovery of compounds without anti-microbial activity is still very significant as the compounds can have application in other industries such as agriculture and cosmetics, not to mention, to gain a better understanding of the SM’s ecological role. Table 5.1 provides a summary of the identified compounds from PE14-07 and PE08-149B.

Table 5. 1. Summary of SMs and bioactivities from PE14-07, PE14-63 and PE08-149B.

Organism	Marine invertebrate (source)	Bioactivities	No. of BGCs (antiSMASH)	Identified / isolated compounds
Pseudovibrio ascidiaceicola PE14-07	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886	*E. coli* 1699, *B. cereus*, *Pseudomonas putida*, *S. epidermidis*, *A. fumigatus*, *C. albicans* and Anti-inflammatory	7	Algoa PV1 – new Algo PV2 – new Algoa PV3 – Known
Bacillus mycoides PE08-149B	*Spongia (Spongia)* sp.001RSASPN	*Escherichia coli* 1699, *B. cereus*, *P. putida*, *Mycobacterium aurum* and *S. epidermidis*	11	Algoa BW1 – new Algoa BW2 – new
Vibrio tasmaniensis PE14-63	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886	*E. coli* 1699, *B. cereus*, *S. epidermidis*, *A. fumigatus* and *C. albicans*	6	None

The current coronavirus outbreak together with previous pandemics such as Ebola and swine flu shows the need to have numerous compounds in chemical libraries and databases available to be ready to screen. Halitunal, a marine diterpene aldehyde has shown to be active against a murine coronavirus (Koehn et al., 1991). This study highlights the need for more studies to screen the potential of marine terpene and other classes as a source of anti-viral (e.g., anti-SARS-CoV-2).

MIA bacteria from the South African coastal line are a potentially good source of diverse anti-microbial compounds as displayed by genome mining and bioassays. Also, the focus on MIA represents an environmentally sustainable solution as once the bacteria are isolated, they can be preserved in the laboratory for further manipulation, eliminating continuous environmental sampling.

6. References

Achan, J., Talisuna, A.O., Erhart, A., Yeka, A., Tibenderana, J.K., Baliraine, F.N., Rosenthal, P.J., D'Alessandro, U., 2011. Quinine, An Old Anti-Malarial Drug In A Modern World: Role In The Treatment Of Malaria. Malar. J. 10, 1475–2875.

Adamek, M., Spohn, M., Stegmann, E., Ziemert, N., 2017. Mining Bacterial Genomes For Secondary Metabolite Gene Clusters, In: Antibiotics, Methods In Molecular Biology. Humana Press, New York, NY, 23–47. https://doi.org/10.1007/978-1-4939-6634-9_2

Adams, D.J., Berecki, G., 2013. Mechanisms Of Conotoxin Inhibition Of N-Type (Cav2.2) Calcium Channels. Biochim. Biophys. Acta BBA - Biomembr., Calcium Channels 1828, 1619–1628. https://doi.org/10.1016/j.bbamem.2013.01.019

Agrawal, S., Acharya, D., Adholeya, A., Barrow, C.J., Deshmukh, S.K., 2017. Nonribosomal Peptides From Marine Microbes And Their Anti-microbial And Anti-cancer Potential. Front. Pharmacol. 8. https://doi.org/10.3389/fphar.2017.00828

Agrawal, S., Adholeya, A., Deshmukh, S.K., 2016. The Pharmacological Potential Of Non-Ribosomal Peptides From Marine Sponge And Tunicates. Front. Pharmacol. 7. https://doi.org/10.3389/fphar.2016.00333

Ahmed, Y., Rebets, Y., Tokovenko, B., Brötz, E., Luzhetskyy, A., 2017. Identification Of Butenolide Regulatory System Controlling Secondary Metabolism In *Streptomyces albus* J1074. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-10316-y

Albarano, L., Esposito, R., Ruocco, N., Costantini, M., 2020. Genome Mining As New Challenge In Natural Products Discovery. Mar. Drugs 18, 199. https://doi.org/10.3390/md18040199

Aleti, G., Sessitsch, A., Brader, G., 2015. Genome Mining: Prediction Of Lipopeptides And Polyketides From *Bacillus* And Related Firmicutes. Comput. Struct. Biotechnol. J. 13, 192–203. https://doi.org/10.1016/j.csbj.2015.03.003

Alex, A., Antunes, A., 2015. Whole Genome Sequencing Of The Symbiont *Pseudovibrio* sp. From The Intertidal Marine Sponge *Polymastia penicillus* Revealed A Gene Repertoire For Host-Switching Permissive Lifestyle. Genome Biol. Evol. 7, 3022–3032. https://doi.org/10.1093/gbe/evv199

Alex, A., Silva, V., Vasconcelos, V., Antunes, A., 2013. Evidence Of Unique And Generalist Microbes In Distantly Related Sympatric Intertidal Marine Sponges (Porifera: Demospongiae). PLoS ONE 8, e80653. https://doi.org/10.1371/journal.pone.0080653

Alexandratos, S.D., 2009. Ion-Exchange Resins: A Retrospective From Industrial And Engineering Chemistry Research. Ind. Eng. Chem. Res. 48, 388–398. https://doi.org/10.1021/ie801242v

Amin, N.M., Bunawan, H., Redzuan, R.A., Jaganath, I.B.S., 2010. *Erwinia mallotivora* sp., A New Pathogen Of Papaya (*Carica Papaya*) In Peninsular Malaysia. Int. J. Mol. Sci. 12, 39–45. https://doi.org/10.3390/ijms12010039

Amin, S.A., Green, D.H., Küpper, F.C., Carrano, C.J., 2009. Vibrioferrin, An Unusual Marine Siderophore: Iron Binding, Photochemistry, And Biological Implications. Inorg. Chem. 48, 11451–11458. https://doi.org/10.1021/ic9016883

Anand, T.P., Bhat, A.W., Shouche, Y.S., Roy, U., Siddharth, J., Sarma, S.P., 2006. Anti-microbial Activity Of Marine Bacteria Associated With Sponges From The Waters Off The Coast Of South East India. Microbiol. Res. 161, 252–262. https://doi.org/10.1016/j.micres.2005.09.002

Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., Bulaj, G., Camarero, J.A., Campopiano, D.J., Challis, G.L., Clardy, J., Cotter, P.D., Craik, D.J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P.C., Entian, K.-D., Fischbach, M.A., Garavelli, J.S., Göransson, U., Gruber, C.W., Haft, D.H., Hemscheidt, T.K., Hertweck, C., Hill, C., Horswill, A.R., Jaspars, M., Kelly, W.L., Klinman, J.P., Kuipers, O.P., Link, A.J., Liu, W., Marahiel, M.A., Mitchell, D.A., Moll, G.N., Moore, B.S., Müller, R., Nair, S.K., Nes, I.F., Norris, G.E., Olivera, B.M., Onaka, H., Patchett, M.L., Piel, J., Reaney, M.J.T., Rebuffat, S., Ross, R.P., Sahl, H.-G., Schmidt, E.W., Selsted, M.E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Süssmuth, R.D., Tagg, J.R., Tang, G.-L., Truman, A.W., Vederas, J.C., Walsh, C.T., Walton, J.D., Wenzel, S.C., Willey, J.M., van der Donk, W.A., 2013. Ribosomally Synthesized And Post-Translationally Modified Peptide Natural Products: Overview And Recommendations For A Universal Nomenclature. Nat. Prod. Rep. 30, 108–160. https://doi.org/10.1039/C2NP20085F

Årstøl, E., Hohmann-Marriott, M.F., 2019. Cyanobacterial Siderophores—Physiology, Structure, Biosynthesis, And Applications. Mar. Drugs 17, 281. https://doi.org/10.3390/md17050281

Ascidiacea World Database, n.d. Ascidiacea World Database - *Pseudodistoma africanum* Millar, 1954 [WWW Document]. URL http://www.marinespecies.org/ascidiacea/aphia.php?p=taxdetails&id=215960 (accessed 2.24.20).

Atta-ur-Rahman, 2017. Frontiers In Natural Product Chemistry. Bentham Science Publishers.

Ayuso, A., Clark, D., González, I., Salazar, O., Anderson, A., Genilloud, O., 2004. A Novel Actinomycete Strain De-Replication Approach Based On The Diversity Of Polyketide Synthase And Nonribosomal Peptide Synthetase Biosynthetic Pathways. Appl. Microbiol. Biotechnol. 67, 795–806.

Ayuso-Sacido, A., Genilloud, O., 2005. New PCR Primers For The Screening Of NRPS And PKS-I Systems In Actinomycetes: Detection And Distribution Of These Biosynthetic Gene Sequences In Major Taxonomic Groups. Microb. Ecol. 49, 10–24. https://doi.org/10.1007/s00248-004-0249-6

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., Miao, J., Lou, C., Zhang, L., 2015. Exploiting A Precise Design Of Universal Synthetic Modular Regulatory Elements To Unlock The Microbial Natural Products In *Streptomyces*. Proc. Natl. Acad. Sci. 112, 12181–12186.

Baker, B.J., 2015. Marine Biomedicine: From Beach To Bedside. CRC Press.

Barnard, A.M.L., Bowden, S.D., Burr, T., Coulthurst, S.J., Monson, R.E., Salmond, G.P.C., 2007. Quorum Sensing, Virulence And Secondary Metabolite Production In Plant Soft-Rotting Bacteria. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 362, 1165–1183. https://doi.org/10.1098/rstb.2007.2042

Bärnighausen, T., Tanser, F., Gqwede, Z., Mbizana, C., Herbst, K., Newell, M.-L., 2008. High HIV Incidence In A Community With High HIV Prevalence In Rural South Africa: Findings From A Prospective Population-Based Study. AIDS 22, 139. https://doi.org/10.1097/QAD.0b013e3282f2ef43

Batdorj, B., Dalgalarrondo, M., Choiset, Y., Pedroche, J., Métro, F., Prévost, H., Chobert, J.-M., Haertlé, T., 2006. Purification And Characterization Of Two Bacteriocins Produced By Lactic Acid Bacteria Isolated From Mongolian Airag. J. Appl. Microbiol. 101, 837–848. https://doi.org/10.1111/j.1365-2672.2006.02966.x

Benkő, R., Gajdács, M., Matuz, M., Bodó, G., Lázár, A., Hajdú, E., Papfalvi, E., Hannauer, P., Erdélyi, P., Pető, Z., 2020. Prevalence And Antibiotic Resistance Of ESKAPE Pathogens Isolated In The Emergency Department Of A Tertiary Care Teaching Hospital In Hungary:

A 5-Year Retrospective Survey. Antibiot. Basel Switz. 9, E624. https://doi.org/10.3390/antibiotics9090624

Berger, M., Neumann, A., Schulz, S., Simon, M., Brinkhoff, T., 2011. Tropodithietic Acid Production In *Phaeobacter gallaeciensis* Is Regulated By N-Acyl Homoserine Lactone-Mediated Quorum Sensing. J. Bacteriol. 193, 6576–6585. https://doi.org/10.1128/JB.05818-11

Bergmann, S., Schümann, J., Scherlach, K., Lange, C., Brakhage, A.A., Hertweck, C., 2007. Genomics-Driven Discovery Of PKS-NRPS Hybrid Metabolites From *Aspergillus nidulans*. Nat. Chem. Biol. 3, 213–217. https://doi.org/10.1038/nchembio869

Bertrand, S., Bohni, N., Schnee, S., Schumpp, O., Gindro, K., Wolfender, J.-L., 2014. Metabolite Induction Via Microorganism Co-Culture: A Potential Way To Enhance Chemical Diversity For Drug Discovery. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2014.03.001

Bhate, D.S., 1955. Pumilin, A New Antibiotic From *Bacillus pumilus*. Nature 175, 816–817. https://doi.org/10.1038/175816a0

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H., 2011. Characterisation Of The *Escherichia coli* Strain Associated With An Outbreak Of Haemolytic Uraemic Syndrome In Germany, 2011: a Microbiological Study. Lancet Infect. Dis. 11, 671–676. https://doi.org/10.1016/S1473-3099(11)70165-7

Bigelis, R., He, H., Yang, H.Y., Chang, L.-P., Greenstein, M., 2006. Production Of Fungal Antibiotics Using Polymeric Solid Supports In Solid-State And Liquid Fermentation. J. Ind. Microbiol. Biotechnol. 33, 815–826. https://doi.org/10.1007/s10295-006-0126-z

Blin, K., Pascal Andreu, V., de los Santos, E.L.C., Del Carratore, F., Lee, S.Y., Medema, M.H., Weber, T., 2019. The Antismash Database Version 2: A Comprehensive Resource On Secondary Metabolite Biosynthetic Gene Clusters. Nucleic Acids Res. 47, D625–D630. https://doi.org/10.1093/nar/gky1060

Blunt, J.W., Copp, B.R., Keyzers, R.A., Munro, M.H.G., Prinsep, M.R., 2014. Marine Natural Products. Nat. Prod. Rep. 31, 160. https://doi.org/10.1039/c3np70117d

Bode, E., Brachmann, A.O., Kegler, C., Simsek, R., Dauth, C., Zhou, Q., Kaiser, M., Klemmt, P., Bode, H.B., 2015. Simple "On-Demand" Production Of Bioactive Natural Products. ChemBioChem 16, 1115–1119. https://doi.org/10.1002/cbic.201500094

Bode, H.B., Bethe, B., Höfs, R., Zeeck, A., 2002. Big Effects From Small Changes: Possible Ways To Explore Nature's Chemical Diversity. ChemBioChem 3, 619–627. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9

Boiteau, R.M., Mende, D.R., Hawco, N.J., McIlvin, M.R., Fitzsimmons, J.N., Saito, M.A., Sedwick, P.N., DeLong, E.F., Repeta, D.J., 2016. Siderophore-Based Microbial Adaptations To Iron Scarcity Across The Eastern Pacific Ocean. Proc. Natl. Acad. Sci. 113, 14237–14242. https://doi.org/10.1073/pnas.1608594113

Bondarev, V., Richter, M., Romano, S., Piel, J., Schwedt, A., Schulz-Vogt, H.N., 2013. The Genus *Pseudovibrio* Contains Metabolically Versatile Bacteria Adapted For Symbiosis: The Genus *Pseudovibrio*. Environ. Microbiol. 15, 2095–2113. https://doi.org/10.1111/1462-2920.12123

Boottanun, P., Potisap, C., Hurdle, J.G., Sermswan, R.W., 2017. Secondary Metabolites From *Bacillus amyloliquefaciens* Isolated From Soil Can Kill *Burkholderia pseudomallei*. AMB Express 7. https://doi.org/10.1186/s13568-016-0302-0

Brady, S.F., Chao, C.J., Handelsman, J., Clardy, J., 2001. Cloning And Heterologous Expression Of A Natural Product Biosynthetic Gene Cluster From Edna. Org. Lett. 3, 1981–1984.

Brady, S.F., Clardy, J., 2004. Palmitoylputrescine, An Antibiotic Isolated From The Heterologous Expression Of DNA Extracted From Bromeliad Tank Water. J. Nat. Prod. 67, 1283–1286. https://doi.org/10.1021/np0499766

Bromley, C.L., Shirley, P.-N., de la Mare, J.-A., Beukes, D., Edkins, Davies-Coleman, M.T., 2013. Halogenated Oxindole And Indoles From The South African Aarine Ascidian *Distaplia skoogi* [WWW Document]. ResearchGate. URL https://www.researchgate.net/publication/279919824_Halogenated_oxindole_and_indoles_from_the_South_African_marine_ascidian_Distaplia_skoogi (accessed 6.23.18).

Bruhn, J.B., Nielsen, K.F., Hjelm, M., Hansen, M., Bresciani, J., Schulz, S., Gram, L., 2005. Ecology, Inhibitory Activity, And Morphogenesis Of A Marine Antagonistic Bacterium Belonging To The *Roseobacter* Clade. Appl. Environ. Microbiol. 71, 7263–7270. https://doi.org/10.1128/AEM.71.11.7263-7270.2005

Bull, A.T., Stach, J.E.M., 2007. Marine Actinobacteria: New Opportunities For Natural Product Search And Discovery. Trends Microbiol. 15, 491–499. https://doi.org/10.1016/j.tim.2007.10.004

Caboche, S., Pupin, M., Leclere, V., Fontaine, A., Jacques, P., Kucherov, G., 2007. NORINE: A Database Of Nonribosomal Peptides. Nucleic Acids Res. 36, D326–D331. https://doi.org/10.1093/nar/gkm792

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: Architecture And Applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cane, D.E., Ikeda, H., 2012. Exploration And Mining Of The Bacterial Terpenome. Acc. Chem. Res. 45, 463–472. https://doi.org/10.1021/ar200198d

Carr, G., Poulsen, M., Klassen, J.L., Hou, Y., Wyche, T.P., Bugni, T.S., Currie, C.R., Clardy, J., 2012. Microtermolides A And B From Termite-Associated *Streptomyces* Sp. And Structural Revision Of Vinylamycin. Org. Lett. 14, 2822–2825. https://doi.org/10.1021/ol301043p

Carroll, A.R., Copp, B.R., Davis, R.A., Keyzers, R.A., Prinsep, M.R., 2019. Marine Natural Products. Nat. Prod. Rep. 36, 122–173. https://doi.org/10.1039/C8NP00092A

Casey, H., Smith, A., Parker, L., Dipper, M., Gould, T., 2015. Pulmonary Tuberculosis In A South African Regional Emergency Centre: Can Infection Control Be Improved To Lower The Risk Of Nosocomial Transmission? SAMJ South Afr. Med. J. 105, 862–865. https://doi.org/10.7196/SAMJNEW.8329

Chairman, K., Singh, A.J.A.R., Ramesh, M., 2012. Screening Twelve Species Of Sponges For Biomedical Activity In Gulf Of Mannar Tuticorin Coast. Int. J. Mar. Sci. 2.

Challis, G.L., 2008. Mining Microbial Genomes For New Natural Products And Biosynthetic Pathways. Microbiology 154, 1555–1569. https://doi.org/10.1099/mic.0.2008/018523-0

Challis, G.L., Bailey, L.M., Deeth, R.J., Lautru, S., 2005. Discovery Of A New Peptide Natural Product By *Streptomyces coelicolor* Genome Mining. Nat. Chem. Biol. 1, 265. https://doi.org/10.1038/nchembio731

Challis, G.L., Naismith, J.H., 2004. Structural Aspects Of Non-Ribosomal Peptide Biosynthesis. Curr. Opin. Struct. Biol. 14, 748–756. https://doi.org/10.1016/j.sbi.2004.10.005

Chan, Y.A., Podevels, A.M., Kevany, B.M., Thomas, M.G., 2009. Biosynthesis Of Polyketide Synthase Extender Units. Nat. Prod. Rep. 26, 90–114.

Chancey, S.T., Zähner, D., Stephens, D.S., 2012. Acquired Inducible Anti-microbial Resistance In Gram-Positive Bacteria. Future Microbiol. 7, 959–978. https://doi.org/10.2217/fmb.12.63

Chen, J., Guo, Y., Lu, Y., Wang, B., Sun, J., Zhang, H., Wang, H., 2019. Chemistry And Biology Of Siderophores From Marine Microbes. Mar. Drugs 17, 562. https://doi.org/10.3390/md17100562

Chen, W.-C., Juang, R.-S., Wei, Y.-H., 2015. Applications Of A Lipopeptide Biosurfactant, Surfactin, Produced By Microorganisms. Biochem. Eng. J. 103, 158–169. https://doi.org/10.1016/j.bej.2015.07.009

Chenard, C., Lauro, F.M. (Eds.), 2017. Microbial Ecology Of Extreme Environments. Springer International Publishing.

Choi, S.-K., Park, S.-Y., Kim, R., Kim, S.-B., Lee, C.-H., Kim, J.F., Park, S.-H., 2009. Identification Of A Polymyxin Synthetase Gene Cluster Of *Paenibacillus polymyxa* and Heterologous Expression Of The Gene In *Bacillus subtilis*. J. Bacteriol. 191, 3350–3358. https://doi.org/10.1128/JB.01728-08

Choudhary, A., Naughton, L.M., Montánchez, I., Dobson, A.D.W., Rai, D.K., 2017. Current Status And Future Prospects Of Marine Natural Products (Mnps) As Anti-microbials. Mar. Drugs 15. https://doi.org/10.3390/md15090272

Chung, G.A., Aktar, Z., Jackson, S., Duncan, K., 1995. High-Throughput Screen For Detecting Antimycobacterial Agents. Antimicrob. Agents Chemother. 39, 2235–2238. https://doi.org/10.1128/AAC.39.10.2235

Churchill, M.E.A., Sibhatu, H.M., Uhlson, C.L., 2011. Defining The Structure And Function Of Acyl-Homoserine Lactone Autoinducers. Methods Mol. Biol. Clifton NJ 692, 159–171. https://doi.org/10.1007/978-1-60761-971-0_12

Ciani, O., Jommi, C., 2014. The Role Of Health Technology Assessment Bodies In Shaping Drug Development. Drug Des. Devel. Ther. 8, 2273. https://doi.org/10.2147/DDDT.S49935

Cimermancic, P., Medema, M.H., Claesen, J., Kurita, K., Wieland Brown, L.C., Mavrommatis, K., Pati, A., Godfrey, P.A., Koehrsen, M., Clardy, J., Birren, B.W., Takano, E., Sali, A., Linington, R.G., Fischbach, M.A., 2014. Insights Into Secondary Metabolism From A Global Analysis Of Prokaryotic Biosynthetic Gene Clusters. Cell 158, 412–421. https://doi.org/10.1016/j.cell.2014.06.034

Cobb, R.E., Zhao, H., 2012. Direct Cloning Of Large Genomic Sequences. Nat. Biotechnol. 30, 405–406.

Collins, F.W.J., O'Connor, P.M., O'Sullivan, O., Gómez-Sala, B., Rea, M.C., Hill, C., Ross, R.P., 2017. Bacteriocin Gene-Trait Matching Across The Complete *Lactobacillus* Pan-Genome. Sci. Rep. 7, 3481. https://doi.org/10.1038/s41598-017-03339-y

Cordeiro, M., Casanova, J.M., Rodrigues, J., Freitas, J., Fonseca, R., Oliveira, R.C. de, Tavares, P.F., 2020. Long-Term Response After 94 Cycles Of Trabectedin In A Patient With Metastatic Leiomyosarcoma Of The Lower Extremity. Case Rep. Oncol. 13, 113–119. https://doi.org/10.1159/000505393

Crowley, S., O'Gara, F., O'Sullivan, O., Cotter, P., Dobson, A., 2014. Marine *Pseudovibrio* sp. As A Novel Source Of Anti-microbials. Mar. Drugs 12, 5916–5929. https://doi.org/10.3390/md12125916

Crüsemann, M., O'Neill, E.C., Larson, C.B., Melnik, A.V., Floros, D.J., da Silva, R.R., Jensen, P.R., Dorrestein, P.C., Moore, B.S., 2016. Prioritizing Natural Product Diversity In A Collection Of 146 Bacterial Strains Based On Growth And Extraction Protocols. J. Nat. Prod. https://doi.org/10.1021/acs.jnatprod.6b00722

Cruz-Morales, P., Kopp, J.F., Martinez-Guerrero, C.E., Yanez-Guerra, L.A., Selem-Mojica, N., Ramos-Aboites, H.E., Feldmann, J., Barona-Gomez, F., 2016. Phylogenomic Analysis Of Natural Products Biosynthetic Gene Clusters Allows Discovery Of Arseno-Organic Metabolites In Model *streptomycetes*. bioRxiv 020503. https://doi.org/10.1101/020503

Cuadrat, R., Ionescu, D., Davila, A.M.R., Grossart, H.-P., 2017. Recovering Genomic Clusters Of Secondary Metabolites From Lakes: A Metagenomics 2.0 Approach. bioRxiv 183061. https://doi.org/10.1101/183061

Cueto, M., Jensen, P.R., Kauffman, C., Fenical, W., Lobkovsky, E., Clardy, J., 2001. Pestalone, A New Antibiotic Produced By A Marine Fungus In Response To Bacterial Challenge. J. Nat. Prod. 64, 1444–1446. https://doi.org/10.1021/np0102713

Cutignano, A., Nuzzo, G., Ianora, A., Luongo, E., Romano, G., Gallo, C., Sansone, C., Aprea, S., Mancini, F., D'Oro, U., Fontana, A., 2015. Development And Application Of A Novel SPE-Method For Bioassay-Guided Fractionation Of Marine Extracts. Mar. Drugs 13, 5736–5749. https://doi.org/10.3390/md13095736

Cwala, Z., Igbinosa, E.O., Okoh, A.I., 2011. Assessment Of Antibiotics Production Potentials In Four Actinomycetes Isolated From Aquatic Environments Of The Eastern Cape Province Of South Africa. Afr. J. Pharm. Pharmacol. 5, 118–124.

Darabpour, E., Roayaei Ardakani, M., Motamedi, H., Ronagh, M.T., Najafzadeh, H., 2012. Purification And Optimization Of Production Conditions Of A Marine-Derived Antibiotic And Ultra-Structural Study On The Effect Of This Antibiotic Against MRSA. Eur. Rev. Med. Pharmacol. Sci. 16, 157–165.

Dashti, Y., Grkovic, T., Abdelmohsen, U.R., Hentschel, U., Quinn, R.J., 2014. Production Of Induced Secondary Metabolites By A Co-Culture Of Sponge-Associated Actinomycetes, *Actinokineospora* sp. EG49 and *Nocardiopsis* sp. RV163. Mar. Drugs 12, 3046–3059. https://doi.org/10.3390/md12053046

Davies-Coleman, M.T., 2005. Bioactive Natural Products From Southern African Marine Invertebrates, In: Atta-Ur-Rahman (Ed.), Studies In Natural Products Chemistry, Bioactive Natural Products (Part L). Elsevier, pp. 61–107. https://doi.org/10.1016/S1572-5995(05)80054-7

Davies-Coleman, M.T., Beukes, D.R., 2005. Ten Years Of Marine Natural Products Research At Rhodes University. ChemInform 36. https://doi.org/10.1002/chin.200537245

Davies-Coleman, M.T., Veale, C.G.L., 2015. Recent Advances In Drug Discovery From South African Marine Invertebrates. Mar. Drugs 13, 6366–6383. https://doi.org/10.3390/md13106366

de Carvalho, C.C.C.R., Fernandes, P., 2010. Production Of Metabolites As Bacterial Responses To The Marine Environment. Mar. Drugs 8, 705–727. https://doi.org/10.3390/md8030705

de Jong, A., van Hijum, S.A.F.T., Bijlsma, J.J.E., Kok, J., Kuipers, O.P., 2006. BAGEL: A Web-Based Bacteriocin Genome Mining Tool. Nucleic Acids Res. 34, W273–W279. https://doi.org/10.1093/nar/gkl237

de Lucas, A., Cañizares, P., García, M.A., Gómez, J., Rodríguez, J.F., 1998. Recovery Of Nicotine From Aqueous Extracts Of Tobacco Wastes By An H+-Form Strong-Acid Ion Exchanger. Ind. Eng. Chem. Res. 37, 4783–4791. https://doi.org/10.1021/ie9801958

Debbab, A., Aly, A.H., Lin, W.H., Proksch, P., 2010. Bioactive Compounds From Marine Bacteria And Fungi: Marine Bioactive Compounds. Microb. Biotechnol. 3, 544–563. https://doi.org/10.1111/j.1751-7915.2010.00179.x

Deleu, M., Paquot, M., Nylander, T., 2008. Effect Of Fengycin, A Lipopeptide Produced By Bacillus Subtilis, On Model Biomembranes. Biophys. J. 94, 2667–2679. https://doi.org/10.1529/biophysj.107.114090

DeLong, E.F., Yayanos, A.A., 1986. Biochemical Function And Ecological Significance Of Novel Bacterial Lipids In Deep-Sea Procaryotes. Appl. Environ. Microbiol. 51, 730–737.

Desauziers, V., Boucharat, C., Cloirec, P.L., 1998. Ion-Exchange Extraction Of Organic Acids In Water Samples. Study Of The Influential Factors And Interactions. Analusis 26, 294–299. https://doi.org/10.1051/analusis:1998176

Devi, P., Wahidullah, S., Rodrigues, C., Souza, L.D., 2010. The Sponge-Associated Bacterium *Bacillus licheniformis* SAB1: A Source Of Anti-microbial Compounds. Mar. Drugs 8, 1203–1212. https://doi.org/10.3390/md8041203

Dittmann, E., Gugger, M., Sivonen, K., Fewer, D.P., 2015. Natural Product Biosynthetic Diversity And Comparative Genomics Of The Cyanobacteria. Trends Microbiol. 23, 642–652. https://doi.org/10.1016/j.tim.2015.07.008

Doroghazi, J.R., Metcalf, W.W., 2013. Comparative Genomics Of Actinomycetes With A Focus On Natural Product Biosynthetic Genes. BMC Genomics 14, 1.

Dorrington, R.A., Hilliar, S.H., Kalinski, J.C.J., Krause, R.W.M., McPhail, K.L., Parker-Nance, S., Wlamsley, T.A., Waterworth, S.C., 2016. *Latrunculid* Sponges, Their Microbial Communities And Secondary Metabolites: Connecting Conserved Bacterial Symbionts To Pyrroloiminoquinone Production, In: Planta Medica. Presented at the Abstracts 9th Joint Meeting of AFERP, ASP, GA, JSP, PSE & SIF, Georg Thieme Verlag KG, p. P596. https://doi.org/10.1055/s-0036-1596655

Dou, X., Dong, B., 2019. Origins And Bioactivities Of Natural Compounds Derived From Marine Ascidians And Their Symbionts. Mar. Drugs 17, 670. https://doi.org/10.3390/md17120670

Dragull, K., Beck, J.J., 2012. Isolation Of Natural Products By Ion-Exchange Methods, In: Sarker, S.D., Nahar, L. (Eds.), Natural Products Isolation, Methods In Molecular Biology. Humana Press, Totowa, NJ, pp. 189–219. https://doi.org/10.1007/978-1-61779-624-1_8

Dupont, S., Carré-Mlouka, A., Descarrega, F., Ereskovsky, A., Longeon, A., Mouray, E., Florent, I., Bourguet-Kondracki, M.L., 2014. Diversity And Biological Activities Of The Bacterial Community Associated With The Marine Sponge *Phorbas tenacior* (Porifera, Demospongiae). Lett. Appl. Microbiol. 58, 42–52. https://doi.org/10.1111/lam.12154

El-Saber Batiha, G., Magdy Beshbishy, A., El-Mleeh, A., M. Abdel-Daim, M., Prasad Devkota, H., 2020. Traditional Uses, Bioactive Chemical Constituents, And Pharmacological And

Toxicological Activities Of *Glycyrrhiza glabra* L. (Fabaceae). Biomolecules 10, 352. https://doi.org/10.3390/biom10030352

Elyakov, G.B., Kuznetsova, T., Mikhailov, V.V., Maltsev, I.I., Voinov, V.G., Fedoreyev, S.A., 1991. Brominated Diphenyl Ethers From A Marine Bacterium Associated With The Sponge *Dysidea* sp. Experientia 47, 632–633. https://doi.org/10.1007/BF01949894

Erba, E., Bergamaschi, D., Bassano, L., Damia, G., Ronzoni, S., Faircloth, G.T., D'Incalci, M., 2001. Ecteinascidin-743 (ET-743), A Natural Marine Compound, With A Unique Mechanism Of Action. Eur. J. Cancer Oxf. Engl. 1990 37, 97–105.

Esteves, A.I.S., Cullen, A., Thomas, T., 2017. Competitive Interactions Between Sponge-Associated Bacteria. FEMS Microbiol. Ecol. 93. https://doi.org/10.1093/femsec/fix008

Farrelly, V., Rainey, F.A., Stackebrandt, E., 1995. Effect Of Genome Size And Rrn Gene Copy Number On PCR Amplification Of 16S rRNA Genes From A Mixture Of Bacterial Species. Appl. Environ. Microbiol. 61, 2798–2801.

Feling, R.H., Buchanan, G.O., Mincer, T.J., Kauffman, C.A., Jensen, P.R., Fenical, W., 2003. Salinosporamide A: A Highly Cytotoxic Proteasome Inhibitor From A Novel Microbial Source, A Marine Bacterium Of The New Genus *Salinospora*. Angew. Chem. Int. Ed Engl. 42, 355–357. https://doi.org/10.1002/anie.200390115

Fiedler, H.-P., Bruntner, C., Bull, A.T., Ward, A.C., Goodfellow, M., Potterat, O., Puder, C., Mihm, G., 2005. Marine Actinomycetes As A Source Of Novel Secondary Metabolites. Antonie Van Leeuwenhoek 87, 37–42. https://doi.org/10.1007/s10482-004-6538-8

Filice, G.A., Pomeroy, C., 1991. Preventing Secondary Infections Among HIV-Positive Persons. Public Health Rep. 106, 503–517.

Finking, R., Marahiel, M.A., 2004. Biosynthesis Of Nonribosomal Peptides [1]. Annu. Rev. Microbiol. 58, 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615

Fischbach, M.A., Walsh, C.T., 2006. Assembly-Line Enzymology For Polyketide And Nonribosomal Peptide Antibiotics:? Logic, Machinery, And Mechanisms. Chem. Rev. 106, 3468–3496. https://doi.org/10.1021/cr0503097

Flemer, B., Kennedy, J., Margassery, L.M., Morrissey, J.P., O'Gara, F., Dobson, A.D.W., 2012. Diversity And Anti-microbial Activities Of Microbes From Two Irish Marine Sponges, *Suberites carnosus* and *Leucosolenia* sp. J. Appl. Microbiol. 112, 289–301. https://doi.org/10.1111/j.1365-2672.2011.05211.x

Friedman, N.D., Temkin, E., Carmeli, Y., 2016. The Negative Impact Of Antibiotic Resistance. Clin. Microbiol. Infect. 22, 416–422. https://doi.org/10.1016/j.cmi.2015.12.002

Frisvad, J.C., 2012. Media And Growth Conditions For Induction Of Secondary Metabolite Production. Methods Mol. Biol. Clifton NJ 944, 47–58. https://doi.org/10.1007/978-1-62703-122-6_3

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., Plaza, A., Xia, L., Müller, R., Stewart, A.F., Zhang, Y., 2012. Full-Length Rece Enhances Linear-Linear Homologous Recombination And Facilitates Direct Cloning For Bioprospecting. Nat. Biotechnol. 30, 440–446. https://doi.org/10.1038/nbt.2183

Fujita, M.J., Kimura, N., sakai, A., Ichikawa, Y., Hanya, T., Otsuka, M., 2011. Cloning And Heterologous Expression Of The Vibrioferrin Biosynthetic Gene Cluster From A Marine Metagenomic Library. Biosci. Biotechnol. Biochem. 75, 2283–2287. https://doi.org/10.1271/bbb.110379

Gabrielsen, C., Brede, D.A., Nes, I.F., Diep, D.B., 2014. Circular Bacteriocins: Biosynthesis And Mode Of Action. Appl. Environ. Microbiol. 80, 6854–6862. https://doi.org/10.1128/AEM.02284-14

Galm, U., Shen, B., 2006. Expression Of Biosynthetic Gene Clusters In Heterologous Hosts For Natural Product Production And Combinatorial Biosynthesis. Expert Opin. Drug Discov. 1, 409–437. https://doi.org/10.1517/17460441.1.5.409

Gao, L., Cai, M., Shen, W., Xiao, S., Zhou, X., Zhang, Y., 2013. Engineered Fungal Polyketide Biosynthesis In *Pichia pastoris*: A Potential Excellent Host For Polyketide Production. Microb. Cell Factories 12, 77.

Gao, S.-S., Shang, Z., Li, X.-M., Li, C.-S., Cui, C.-M., Wang, B.-G., 2012. Secondary Metabolites Produced By Solid Fermentation Of The Marine-Derived Fungus *Penicillium commune* QSD-17. Biosci. Biotechnol. Biochem. 76, 358–360. https://doi.org/10.1271/bbb.110332

Garate, L., Sureda, J., Agell, G., Uriz, M.J., 2017. Endosymbiotic Calcifying Bacteria Across Sponge Species And Oceans. Sci. Rep. 7, 43674. https://doi.org/10.1038/srep43674

Gaudêncio, S.P., Pereira, F., 2015. Dereplication: Racing To Speed Up The Natural Products Discovery Process. Nat. Prod. Rep. 32, 779–810. https://doi.org/10.1039/C4NP00134F

Geng, H., Belas, R., 2010. Expression Of Tropodithietic Acid Biosynthesis Is Controlled By A Novel Autoinducer. J. Bacteriol. 192, 4377–4387. https://doi.org/10.1128/JB.00410-10

Ghareeb, M.A., Tammam, M.A., El-Demerdash, A., Atanasov, A.G., 2020. Insights About Clinically Approved And Preclinically Investigated Marine Natural Products. Curr. Res. Biotechnol. 2, 88–102. https://doi.org/10.1016/j.crbiot.2020.09.001

Gillespie, D.E., Brady, S.F., Bettermann, A.D., Cianciotto, N.P., Liles, M.R., Rondon, M.R., Clardy, J., Goodman, R.M., Handelsman, J., 2002. Isolation Of Antibiotics Turbomycin A And B From A Metagenomic Library Of Soil Microbial DNA. Appl. Environ. Microbiol. 68, 4301–4306. https://doi.org/10.1128/AEM.68.9.4301-4306.2002

Gomez-Escribano, J.P., Bibb, M.J., 2011. Engineering *Streptomyces coelicolor* For Heterologous Expression Of Secondary Metabolite Gene Clusters: *Streptomyces* Host For Heterologous Expression Of Gene Clusters. Microb. Biotechnol. 4, 207–215. https://doi.org/10.1111/j.1751-7915.2010.00219.x

Gonzalez, J.B., Fernandez, F.J., Tomasini, A., 2003. Microbial Secondary Metabolites Production And Strain Improvement. Indian J. Biotechnol. 2, 322–333.

Goodwin, R.J.A., Bunch, J., McGinnity, D.F., 2017. Chapter Six - Mass Spectrometry Imaging In Oncology Drug Discovery, In: Drake, R.R., Mcdonnell, L.A. (Eds.), Advances In Cancer Research, Applications Of Mass Spectrometry Imaging To Cancer. Academic Press, pp. 133–171. https://doi.org/10.1016/bs.acr.2016.11.005

Graça, A.P., Bondoso, J., Gaspar, H., Xavier, J.R., Monteiro, M.C., de la Cruz, M., Oves-Costales, D., Vicente, F., Lage, O.M., 2013. Anti-microbial Activity Of Heterotrophic Bacterial Communities From The Marine Sponge *Erylus discophorus* (Astrophorida, Geodiidae). PLoS ONE 8, e78992. https://doi.org/10.1371/journal.pone.0078992

Gram, L., Melchiorsen, J., Bruhn, J.B., 2010. Antibacterial Activity Of Marine Culturable Bacteria Collected From A Global Sampling Of Ocean Surface Waters And Surface Swabs Of Marine Organisms. Mar. Biotechnol. N. Y. N 12, 439–451. https://doi.org/10.1007/s10126-009-9233-y

Griffiths, C.L., Robinson, T.B., Lange, L., Mead, A., 2010. Marine Biodiversity In South Africa: An Evaluation Of Current States Of Knowledge. PLoS One 5, e12008.

Gross, F., Luniak, N., Perlova, O., Gaitatzis, N., Jenke-Kodama, H., Gerth, K., Gottschalk, D., Dittmann, E., Müller, R., 2006. Bacterial Type III Polyketide Synthases: Phylogenetic Analysis And Potential For The Production Of Novel Secondary Metabolites By Heterologous Expression In *Pseudomonads*. Arch. Microbiol. 185, 28–38. https://doi.org/10.1007/s00203-005-0059-3

Gross, H., 2007. Strategies To Unravel The Function Of Orphan Biosynthesis Pathways: Recent Examples And Future Prospects. Appl. Microbiol. Biotechnol. 75, 267–277. https://doi.org/10.1007/s00253-007-0900-5

Gross, H., Stockwell, V.O., Henkels, M.D., Nowak-Thompson, B., Loper, J.E., Gerwick, W.H., 2007. The Genomisotopic Approach: A Systematic Method To Isolate Products Of Orphan Biosynthetic Gene Clusters. Chem. Biol. 14, 53–63. https://doi.org/10.1016/j.chembiol.2006.11.007

Grubbs, K.J., Bleich, R.M., Santa Maria, K.C., Allen, S.E., Farag, S., Shank, E.A., Bowers, A.A., 2017. Large-Scale Bioinformatics Analysis Of *Bacillus* Genomes Uncovers Conserved Roles Of Natural Products In Bacterial Physiology. mSystems 2. https://doi.org/10.1128/mSystems.00040-17

Gulder, T.A.M., Moore, B.S., 2010. Salinosporamide Natural Products: Potent 20S Proteasome Inhibitors As Promising Cancer Chemotherapeutics. Angew. Chem. Int. Ed Engl. 49, 9346–9367. https://doi.org/10.1002/anie.201000728

Gunasekera, S.P., Cross, S.S., 1992. Fistularin 3 And 11-Ketofistularin 3. Feline Leukemia Virus Active Bromotyrosine Metabolites From The Marine Sponge *Aplysina archeri*. J. Nat. Prod. 55, 509–512. https://doi.org/10.1021/np50082a020

Gupta, A., Bhakta, S., Kundu, S., Gupta, M., Srivastava, B.S., Srivastava, R., 2009. Fast-Growing, Non-Infectious And Intracellularly Surviving Drug-Resistant *Mycobacterium aurum*: A Model For High-Throughput Antituberculosis Drug Screening. J. Antimicrob. Chemother. 64, 774–781. https://doi.org/10.1093/jac/dkp279

Hammami, R., Zouhir, A., Ben Hamida, J., Fliss, I., 2007. BACTIBASE: A New Web-Accessible Database For Bacteriocin Characterization. BMC Microbiol. 7, 89. https://doi.org/10.1186/1471-2180-7-89

Han, A.W., Sandy, M., Fishman, B., Trindade-Silva, A.E., Soares, C.A.G., Distel, D.L., Butler, A., Haygood, M.G., 2013. Turnerbactin, A Novel Triscatecholate Siderophore From The Shipworm Endosymbiont *Teredinibacter Turnerae* T7901. PloS One 8, e76151. https://doi.org/10.1371/journal.pone.0076151

Hardoim, C.C.P., Costa, R., 2014. Microbial Communities And Bioactive Compounds In Marine Sponges Of The Family Irciniidae—A Review. Mar. Drugs 12, 5089–5122. https://doi.org/10.3390/md12105089

Harrington, C., Reen, F.J., Mooij, M.J., Stewart, F.A., Chabot, J.-B., Guerra, A.F., Glöckner, F.O., Nielsen, K.F., Gram, L., Dobson, A.D.W., Adams, C., O'Gara, F., 2014. Characterisation of Non-Autoinducing Tropodithietic Acid (TDA) Production From Marine Sponge *Pseudovibrio* Species. Mar. Drugs 12, 5960–5978. https://doi.org/10.3390/md12125960

Harvey, A., 2008. Natural Products In Drug Discovery. Drug Discov. Today 13, 894–901. https://doi.org/10.1016/j.drudis.2008.07.004

Hatherley, R., Brown, D.K., Musyoka, T.M., Penkler, D.L., Faya, N., Lobb, K.A., Tastan Bishop, Ö., 2015. SANCDB: A South African Natural Compound Database. J. Cheminformatics 7. https://doi.org/10.1186/s13321-015-0080-8

Hentschel, U., Schmid, M., Wagner, M., Fieseler, L., Gernert, C., Hacker, J., 2001. Isolation And Phylogenetic Analysis Of Bacteria With Anti-microbial Activities From The Mediterranean Sponges *Aplysina aerophoba* and *Aplysina cavernicola*. FEMS Microbiol. Ecol. 35, 305–312. https://doi.org/10.1111/j.1574-6941.2001.tb00816.x

Herr, D.S., 1945. Synthetic Ion Exchange Resins In The Separation, Recovery, And Concentration Of Thiamine. Ind. Eng. Chem. 37, 631–634. https://doi.org/10.1021/ie50427a011

Higgs, R.E., Zahn, J.A., Gygi, J.D., Hilton, M.D., 2001. Rapid Method To Estimate The Presence Of Secondary Metabolites In Microbial Extracts. Appl. Environ. Microbiol. 67, 371–376. https://doi.org/10.1128/AEM.67.1.371-376.2001

Himes, P.M., Allen, S.E., Hwang, S., Bowers, A.A., 2016. Production Of Sactipeptides In *Escherichia coli*: Probing The Substrate Promiscuity Of Subtilosin A Biosynthesis. ACS Chem. Biol. 11, 1737–1744. https://doi.org/10.1021/acschembio.6b00042

Hirschmann, M., Grundmann, F., Bode, H.B., 2017. Identification And Occurrence Of The Hydroxamate Siderophores Aerobactin, Putrebactin, Avaroferrin And Ochrobactin C As Virulence Factors From Entomopathogenic Bacteria. Environ. Microbiol. 19, 4080–4090. https://doi.org/10.1111/1462-2920.13845

Hochmuth, T., Piel, J., 2009. Polyketide Synthases Of Bacterial Symbionts In Sponges – Evolution-Based Applications In Natural Products Research. Phytochemistry 70, 1841–1849. https://doi.org/10.1016/j.phytochem.2009.04.010

Hoppers, A., Stoudenmire, J., Wu, S., Lopanik, N.B., 2015. Antibiotic Activity And Microbial Community Of The Temperate Sponge, *Haliclona* sp. J. Appl. Microbiol. 118, 419–430. https://doi.org/10.1111/jam.12709

Horinouchi, S., Ueda, K., Nakayama, J., Ikeda, T., 2010. 4.07 - Cell-To-Cell Communications Among Microorganisms, In: Liu, H.-W. (Ben), Mander, L. (Eds.), Comprehensive Natural Products II. Elsevier, Oxford, pp. 283–337. https://doi.org/10.1016/B978-008045382-8.00098-8

Horwood, P., 2004. Evidence For Non-Ribosomal Peptide Synthetase Production Of Cereulide (The Emetic Toxin) In *Bacillus cereus*. FEMS Microbiol. Lett. 236, 319–324. https://doi.org/10.1016/j.femsle.2004.06.004

Hou, Y., Braun, D.R., Michel, C.R., Klassen, J.L., Adnani, N., Wyche, T.P., Bugni, T.S., 2012. Microbial Strain Prioritization Using Metabolomics Tools For The Discovery Of Natural Products. Anal. Chem. 84, 4277–4283. https://doi.org/10.1021/ac202623g

Hu, Y., Chou, W.K.W., Hopson, R., Cane, D.E., 2011. Genome Mining In *Streptomyces clavuligerus*: Expression And Biochemical Characterization Of Two New Cryptic Sesquiterpene Synthases. Chem. Biol. 18, 32–37. https://doi.org/10.1016/j.chembiol.2010.11.008

Huang, C., Sun, C., Lou, I., Tuo, L., Guo, L., Guo, M., Leung, R.K.-K., Lee, S.M.Y., Yew, W.W., 2016. Genome-Guided Investigation Of Antibiotic Substances Produced By *Allosalinactinospora lopnorensis* CA15-2^{T} From Lop Nor Region, China. Sci. Rep. 6, srep20667. https://doi.org/10.1038/srep20667

Hughes, J.P., Rees, S., Kalindjian, S.B., Philpott, K.L., 2011. Principles Of Early Drug Discovery. Br. J. Pharmacol. 162, 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

Huo, L., Rachid, S., Stadler, M., Wenzel, S.C., Müller, R., 2012. Synthetic Biotechnology To Study And Engineer Ribosomal Bottromycin Biosynthesis. Chem. Biol. 19, 1278–1287. https://doi.org/10.1016/j.chembiol.2012.08.013

Ismail, A., Ktari, L., Ahmed, M., Bolhuis, H., Boudabbous, A., Stal, L.J., Cretoiu, M.S., El Bour, M., 2016. Anti-microbial Activities Of Bacteria Associated With The Brown Alga *Padina pavonica*. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.01072

Jackson, S.A., Kennedy, J., Morrissey, J.P., O'Gara, F., Dobson, A.D.W., 2012. Pyrosequencing Reveals Diverse And Distinct Sponge-Specific Microbial Communities In Sponges From A Single Geographical Location In Irish Waters. Microb. Ecol. 64, 105–116. https://doi.org/10.1007/s00248-011-0002-x

Jantan, I., Bukhari, S.N.A., Mohamed, M.A.S., Wai, L.K., Mesaik, M.A., 2015. The Evolving Role Of Natural Products From The Tropical Rainforests As A Replenishable Source Of New Drug Leads, In: Vallisuta, O., Olimat, S. (Eds.), Drug Discovery And Development - From Molecules To Medicine. InTech.

Jaspars, M., Pascale, D.D., Andersen, J.H., Reyes, F., Crawford, A.D., Ianora, A., 2016. The Marine Biodiscovery Pipeline And Ocean Medicines Of Tomorrow. J. Mar. Biol. Assoc. U. K. 96, 151–158. https://doi.org/10.1017/S0025315415002106

Jeganathan, P., Rajasekaran, K.M., Asha Devi, N.K., Karuppusamy, S., 2013. Anti-microbial Activity And Characterization Of Marine Bacteria.

Jenke-Kodama, H., Dittmann, E., 2009. Evolution Of Metabolic Diversity: Insights From Microbial Polyketide Synthases. Phytochemistry 70, 1858–1866. https://doi.org/10.1016/j.phytochem.2009.05.021

Jensen, P.R., Fenical, W., 1994. Strategies For The Discovery Of Secondary Metabolites From Marine Bacteria: Ecological Perspectives. Annu. Rev. Microbiol. 48, 559–584. https://doi.org/10.1146/annurev.mi.48.100194.003015

Ji, H.-F., Li, X.-J., Zhang, H.-Y., 2009. Natural Products And Drug Discovery. EMBO Rep. 10, 194–200.

Joerger, R.D., 2003. Alternatives To Antibiotics: Bacteriocins, Anti-microbial Peptides And Bacteriophages. Poult. Sci. 82, 640–647.

Johnston, I., Osborn, L.J., Markley, R.L., McManus, E.A., Kadam, A., Schultz, K.B., Nagajothi, N., Ahern, P.P., Brown, J.M., Claesen, J., 2021. Identification Of Essential Genes For *Escherichia coli* Aryl Polyene Biosynthesis And Function In Biofilm Formation. Npj Biofilms Microbiomes 7, 1–10. https://doi.org/10.1038/s41522-021-00226-3

Jooste, C.M., Oliver, J., Emami-Khoyi, A., Teske, P.R., 2018. Is The Wild Coast In Eastern South Africa A Distinct Marine Bioregion? Helgol. Mar. Res. 72, 6. https://doi.org/10.1186/s10152-018-0509-3

Ju, K.-S., Gao, J., Doroghazi, J.R., Wang, K.-K.A., Thibodeaux, C.J., Li, S., Metzger, E., Fudala, J., Su, J., Zhang, J.K., others, 2015. Discovery Of Phosphonic Acid Natural Products By Mining The Genomes Of 10,000 Actinomycetes. Proc. Natl. Acad. Sci. 112, 12175–12180.

Julien, B., Shah, S., 2002. Heterologous Expression Of Epothilone Biosynthetic Genes In *Myxococcus xanthus*. Antimicrob. Agents Chemother. 46, 2772–2778. https://doi.org/10.1128/AAC.46.9.2772-2778.2002

Kahn, J.G., Marseille, E., Auvert, B., 2006. Cost-Effectiveness Of Male Circumcision For HIV Prevention In A South African Setting. PLoS Med. 3. https://doi.org/10.1371/journal.pmed.0030517

Kameyama, T., Takahashi, A., Kurasawa, S., Ishizuka, M., Okami, Y., Takeuchi, T., Umezawa, H., 1987. Bisucaberin, A New Siderophore, Sensitizing Tumor Cells To Macrophage-Mediated Cytolysis. I. Taxonomy Of The Producing Organism, Isolation And Biological Properties. J. Antibiot. (Tokyo) 40, 1664–1670. https://doi.org/10.7164/antibiotics.40.1664

Karimi, E., Ramos, M., Gonçalves, J.M.S., Xavier, J.R., Reis, M.P., Costa, R., 2017. Comparative Metagenomics Reveals The Distinctive Adaptive Features Of The *Spongia officinalis* Endosymbiotic Consortium. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.02499

Kato, C., Nogi, Y., 2001. Correlation Between Phylogenetic Structure And Function: Examples From Deep-Sea *Shewanella*. FEMS Microbiol. Ecol. 35, 223–230. https://doi.org/10.1111/j.1574-6941.2001.tb00807.x

Kaur, D.C., Chate, S.S., 2015. Study Of Antibiotic Resistance Pattern In Methicillin-Resistant *Staphylococcus aureus* With Special Reference To Newer Antibiotic. J. Glob. Infect. Dis. 7, 78–84. https://doi.org/10.4103/0974-777X.157245

Kaysser, L., Tang, X., Wemakor, E., Sedding, K., Hennig, S., Siebenberg, S., Gust, B., 2011. Identification Of A Napsamycin Biosynthesis Gene Cluster By Genome Mining. ChemBioChem 12, 477–487. https://doi.org/10.1002/cbic.201000460

Kelecom, A., 2002. Secondary Metabolites From Marine Microorganisms. An. Acad. Bras. Cienc. 74, 151–170.

Kennedy, J., Baker, P., Piper, C., Cotter, P.D., Walsh, M., Mooij, M.J., Bourke, M.B., Rea, M.C., O'Connor, P.M., Ross, R.P., Hill, C., O'Gara, F., Marchesi, J.R., Dobson, A.D.W., 2009. Isolation And Analysis Of Bacteria With Anti-microbial Activities From The Marine Sponge *Haliclona simulans* Collected From Irish Waters. Mar. Biotechnol. 11, 384–396. https://doi.org/10.1007/s10126-008-9154-1

Kennedy, J., Codling, C.E., Jones, B.V., Dobson, A.D.W., Marchesi, J.R., 2008. Diversity Of Microbes Associated With The Marine Sponge, *Haliclona simulans*, Isolated From Irish Waters And Identification Of Polyketide Synthase Genes From The Sponge Metagenome. Environ. Microbiol. 10, 1888–1902. https://doi.org/10.1111/j.1462-2920.2008.01614.x

Kernan, M.R., Cambie, R.C., Bergquist, P.R., 1990. Chemistry of Sponges, VII. 11,19-Dideoxyfistularin 3 And 11-Hydroxyaerothionin, Bromotyrosine Derivatives From *Pseudoceratina durissima*. J. Nat. Prod. 53, 615–622. https://doi.org/10.1021/np50069a012

Khaldi, N., Seifuddin, F.T., Turner, G., Haft, D., Nierman, W.C., Wolfe, K.H., Fedorova, N.D., 2010. SMURF: Genomic Mapping Of Fungal Secondary Metabolite Clusters. Fungal Genet. Biol. FG B 47, 736–741. https://doi.org/10.1016/j.fgb.2010.06.003

Khan, H., Flint, S.H., Yu, P.-L., 2013. Development Of A Chemically Defined Medium For The Production Of Enterolysin A From *Enterococcus faecalis* B9510. J. Appl. Microbiol. 114, 1092–1102. https://doi.org/10.1111/jam.12115

Kildgaard, S., Mansson, M., Dosen, I., Klitgaard, A., Frisvad, J., Larsen, T., Nielsen, K., 2014. Accurate Dereplication Of Bioactive Secondary Metabolites From Marine-Derived Fungi By UHPLC-DAD-QTOFMS And A MS/HRMS Library. Mar. Drugs 12, 3681–3705. https://doi.org/10.3390/md12063681

Kim, H.Y., Heo, D.Y., Park, H.M., Singh, D., Lee, C.H., 2016. Metabolomic And Transcriptomic Comparison Of Solid-State And Submerged Fermentation Of *Penicillium expansum* KACC 40815. PLOS ONE 11, e0149012. https://doi.org/10.1371/journal.pone.0149012

Kim, J.H., Feng, Z., Bauer, J.D., Kallifidas, D., Calle, P.Y., Brady, S.F., 2010. Cloning Large Natural Product Gene Clusters From The Environment: Piecing Environmental DNA Gene Clusters Back Together With TAR. Biopolymers 93, 833–844. https://doi.org/10.1002/bip.21450

Kim, T.K., Fuerst, J.A., 2006. Diversity Of Polyketide Synthase Genes From Bacteria Associated With The Marine Sponge *Pseudoceratina clavata*: Culture-Dependent And Culture-Independent Approaches. Environ. Microbiol. 8, 1460–1470. https://doi.org/10.1111/j.1462-2920.2006.01040.x

Kiran, G., Seghal, Priyadharsini, S., Sajayan, A., Ravindran, A., Selvin, J., 2018. An Antibiotic Agent Pyrrolo[1,2- A]Pyrazine-1,4-Dione, Hexahydro Isolated From A Marine Bacteria *Bacillus* Tequilensis MSI45 Effectively Controls Multi-Drug Resistant *Staphylococcus aureus*. RSC Adv. 8, 17837–17846. https://doi.org/10.1039/C8RA00820E

Kirimura, K., Watanabe, S., Kobayashi, K., 2016. Heterologous Gene Expression And Functional Analysis Of A Type III Polyketide Synthase From *Aspergillus niger* NRRL 328. Biochem. Biophys. Res. Commun. https://doi.org/10.1016/j.bbrc.2016.04.023

Kjeldsen, K.U., Obst, M., Nakano, H., Funch, P., Schramm, A., 2010. Two Types Of Endosymbiotic Bacteria In The Enigmatic Marine Worm *Xenoturbella bocki*. Appl. Environ. Microbiol. 76, 2657–2662. https://doi.org/10.1128/AEM.01092-09

Knappe, T.A., Linne, U., Zirah, S., Rebuffat, S., Xie, X., Marahiel, M.A., 2008. Isolation And Structural Characterization Of Capistruin, A Lasso Peptide Predicted From The Genome

Sequence Of *Burkholderia thailandensis* E264. J. Am. Chem. Soc. 130, 11446–11454. https://doi.org/10.1021/ja802966g

Knight, G.M., Dodd, P.J., Grant, A.D., Fielding, K.L., Churchyard, G.J., White, R.G., 2015. Tuberculosis Prevention In South Africa. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0122514

Koehler, T.M., 2009. *Bacillus anthracis* Physiology And Genetics. Mol. Aspects Med. 30, 386–396. https://doi.org/10.1016/j.mam.2009.07.004

Koehn, F.E., Sarath, G.P., Neil, D.N., Cross, S.S., 1991. Halitunal, An Unusual Diterpene Aldehyde From The Marine Alga *Halimeda tuna*. Tetrahedron Lett. 32, 169–172. https://doi.org/10.1016/0040-4039(91)80845-W

Kronheim, S., Daniel-Ivad, M., Duan, Z., Hwang, S., Wong, A.I., Mantel, I., Nodwell, J.R., Maxwell, K.L., 2018. A Chemical Defence Against Phage Infection. Nature. https://doi.org/10.1038/s41586-018-0767-x

Kudo, F., Matsuura, Y., Hayashi, T., Fukushima, M., Eguchi, T., 2016. Genome Mining Of The Sordarin Biosynthetic Gene Cluster From *Sordaria araneosa Cain* ATCC 36386: characterization Of Cycloaraneosene Synthase And GDP-6-Deoxyaltrose Transferase. J. Antibiot. (Tokyo) 69, 541–548. https://doi.org/10.1038/ja.2016.40

Kuhlmann, A.U., Hoffmann, T., Bursy, J., Jebbar, M., Bremer, E., 2011. Ectoine And Hydroxyectoine As Protectants Against Osmotic And Cold Stress: Uptake Through The Sigb-Controlled Betaine-Choline- Carnitine Transporter-Type Carrier Ectt From *Virgibacillus pantothenticus*. J. Bacteriol. 193, 4699. https://doi.org/10.1128/JB.05270-11

Kumar, D., 2015. Nuclear Magnetic Resonance (NMR) Spectroscopy: Metabolic Profiling Of Medicinal Plants And Their Products. Crit. Rev. Anal. Chem. 00–00. https://doi.org/10.1080/10408347.2015.1106932

Kunwar, A., Priyadarsini, K.I., 2011. Free Radicals, Oxidative Stress And Importance Of Antioxidants In Human Health - Proquest [WWW Document]. URL https://search.proquest.com/openview/f2055f0ed1e921ddf53dcdfc94a625f9/1?pq-origsite=gscholar&cbl=706338 (accessed 12.6.17).

Kurutas, E.B., 2016. The Importance Of Antioxidants Which Play The Role In Cellular Response Against Oxidative/Nitrosative Stress: Current State. Nutr. J. 15, 71. https://doi.org/10.1186/s12937-016-0186-5

Laux-Biehlmann, A., Mouheiche, J., Vérièpe, J., Goumon, Y., 2013. Endogenous Morphine And Its Metabolites In Mammals: History, Synthesis, Localization And Perspectives. Neuroscience 233, 95–117. https://doi.org/10.1016/j.neuroscience.2012.12.013

Leal, M.C., Puga, J., Serôdio, J., Gomes, N.C.M., Calado, R., 2012. Trends In The Discovery Of New Marine Natural Products From Invertebrates Over The Last Two Decades – Where And What Are We Bioprospecting? PLOS ONE 7, e30580. https://doi.org/10.1371/journal.pone.0030580

Leal, M.C., Sheridan, C., Osinga, R., Dionísio, G., Rocha, R.J.M., Silva, B., Rosa, R., Calado, R., 2014. Marine Microorganism-Invertebrate Assemblages: Perspectives To Solve The "Supply Problem" In The Initial Steps Of Drug Discovery. Mar. Drugs 12, 3929–3952. https://doi.org/10.3390/md12073929

Lee, J.-K., Lim, H.-S., Kim, J.-H., 2002. Cytotoxic Activity Of Aminoderivatized Cationic Chitosan Derivatives. Bioorg. Med. Chem. Lett. 12, 2949–2951.

Lee, J.Y., Passalacqua, K.D., Hanna, P.C., Sherman, D.H., 2011. Regulation Of Petrobactin And Bacillibactin Biosynthesis In *Bacillus anthracis* Under Iron And Oxygen Variation. PLOS ONE 6, e20777. https://doi.org/10.1371/journal.pone.0020777

Leisch, M., Egle, A., Greil, R., 2019. Plitidepsin: A Potential New Treatment For Relapsed/Refractory Multiple Myeloma. Future Oncol. Lond. Engl. 15, 109–120. https://doi.org/10.2217/fon-2018-0492

Letzel, A.-C., Pidot, S.J., Hertweck, C., 2014. Genome Mining For Ribosomally Synthesized And Post-Translationally Modified Peptides (RiPPs) in Anaerobic Bacteria. BMC Genomics 15, 983.

Levatin, J.L., n.d. Pediatric Diagnosis And Treatment 5.

Li, H., Huang, H., Hou, L., Ju, J., Li, W., 2017. Discovery Of Antimycin-Type Depsipeptides From A Wbl Gene Mutant Strain Of Deepsea-Derived *Streptomyces somaliensis* SCSIO ZH66 And Their Effects On Pro-Inflammatory Cytokine Production. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.00678

Li, K., Chung-Davidson, Y.-W., Bussy, U., Li, W., 2015. Recent Advances And Applications Of Experimental Technologies In Marine Natural Product Research. Mar. Drugs 13, 2694–2713. https://doi.org/10.3390/md13052694

Lin, Y., Schiavo, S., Orjala, J., Vouros, P., Kautz, R., 2008. Microscale LC-MS-NMR Platform Applied To The Identification Of Active Cyanobacterial Metabolites. Anal. Chem. 80, 8045–8054. https://doi.org/10.1021/ac801049k

Lindequist, U., 2016. Marine-Derived Pharmaceuticals – Challenges And Opportunities. Biomol. Ther. 24, 561–571. https://doi.org/10.4062/biomolther.2016.181

Liu, M., Li, A., 2011. [Quorum Sensing Involved In The Regulation Of Secondary Metabolism In *Streptomycetes*--a review]. Wei Sheng Wu Xue Bao 51, 571–578.

Lorente, A., Makowski, K., Albericio, F., Álvarez, M., 2014. Bioactive Marine Polyketides As Potential And Promising Drugs. Ann. Mar. Biol. Res. 1, 1003.

Low, C.-Y., Rotstein, C., 2011. Emerging Fungal Infections In Immunocompromised Patients. F1000 Med. Rep. 3, 14. https://doi.org/10.3410/M3-14

Machado, H., Sonnenschein, E.C., Melchiorsen, J., Gram, L., 2015. Genome Mining Reveals Unlocked Bioactive Potential Of Marine Gram-Negative Bacteria. BMC Genomics 16, 158. https://doi.org/10.1186/s12864-015-1365-z

Macintyre, L., Zhang, T., Viegelmann, C., Martinez, I., Cheng, C., Dowdells, C., Abdelmohsen, U., Gernert, C., Hentschel, U., Edrada-Ebel, R., 2014. Metabolomic Tools For Secondary Metabolite Discovery From Marine Microbial Symbionts. Mar. Drugs 12, 3416–3448. https://doi.org/10.3390/md12063416

Maduray, S., 2013. Studies On The Diversity And Spatial Distribution Of Deep-Water Sponges Along The West And South Coasts Of South Africa (Thesis). University of the Western Cape.

Maksimov, M.O., Pelczer, I., Link, A.J., 2012. Precursor-Centric Genome-Mining Approach For Lasso Peptide Discovery. Proc. Natl. Acad. Sci. U. S. A. 109, 15223–15228. https://doi.org/10.1073/pnas.1208978109

Malve, H., 2016. Exploring The Ocean For New Drug Developments: Marine Pharmacology. J. Pharm. Bioallied Sci. 8, 83–91. https://doi.org/10.4103/0975-7406.171700

Maneveldt, G.W., Chamberlain, Y.M., Keats, D.W., 2008. A Catalogue With Keys To The Non-Geniculate Coralline Algae (*Corallinales*, Rhodophyta) of South Africa. South Afr. J. Bot. 74, 555–566. https://doi.org/10.1016/j.sajb.2008.02.002

Mansson, M., Gram, L., Larsen, T.O., 2011. Production Of Bioactive Secondary Metabolites By Marine Vibrionaceae. Mar. Drugs 9, 1440–1468. https://doi.org/10.3390/md9091440

Margassery, L.M., Kennedy, J., O'Gara, F., Dobson, A.D., Morrissey, J.P., 2012. Diversity And Antibacterial Activity Of Bacteria Isolated From The Coastal Marine Sponges Amphilectus Fucorum And Eurypon Major: Diversity And Antibacterial Activity Of Bacteria. Lett. Appl. Microbiol. 55, 2–8. https://doi.org/10.1111/j.1472-765X.2012.03256.x

Martín, J., da S. Sousa, T., Crespo, G., Palomo, S., González, I., Tormo, J., de la Cruz, M., Anderson, M., Hill, R., Vicente, F., Genilloud, O., Reyes, F., 2013. Kocurin, The True Structure Of PM181104, An Anti-Methicillin-Resistant *Staphylococcus aureus* (MRSA) Thiazolyl Peptide From The Marine-Derived Bacterium *Kocuria palustris*. Mar. Drugs 11, 387–398. https://doi.org/10.3390/md11020387

Martinez, J.L., 2014. General Principles Of Antibiotic Resistance In Bacteria. Drug Discov. Today Technol. 11, 33–39. https://doi.org/10.1016/j.ddtec.2014.02.001

Martins, A., Vieira, H., Gaspar, H., Santos, S., 2014. Marketed Marine Natural Products In The Pharmaceutical And Cosmeceutical Industries: Tips For Success. Mar. Drugs 12, 1066–1101. https://doi.org/10.3390/md12021066

Marwick, null, Wright, null, Burgess, null, 1999. Bioprocess Intensification For Production Of Novel Marine Bacterial Antibiotics Through Bioreactor Operation And Design. Mar. Biotechnol. N. Y. N 1, 495–0507. https://doi.org/10.1007/pl00011806

Maslin, M., Gaertner-Mazouni, N., Debitus, C., Joy, N., Ho, R., 2021. Marine Sponge Aquaculture Towards Drug Development: An Ongoing History Of Technical, Ecological, Chemical Considerations And Challenges. Aquac. Rep. 21, 100813. https://doi.org/10.1016/j.aqrep.2021.100813

Matcher, G.F., Waterworth, S.C., Walmsley, T.A., Matsatsa, T., Parker-Nance, S., Davies-Coleman, M.T., Dorrington, R.A., 2017. Keeping It In The Family: Coevolution Of Latrunculid Sponges And Their Dominant Bacterial Symbionts. MicrobiologyOpen 6, n/a-n/a. https://doi.org/10.1002/mbo3.417

Matobole, R., van Zyl, L., Parker-Nance, S., Davies-Coleman, M., Trindade, M., 2017. Antibacterial Activities Of Bacteria Isolated From The Marine Sponges *Isodictya compressa* and *Higginsia bidentifera* Collected from Algoa Bay, South Africa. Mar. Drugs 15, 47. https://doi.org/10.3390/md15020047

Matsuda, Y., Mitsuhashi, T., Quan, Z., Abe, I., 2015. Molecular Basis For Stellatic Acid Biosynthesis: A Genome Mining Approach For Discovery Of Sesterterpene Synthases. Org. Lett. 150909104027006. https://doi.org/10.1021/acs.orglett.5b02404

Máximo, P., Ferreira, L.M., Branco, P., Lima, P., Lourenço, A., 2016. The Role Of Spongia Sp. In The Discovery Of Marine Lead Compounds. Mar. Drugs 14. https://doi.org/10.3390/md14080139

McAlpine, J.B., Bachmann, B.O., Piraee, M., Tremblay, S., Alarco, A.-M., Zazopoulos, E., Farnet, C.M., 2005. Microbial Genomics As A Guide To Drug Discovery And Structural Elucidation: ECO-02301, A Novel Antifungal Agent, As An Example †. J. Nat. Prod. 68, 493–496. https://doi.org/10.1021/np0401664

McClure, R.A., Goering, A.W., Ju, K.-S., Baccile, J.A., Schroeder, F.C., Metcalf, W.W., Thomson, R.J., Kelleher, Neil.L., 2016. Elucidating The Rimosamide-Detoxin Natural Product Families And Their Biosynthesis Using Metabolite/Gene Cluster Correlations. ACS Chem. Biol. 11, 3452–3460. https://doi.org/10.1021/acschembio.6b00779

Mehbub, M., Lei, J., Franco, C., Zhang, W., 2014. Marine Sponge Derived Natural Products Between 2001 And 2010: Trends And Opportunities For Discovery Of Bioactives. Mar. Drugs 12, 4539–4577. https://doi.org/10.3390/md12084539

Memoli, M.J., Athota, R., Reed, S., Czajkowski, L., Bristol, T., Proudfoot, K., Hagey, R., Voell, J., Fiorentino, C., Ademposi, A., Shoham, S., Taubenberger, J.K., 2014. The Natural History Of Influenza Infection In The Severely Immunocompromised Vs Nonimmunocompromised Hosts. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 58, 214–224. https://doi.org/10.1093/cid/cit725

Mendelson, M., n.d. Anti-microbial Resistance Background Document. Antimicrob. Resist. 16.

Millar, D.R.H., 1954. XXI.—*Pseudodistoma africanum*, sp. n., A New Compound Ascidian From South Africa. Ann. Mag. Nat. Hist. 7, 128–132. https://doi.org/10.1080/00222935408651705

Mishra, B.B., Tiwari, V.K., 2011. Natural Products: An Evolving Role In Future Drug Discovery. Eur. J. Med. Chem. 46, 4769–4807. https://doi.org/10.1016/j.ejmech.2011.07.057

Molinari, G., 2009. Natural Products In Drug Discovery: Present Status And Perspectives, In: Pharmaceutical Biotechnology, Advances In Experimental Medicine And Biology. Springer, New York, NY, pp. 13–27. https://doi.org/10.1007/978-1-4419-1132-2_2

Mondol, M.A.M., Shin, H.J., Islam, M.T., 2013. Diversity Of Secondary Metabolites From Marine Bacillus Species: Chemistry And Biological Activity. Mar. Drugs 11, 2846–2872. https://doi.org/10.3390/md11082846

Monteiro, M.C., de la Cruz, M., Cantizani, J., Moreno, C., Tormo, J.R., Mellado, E., De Lucas, J.R., Asensio, F., Valiante, V., Brakhage, A.A., Latge, J.-P., Genilloud, O., Vicente, F., 2012. A New Approach To Drug Discovery: High-Throughput Screening Of Microbial Natural Extracts Against *Aspergillus fumigatus* Using Resazurin. J. Biomol. Screen. 17, 542–549. https://doi.org/10.1177/1087057111433459

Müller, S., Garcia-Gonzalez, E., Mainz, A., Hertlein, G., Heid, N.C., Mösker, E., van den Elst, H., Overkleeft, H.S., Genersch, E., Süssmuth, R.D., 2014. Paenilamicin: Structure And Biosynthesis Of A Hybrid Nonribosomal Peptide/Polyketide Antibiotic From The Bee Pathogen *Paenibacillus larvae*. Angew. Chem. Int. Ed. 53, 10821–10825. https://doi.org/10.1002/anie.201404572

Nah, H.-J., Pyeon, H.-R., Kang, S.-H., Choi, S.-S., Kim, E.-S., 2017. Cloning And Heterologous Expression Of A Large-Sized Natural Product Biosynthetic Gene Cluster In *Streptomyces* Species. Front. Microbiol. 8, 394. https://doi.org/10.3389/fmicb.2017.00394

Nakashima, Y., Egami, Y., Kimura, M., Wakimoto, T., Abe, I., 2016. Metagenomic Analysis Of The Sponge *Discodermia* Reveals The Production Of The Cyanobacterial Natural Product Kasumigamide By 'Entotheonella.' PLOS ONE 11, e0164468. https://doi.org/10.1371/journal.pone.0164468

Naughton, L.M., Romano, S., O'Gara, F., Dobson, A.D.W., 2017. Identification Of Secondary Metabolite Gene Clusters In The *Pseudovibrio* Genus Reveals Encouraging Biosynthetic Potential Toward The Production Of Novel Bioactive Compounds. Front. Microbiol. 8, 1494. https://doi.org/10.3389/fmicb.2017.01494

Neilan, B.A., Dittmann, E., Rouhiainen, L., Bass, R.A., Schaub, V., Sivonen, K., Börner, T., 1999. Nonribosomal Peptide Synthesis And Toxigenicity Of Cyanobacteria. J. Bacteriol. 181, 4089–4097.

Nemeth, J., Oesch, G., Kuster, S.P., 2015. Bacteriostatic Versus Bactericidal Antibiotics For Patients With Serious Bacterial Infections: Systematic Review And Meta-Analysis. J. Antimicrob. Chemother. 70, 382–395. https://doi.org/10.1093/jac/dku379

Newman, D., 2017. Screening And Identification Of Novel Biologically Active Natural Compounds. F1000Research 6. https://doi.org/10.12688/f1000research.11221.1

Newman, D., Cragg, G., 2014. Marine-Sourced Anti-Cancer And Cancer Pain Control Agents In Clinical And Late Preclinical Development. Mar. Drugs 12, 255–278. https://doi.org/10.3390/md12010255

Newman, D.J., Cragg, G.M., 2020. Natural Products As Sources Of New Drugs Over The Nearly Four Decades From 01/1981 To 09/2019. J. Nat. Prod. 83, 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285

Newman, D.J., Cragg, G.M., 2016. Natural Products As Sources Of New Drugs From 1981 To 2014. J. Nat. Prod. 79, 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055

Newman, D.J., Cragg, G.M., 2012. Natural Products As Sources Of New Drugs Over The 30 Years From 1981 To 2010. J. Nat. Prod. 75, 311–335. https://doi.org/10.1021/np200906s

Newman, D.J., Cragg, G.M., 2007. Natural Products As Sources Of New Drugs Over The Last 25 Years. J. Nat. Prod. 70, 461–477. https://doi.org/10.1021/np068054v

Newman, D.J., Cragg, G.M., Snader, K.M., 2000. The Influence Of Natural Products Upon Drug Discovery. Nat. Prod. Rep. 17, 215–234. https://doi.org/10.1039/a902202c

Nicacio, K.J., Ióca, L.P., Fróes, A.M., Leomil, L., Appolinario, L.R., Thompson, C.C., Thompson, F.L., Ferreira, A.G., Williams, D.E., Andersen, R.J., Eustaquio, A.S., Berlinck, R.G.S., 2017. Cultures Of The Marine Bacterium *Pseudovibrio denitrificans* Ab134 Produce Bromotyrosine-Derived Alkaloids Previously Only Isolated From Marine Sponges. J. Nat. Prod. 80, 235–240. https://doi.org/10.1021/acs.jnatprod.6b00838

Nichols, D.S., 2003. Prokaryotes And The Input Of Polyunsaturated Fatty Acids To The Marine Food Web. FEMS Microbiol. Lett. 219, 1–7.

Nielsen, K.F., Månsson, M., Rank, C., Frisvad, J.C., Larsen, T.O., 2011. Dereplication Of Microbial Natural Products by LC-DAD-TOFMS. J. Nat. Prod. 74, 2338–2348. https://doi.org/10.1021/np200254t

Nightingale, K., 2019. The Bite That Cures: How We're Turning Venom Into Medicine [WWW Document]. BBC Sci. Focus Mag. URL https://www.sciencefocus.com/nature/the-bite-that-cures-how-were-turning-venom-into-medicine/ (accessed 7.2.20).

Noyer, C., Thomas, O.P., Becerro, M.A., 2011. Patterns Of Chemical Diversity In The Mediterranean Sponge *Spongia lamella*. PLoS ONE 6. https://doi.org/10.1371/journal.pone.0020844

Nuñez, C.V., de Almeida, E.V.R., Granato, A.C., Marques, S.O., Santos, K.O., Pereira, F.R., Macedo, M.L., Ferreira, A.G., Hajdu, E., Pinheiro, U.S., Muricy, G., Peixinho, S., Freeman, C.J., Gleason, D.F., Berlinck, R.G.S., 2008. Chemical Variability Within The Marine Sponge *Aplysina fulva*. Biochem. Syst. Ecol. 36, 283–296. https://doi.org/10.1016/j.bse.2007.09.008

Odom, A.R., 2011. Five Questions About Non-Mevalonate Isoprenoid Biosynthesis. PLoS Pathog. 7, e1002323. https://doi.org/10.1371/journal.ppat.1002323

Oh, D.-C., Kauffman, C.A., Jensen, P.R., Fenical, W., 2007. Induced Production Of Emericellamides A And B From The Marine-Derived Fungus *Emericella* sp. In Competing Co-Culture. J. Nat. Prod. 70, 515–520. https://doi.org/10.1021/np060381f

Ojima, I., 2008. Modern Natural Products Chemistry And Drug Discovery. J. Med. Chem. 51, 2587–2588. https://doi.org/10.1021/jm701291u

Okamoto-Hosoya, Y., Okamoto, S., Ochi, K., 2003. Development Of Antibiotic-Overproducing Strains By Site-Directed Mutagenesis Of The rpsL Gene in *Streptomyces lividans*. Appl. Environ. Microbiol. 69, 4256–4259. https://doi.org/10.1128/AEM.69.7.4256-4259.2003

Orhan, I.E., Tumen, I., 2015. Potential Of *Cupressus sempervirens* (Mediterranean Cypress) In Health. Mediterr. Diet 639–647. https://doi.org/10.1016/B978-0-12-407849-9.00057-9

Ostroumova, O.S., Efimova, S.S., Malev, V.V., 2015. Chapter Six - Modifiers Of Membrane Dipole Potentials As Tools For Investigating Ion Channel Formation And Functioning, In: Jeon, K.W. (Ed.), International Review of Cell and Molecular Biology. Academic Press, pp. 245–297. https://doi.org/10.1016/bs.ircmb.2014.12.001

Oyon, B., 2018. What Are The Pharmaceutical Sources Of Drugs? [WWW Document]. HubPages. URL https://hubpages.com/education/Where-do-drugs-come-from-Sources-of-Drugs (accessed 7.2.20).

Padmalayam, I., 2016. Assay Development In Drug Discovery 40.

Pappas, P.G., Alexander, B.D., Andes, D.R., Hadley, S., Kauffman, C.A., Freifeld, A., Anaissie, E.J., Brumble, L.M., Herwaldt, L., Ito, J., Kontoyiannis, D.P., Lyon, G.M., Marr, K.A., Morrison, V.A., Park, B.J., Patterson, T.F., Perl, T.M., Oster, R.A., Schuster, M.G., Walker, R., Walsh, T.J., Wannemuehler, K.A., Chiller, T.M., 2010. Invasive Fungal Infections Among Organ Transplant Recipients: Results Of The Transplant-Associated Infection Surveillance Network (TRANSNET). Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 50, 1101–1111. https://doi.org/10.1086/651262

Papp-Wallace, K.M., Endimiani, A., Taracila, M.A., Bonomo, R.A., 2011. Carbapenems: Past, Present, And Future ▿. Antimicrob. Agents Chemother. 55, 4943–4960. https://doi.org/10.1128/AAC.00296-11

Pattanaik, B., Lindberg, P., 2015. Terpenoids And Their Biosynthesis In Cyanobacteria. Life 5, 269–293. https://doi.org/10.3390/life5010269

Paul, D., Sinha, S.N., 2016. Isolation And Characterization Of Phosphate Solubilizing Bacterium *Pseudomonas aeruginosa* KUPSB12 With Antibacterial Potential From River Ganga, India. Ann. Agrar. Sci. https://doi.org/10.1016/j.aasci.2016.10.001

Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S.A., Mavrodi, D.V., DeBoy, R.T., Seshadri, R., Ren, Q., Madupu, R., Dodson, R.J., Durkin, A.S., Brinkac, L.M., Daugherty, S.C., Sullivan, S.A., Rosovitz, M.J., Gwinn, M.L., Zhou, L., Schneider, D.J., Cartinhour, S.W., Nelson, W.C., Weidman, J., Watkins, K., Tran, K., Khouri, H., Pierson, E.A., Pierson, L.S., Thomashow, L.S., Loper, J.E., 2005. Complete Genome Sequence Of The Plant Commensal *Pseudomonas fluorescens* Pf-5. Nat. Biotechnol. 23, 873–878. https://doi.org/10.1038/nbt1110

Payne, D.J., Gwynn, M.N., Holmes, D.J., Pompliano, D.L., 2007. Drugs For Bad Bugs: Confronting The Challenges Of Antibacterial Discovery. Nat. Rev. Drug Discov. 6, 29–40. https://doi.org/10.1038/nrd2201

Peng, H., Ouyang, Y., Bilal, M., Wang, W., Hu, H., Zhang, X., 2018. Identification, Synthesis And Regulatory Function Of The N-Acylated Homoserine Lactone Signals Produced By *Pseudomonas chlororaphis* HT66. Microb. Cell Factories 17, 9. https://doi.org/10.1186/s12934-017-0854-y

Peralta-Yahya, P.P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J.D., Lee, T.S., 2011. Identification And Microbial Production Of A Terpene-Based Advanced Biofuel. Nat. Commun. 2, 483. https://doi.org/10.1038/ncomms1494

Pereira, F., 2019. Have Marine Natural Product Drug Discovery Efforts Been Productive And How Can We Improve Their Efficiency? Expert Opin. Drug Discov. 14, 717–722. https://doi.org/10.1080/17460441.2019.1604675

Pérez, J., Muñoz-Dorado, J., Braña, A.F., Shimkets, L.J., Sevillano, L., Santamaría, R.I., 2011. *Myxococcus xanthus* Induces Actinorhodin Overproduction And Aerial Mycelium Formation By *Streptomyces coelicolor*. Microb. Biotechnol. 4, 175–183. https://doi.org/10.1111/j.1751-7915.2010.00208.x

Pérez, L.B., Fenical, W., 2017. Chapter 8 - Accessing Marine Microbial Diversity For Drug Discovery, In: Kurtböke, I. (Ed.), Microbial Resources. Academic Press, pp. 169–187. https://doi.org/10.1016/B978-0-12-804765-1.00008-4

Perovic, O., Ismail, H., Schalkwyk, E.V., 2018. Anti-microbial Resistance Surveillance In The South African Public Sector. South. Afr. J. Infect. Dis. 33, 118–129. https://doi.org/10.1080/23120053.2018.1469851

Pettit, G.R., Herald, C.L., Cichacz, Z.A., Gao, F., Schmidt, J.M., Boyd, M.R., Christie, N.D., Boettner, F.E., 1993. Isolation And Structure Of The Powerful Human Cancer Cell Growth Inhibitors Spongistatins 4 And 5 From An African *Spirastrella spinispirulifera* (porifera). J. Chem. Soc. Chem. Commun. 1805–1807. https://doi.org/10.1039/C39930001805

Pettit, G.R., Inoue, Masuo., Kamano, Yoshiaki., Herald, D.L., Arm, Christoph., Dufresne, Claude., Christie, N.D., Schmidt, J.M., Doubek, D.L., Krupa, T.S., 1988. Antineoplastic Agents. 147. Isolation And Structure Of The Powerful Cell Growth Inhibitor Cephalostatin 1. J. Am. Chem. Soc. 110, 2006–2007. https://doi.org/10.1021/ja00214a078

Pfeifer, B.A., Wang, C.C.C., Walsh, C.T., Khosla, C., 2003. Biosynthesis Of Yersiniabactin, A Complex Polyketide-Nonribosomal Peptide, Using *Escherichia coli* As A Heterologous Host. Appl. Environ. Microbiol. 69, 6698–6702. https://doi.org/10.1128/AEM.69.11.6698-6702.2003

Phelan, J., Maitra, A., McNerney, R., Nair, M., Gupta, A., Coll, F., Pain, A., Bhakta, S., Clark, T.G., 2015. The Draft Genome Of Mycobacterium Aurum, A Potential Model Organism For Investigating Drugs Against *Mycobacterium tuberculosis* and *Mycobacterium leprae*. Int. J. Mycobacteriology 4, 207–216. https://doi.org/10.1016/j.ijmyco.2015.05.001

Phelan, R.W., O'Halloran, J.A., Kennedy, J., Morrissey, J.P., Dobson, A.D.W., O'Gara, F., Barbosa, T.M., 2012. Diversity And Bioactive Potential Of Endospore-Forming Bacteria Cultured From The Marine Sponge *Haliclona simulans*: Sporeformers from the marine sponge *H. simulans*. J. Appl. Microbiol. 112, 65–78. https://doi.org/10.1111/j.1365-2672.2011.05173.x

Phillips, T., Chase, M., Wagner, S., Renzi, C., Powell, M., DeAngelo, J., Michels, P., 2013. Use Of In Situ Solid-Phase Adsorption In Microbial Natural Product Fermentation

Development. J. Ind. Microbiol. Biotechnol. 40, 411–425. https://doi.org/10.1007/s10295-013-1247-9

Phillipson, J.D., 2001. Phytochemistry And Medicinal Plants. Phytochemistry 56, 237–243.

Picot, A., Grenouillet, P., 1994. Safety In The Chemistry And Biochemistry Laboratory. John Wiley & Sons.

Piel, J., 2009. Metabolites From Symbiotic Bacteria. Nat. Prod. Rep. 26, 338–362. https://doi.org/10.1039/b703499g

Piel, J., 2006. Bacterial Symbionts: Prospects For The Sustainable Production Of Invertebrate-Derived Pharmaceuticals. Curr. Med. Chem. 13, 39–50.

Piel, J., 2004. Metabolites From Symbiotic Bacteria. Nat. Prod. Rep. 21, 519–538. https://doi.org/10.1039/B310175B

Poelarends, G.J., Kulakov, L.A., Larkin, M.J., van Hylckama Vlieg, J.E., Janssen, D.B., 2000. Roles Of Horizontal Gene Transfer And Gene Integration In Evolution Of 1,3-Dichloropropene- And 1,2-Dibromoethane-Degradative Pathways. J. Bacteriol. 182, 2191–2199. https://doi.org/10.1128/jb.182.8.2191-2199.2000

Pohle, S., Appelt, C., Roux, M., Fiedler, H.-P., Süssmuth, R.D., 2011. Biosynthetic Gene Cluster Of The Non-Ribosomally Synthesized Cyclodepsipeptide Skyllamycin: Deciphering Unprecedented Ways Of Unusual Hydroxylation Reactions. J. Am. Chem. Soc. 133, 6194–6205. https://doi.org/10.1021/ja108971p

Poplawsky, A.R., Urban, S.C., Chun, W., 2000. Biological Role Of Xanthomonadin Pigments In *Xanthomonas campestris* pv. *campestris*. Appl. Environ. Microbiol. 66, 5123–5127.

Porsby, C.H., Webber, M.A., Nielsen, K.F., Piddock, L.J.V., Gram, L., 2011. Resistance And Tolerance To Tropodithietic Acid, An Anti-microbial In Aquaculture, Is Hard To Select. Antimicrob. Agents Chemother. 55, 1332–1337. https://doi.org/10.1128/AAC.01222-10

Puebla, L.E.J., 2012. Fungal Infections In Immunosuppressed Patients, Immunodeficiency. IntechOpen. https://doi.org/10.5772/51512

Raina, J.-B., Tapiolas, D., Motti, C.A., Foret, S., Seemann, T., Tebben, J., Willis, B.L., Bourne, D.G., 2016. Isolation Of An Anti-microbial Compound Produced By Bacteria Associated With Reef-Building Corals. PeerJ 4, e2275. https://doi.org/10.7717/peerj.2275

Rajan, B.M., Kannabiran, K., 2014. Extraction And Identification Of Antibacterial Secondary Metabolites From Marine *Streptomyces* sp. VITBRK2. Int. J. Mol. Cell. Med. 3, 130–137.

Ramachandran, S., Singh, M., Mandal, M., 2012. Purification Of Azurin From *Pseudomonas Aeuroginosa*, in: Calderon, L. (Ed.), Chromatography - The Most Versatile Method of Chemical Analysis. InTech.

Rateb, M.E., Houssen, W.E., Arnold, M., Abdelrahman, M.H., Deng, H., Harrison, W.T.A., Okoro, C.K., Asenjo, J.A., Andrews, B.A., Ferguson, G., Bull, A.T., Goodfellow, M., Ebel, R., Jaspars, M., 2011. Chaxamycins A–D, Bioactive Ansamycins From A Hyper-Arid Desert *Streptomyces* sp. J. Nat. Prod. 74, 1491–1499. https://doi.org/10.1021/np200320u

Rausch, C., Hoof, I., Weber, T., Wohlleben, W., Huson, D.H., 2007. Phylogenetic Analysis Of Condensation Domains In NRPS Sheds Light On Their Functional Evolution. BMC Evol. Biol. 7, 78. https://doi.org/10.1186/1471-2148-7-78

Reddy, B.V.B., Kallifidas, D., Kim, J.H., Charlop-Powers, Z., Feng, Z., Brady, S.F., 2012. Natural Product Biosynthetic Gene Diversity In Geographically Distinct Soil Microbiomes. Appl. Environ. Microbiol. 78, 3744–3752. https://doi.org/10.1128/AEM.00102-12

Reshetnikov, A.S., Khmelenina, V.N., Mustakhimov, I.I., Trotsenko, Y.A., 2011. Genes And Enzymes Of Ectoine Biosynthesis In Halotolerant Methanotrophs. Methods Enzymol. 495, 15–30. https://doi.org/10.1016/B978-0-12-386905-0.00002-4

Reygaert, W.C., 2018. An Overview Of The Anti-microbial Resistance Mechanisms Of Bacteria. AIMS Microbiol. 4, 482–501. https://doi.org/10.3934/microbiol.2018.3.482

Reysenbach, A.-L., Giver, L.J., Wickham, G.S., Pace, N.R., 1992. Differential Amplification Of rRNA Genes By Polymerase Chain Reaction. Appl. Environ. Microbiol. 58, 3417–3418.

Richter, M., Rosselló-Móra, R., 2009. Shifting The Genomic Gold Standard For The Prokaryotic Species Definition. Proc. Natl. Acad. Sci. U. S. A. 106, 19126–19131. https://doi.org/10.1073/pnas.0906412106

Ritacco, F.V., Haltli, B., Janso, J.E., Greenstein, M., Bernan, V.S., 2003. Dereplication Of *Streptomyces* Soil Isolates And Detection Of Specific Biosynthetic Genes Using An Automated Ribotyping Instrument. J. Ind. Microbiol. Biotechnol. 30, 472–479. https://doi.org/10.1007/s10295-003-0038-0

Roca, I., Akova, M., Baquero, F., Carlet, J., Cavaleri, M., Coenen, S., Cohen, J., Findlay, D., Gyssens, I., Heure, O.E., Kahlmeter, G., Kruse, H., Laxminarayan, R., Liébana, E., López-Cerero, L., MacGowan, A., Martins, M., Rodríguez-Baño, J., Rolain, J.-M., Segovia, C., Sigauque, B., Tacconelli, E., Wellington, E., Vila, J., 2015. The Global Threat Of Anti-microbial Resistance: Science For Intervention. New Microbes New Infect. 6, 22–29. https://doi.org/10.1016/j.nmni.2015.02.007

Rocha-Roa, C., Molina, D., Cardona, N., 2018. A Perspective On Thiazolidinone Scaffold Development As A New Therapeutic Strategy For Toxoplasmosis. Front. Cell. Infect. Microbiol. 8, 360. https://doi.org/10.3389/fcimb.2018.00360

Romano, S., 2018. Ecology And Biotechnological Potential Of Bacteria Belonging To The Genus *Pseudovibrio*. Appl. Environ. Microbiol. 84, e02516-17. https://doi.org/10.1128/AEM.02516-17

Rosa, M.D., Gambacorta, A., Minale, L., Bu'Lock, J.D., 1971. Bacterial Triterpenes. J. Chem. Soc. Chem. Commun. 619–620. https://doi.org/10.1039/C29710000619

Rosas-Ballina, M., Goldstein, R.S., Gallowitsch-Puerta, M., Yang, L., Valdés-Ferrer, S.I., Patel, N.B., Chavan, S., Al-Abed, Y., Yang, H., Tracey, K.J., 2009. The Selective A7 Agonist GTS-21 Attenuates Cytokine Production In Human Whole Blood And Human Monocytes Activated By Ligands For TLR2, TLR3, TLR4, TLR9, And RAGE. Mol. Med. 15, 195–202. https://doi.org/10.2119/molmed.2009.00039

Ross, A.C., Liu, H., Pattabiraman, V.R., Vederas, J.C., 2010. Synthesis Of The Lantibiotic Lactocin S Using Peptide Cyclizations On Solid Phase. J. Am. Chem. Soc. 132, 462–463. https://doi.org/10.1021/ja9095945

Rottig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., Kohlbacher, O., 2011. Nrpspredictor2--A Web Server For Predicting NRPS Adenylation Domain Specificity. Nucleic Acids Res. 39, W362–W367. https://doi.org/10.1093/nar/gkr323

Rueda, A., Zubía, E., Ortega, M.J., Carballo, J.L., Salvá, J., 1998. New Metabolites From The Sponge *Spongia agaricina*. J. Nat. Prod. 61, 258–261. https://doi.org/10.1021/np970390b

Ruiz, B., Chávez, A., Forero, A., García-Huante, Y., Romero, A., Sánchez, M., Rocha, D., Sánchez, B., Rodríguez-Sanoja, R., Sánchez, S., Langley, E., 2010. Production Of Microbial Secondary Metabolites: Regulation By The Carbon Source. Crit. Rev. Microbiol. 36, 146–167. https://doi.org/10.3109/10408410903489576

S. Rangnekar, S., Khan, T., 2015. Novel Anti-Inflammatory Drugs From Marine Microbes. Nat. Prod. J. 5, 206–218. https://doi.org/10.2174/2210315505666150827212323

Sabdono, A., Radjasa, O.K., 2011. Microbial Symbionts In Marine Sponges: Marine Natural Product Factory. J. Coast. Dev. 11, 57–61.

Sanchez, L.M., Wong, W.R., Riener, R.M., Schulze, C.J., Linington, R.G., 2012. Examining The Fish Microbiome: Vertebrate-Derived Bacteria As An Environmental Niche For The Discovery Of Unique Marine Natural Products. PLoS ONE 7, e35398. https://doi.org/10.1371/journal.pone.0035398

Santajit, S., Indrawattana, N., 2016. Mechanisms Of Anti-microbial Resistance In ESKAPE Pathogens. BioMed Res. Int. 2016, 2475067. https://doi.org/10.1155/2016/2475067

Sauberan, J.B., Bradley, J.S., 2018. 292 - Anti-microbial Agents, In: Long, S.S., Prober, C.G., Fischer, M. (Eds.), Principles And Practice Of Pediatric Infectious Diseases (Fifth Edition). Elsevier, pp. 1499-1531.e3. https://doi.org/10.1016/B978-0-323-40181-4.00292-9

Scherlach, K., Hertweck, C., 2009. Triggering Cryptic Natural Product Biosynthesis In Microorganisms. Org. Biomol. Chem. 7, 1753. https://doi.org/10.1039/b821578b

Scherlach, K., Hertweck, C., 2006. Discovery Of Aspoquinolones A-D, Prenylated Quinoline-2-One Alkaloids From *Aspergillus nidulans*, Motivated By Genome Mining. Org. Biomol. Chem. 4, 3517–3520. https://doi.org/10.1039/b607011f

Schinke, C., Martins, T., Queiroz, S.C.N., Melo, I.S., Reyes, F.G.R., 2017. Antibacterial Compounds From Marine Bacteria, 2010–2015. J. Nat. Prod. https://doi.org/10.1021/acs.jnatprod.6b00235

Schmidt-Dannert, C., 2015. Nextgen Microbial Natural Products Discovery. Microb. Biotechnol. 8, 26–28. https://doi.org/10.1111/1751-7915.12184

Schofield, M.M., Jain, S., Porat, D., Dick, G.J., Sherman, D.H., 2015. Identification And Analysis Of The Bacterial Endosymbiont Specialized For Production Of The Chemotherapeutic Natural Product ET-743. Environ. Microbiol. 17, 3964–3975. https://doi.org/10.1111/1462-2920.12908

Shelest, E., Heimerl, N., Fichtner, M., Sasso, S., 2015. Multimodular Type I Polyketide Synthases In Algae Evolve By Module Duplications And Displacement Of AT Domains In Trans. BMC Genomics 16. https://doi.org/10.1186/s12864-015-2222-9

Shen, X., Mustafa, M., Chen, Y., Cao, Y., Gao, J., 2019. Natural Thiopeptides As A Privileged Scaffold For Drug Discovery And Therapeutic Development. Med. Chem. Res. 28, 1063–1098. https://doi.org/10.1007/s00044-019-02361-1

Silver, L.L., 2016. A Gestalt Approach To Gram-Negative Entry. Bioorg. Med. Chem. 24, 6379–6389. https://doi.org/10.1016/j.bmc.2016.06.044

Singh, M.P., Mroczenski-Wildey, M.J., Steinberg, D.A., Andersen, R.J., Maiese, W.M., Greenstein, M., 1997. Biological Activity And Mechanistic Studies Of Andrimid. J. Antibiot. (Tokyo) 50, 270–273. https://doi.org/10.7164/antibiotics.50.270

Sipkema, D., Schippers, K., Maalcke, W.J., Yang, Y., Salim, S., Blanch, H.W., 2011. Multiple Approaches To Enhance The Cultivability Of Bacteria Associated With The Marine

Sponge *Haliclona* (*gellius*) sp. Appl. Environ. Microbiol. 77, 2130–2140. https://doi.org/10.1128/AEM.01203-10

Skariyachan, S., G. Rao, A., Patil, M.R., Saikia, B., Bharadwaj KN, V., Rao GS, J., 2014. Anti-microbial Potential Of Metabolites Extracted From Bacterial Symbionts Associated With Marine Sponges In Coastal Area Of Gulf Of Mannar Biosphere, India. Lett. Appl. Microbiol. 58, 231–241. https://doi.org/10.1111/lam.12178

Skinnider, M.A., Dejong, C.A., Rees, P.N., Johnston, C.W., Li, H., Webster, A.L.H., Wyatt, M.A., Magarvey, N.A., 2015. Genomes To Natural Products Prediction Informatics For Secondary Metabolomes (PRISM). Nucleic Acids Res. gkv1012. https://doi.org/10.1093/nar/gkv1012

Skinnider, M.A., Johnston, C.W., Gunabalasingam, M., Merwin, N.J., Kieliszek, A.M., MacLellan, R.J., Li, H., Ranieri, M.R.M., Webster, A.L.H., Cao, M.P.T., Pfeifle, A., Spencer, N., To, Q.H., Wallace, D.P., Dejong, C.A., Magarvey, N.A., 2020. Comprehensive Prediction Of Secondary Metabolite Structure And Biological Activity From Microbial Genome Sequences. Nat. Commun. 11, 6058. https://doi.org/10.1038/s41467-020-19986-1

Skinnider, M.A., Merwin, N.J., Johnston, C.W., Magarvey, N.A., 2017. PRISM 3: Expanded Prediction Of Natural Product Chemical Structures From Microbial Genomes. Nucleic Acids Res. 45, W49–W54. https://doi.org/10.1093/nar/gkx320

Song, C., Schmidt, R., de Jager, V., Krzyzanowska, D., Jongedijk, E., Cankar, K., Beekwilder, J., van Veen, A., de Boer, W., van Veen, J.A., Garbeva, P., 2015. Exploring The Genomic Traits Of Fungus-Feeding Bacterial Genus *Collimonas*. BMC Genomics 16, 1103. https://doi.org/10.1186/s12864-015-2289-3

Soysal, A., 2015. Prevention Of Invasive Fungal Infections In Immunocompromised Patients: The Role Of Delayed-Release Posaconazole. Infect. Drug Resist. 8, 321–331. https://doi.org/10.2147/IDR.S65592

Srilekha, V., Krishna, G., Seshasrinivas, V., Charya, M.A.S., 2017. Antibacterial And Anti-Inflammatory Activities Of Marine *Brevibacterium* sp. Res. Pharm. Sci. 12, 283–289. https://doi.org/10.4103/1735-5362.212045

Steinert, G., Whitfield, S., Taylor, M.W., Thoms, C., Schupp, P.J., 2014. Application Of Diffusion Growth Chambers For The Cultivation Of Marine Sponge-Associated Bacteria. Mar. Biotechnol. N. Y. N 16, 594–603. https://doi.org/10.1007/s10126-014-9575-y

Stepper, J., Shastri, S., Loo, T.S., Preston, J.C., Novak, P., Man, P., Moore, C.H., Havlíček, V., Patchett, M.L., Norris, G.E., 2011. Cysteine *S* -Glycosylation, A New Post-Translational Modification Found In Glycopeptide Bacteriocins. FEBS Lett. 585, 645–650. https://doi.org/10.1016/j.febslet.2011.01.023

Stirk, W.A., Schwalb, A.N., Light, M.E., Medková, J., Lenobel, R., Strnad, M., van Staden, J., 2003. Potential Medicinal Value Of Some South African Seaweeds. South Afr. J. Bot. 69, 462–468. https://doi.org/10.1016/S0254-6299(15)30282-9

Streit, W.R., Daniel, R., Jaeger, K.-E., 2004. Prospecting For Biocatalysts And Drugs In The Genomes Of Non-Cultured Microorganisms. Curr. Opin. Biotechnol. 15, 285–290. https://doi.org/10.1016/j.copbio.2004.05.006

Strieker, M., Tanović, A., Marahiel, M.A., 2010. Nonribosomal Peptide Synthetases: Structures And Dynamics. Curr. Opin. Struct. Biol. 20, 234–240. https://doi.org/10.1016/j.sbi.2010.01.009

Subramani, R., Aalbersberg, W., 2012. Marine Actinomycetes: An Ongoing Source Of Novel Bioactive Metabolites. Microbiol. Res. 167, 571–580. https://doi.org/10.1016/j.micres.2012.06.005

Sule, P., Belas, R., 2013. A Novel Inducer Of *Roseobacter* Motility Is Also A Disruptor Of Algal Symbiosis. J. Bacteriol. 195, 637–646. https://doi.org/10.1128/JB.01777-12

Sun, M.-K., Alkon, D.L., 2006. Bryostatin-1: Pharmacology And Therapeutic Potential As A CNS Drug. CNS Drug Rev. 12, 1–8. https://doi.org/10.1111/j.1527-3458.2006.00001.x

Svenson, J., 2013. MabCent: Arctic marine bioprospecting in Norway. Phytochem. Rev. 12, 567–578. https://doi.org/10.1007/s11101-012-9239-3

The Editors of Encyclopaedia Britannica, 2019. Penicillin | Discovery, Mechanism Of Action, & Uses [WWW Document]. Encycl. Br. URL https://www.britannica.com/science/penicillin (accessed 7.2.20).

Thomas, T.R.A., Kavlekar, D.P., LokaBharathi, P.A., 2010. Marine Drugs From Sponge-Microbe Association—A Review. Mar. Drugs 8, 1417–1468. https://doi.org/10.3390/md8041417

Thomson, N.R., Crow, M.A., McGowan, S.J., Cox, A., Salmond, G.P., 2000. Biosynthesis Of Carbapenem Antibiotic And Prodigiosin Pigment In Serratia Is Under Quorum Sensing control. Mol. Microbiol. 36, 539–556.

Timmermans, M.L., Paudel, Y.P., Ross, A.C., 2017. Investigating The Biosynthesis Of Natural Products From Marine Proteobacteria: A Survey Of Molecules And Strategies. Mar. Drugs 15. https://doi.org/10.3390/md15080235

Trauger, J.W., Walsh, C.T., 2000. Heterologous Expression In *Escherichia coli* of The First Module Of The Nonribosomal Peptide Synthetase For Chloroeremomycin, A Vancomycin-Type Glycopeptide Antibiotic. Proc. Natl. Acad. Sci. 97, 3112–3117.

Unsin, C.E.-M., Rajski, S.R., Shen, B., 2013. The Role Of Genetic Engineering In Natural Product-Based Anti-cancer Drug Discovery, In: Koehn, F.E. (Ed.), Natural Products and Cancer Drug Discovery. Springer New York, New York, NY, pp. 175–191.

van der Meij, A., Worsley, S.F., Hutchings, M.I., van Wezel, G.P., 2017. Chemical Ecology Of Antibiotic Production By Actinomycetes. FEMS Microbiol. Rev. 41, 392–416. https://doi.org/10.1093/femsre/fux005

van Heel, A.J., de Jong, A., Montalban-Lopez, M., Kok, J., Kuipers, O.P., 2013. BAGEL3: Automated Identification Of Genes Encoding Bacteriocins And (Non-)Bactericidal Posttranslationally Modified Peptides. Nucleic Acids Res. 41, W448–W453. https://doi.org/10.1093/nar/gkt391

van Zyl, L.J., Matobole, R., Augustin Nsole Biteghe, F., Klein, T., Kirby, B., Trindade, M., 2016. Draft Genome Sequences Of Three *Bacillus* Species From South African Marine Sponges. Genome Announc. 4. https://doi.org/10.1128/genomeA.00143-16

VanderMolen, K.M., Raja, H.A., El-Elimat, T., Oberlies, N.H., 2013. Evaluation Of Culture Media For The Production Of Secondary Metabolites In A Natural Products Screening Program. AMB Express 3, 1.

Vandormael, A., Oliveira, T. de, Tanser, F., Bärnighausen, T., Herbeck, J.T., 2018. High Percentage Of Undiagnosed HIV Cases Within A Hyperendemic South African Community: A Population-Based Study. J Epidemiol Community Health 72, 168–172. https://doi.org/10.1136/jech-2017-209713

Versluis, D., Nijsse, B., Naim, M.A., Koehorst, J.J., Wiese, J., Imhoff, J.F., Schaap, P.J., van Passel, M.W.J., Smidt, H., Sipkema, D., 2018. Comparative Genomics Highlights Symbiotic Capacities And High Metabolic Flexibility Of The Marine Genus *Pseudovibrio*. Genome Biol. Evol. 10, 125–142. https://doi.org/10.1093/gbe/evx271

Vizcaino, M., Isabel, 2011. Chemical Defense Of Pseudopterogorgia Americana: A Focus On The Anti-microbial Potential Of A *Pseudovibrio* sp :: MUSC Theses and Dissertations [WWW Document]. URL http://cdm-library-v.musc.edu/cdm/ref/collection/mtd/id/672 (accessed 11.22.17).

Vizcaino, M., Morris, P., Moeller, P., 2010. Evidence For The Production Of Novel Macrocyclic Antibiotics From The Marine Bacterium *Pseudovibrio denitrificans* [WWW Document]. URL https://scifinder-n.cas.org/searchDetail/reference/6215d48ef108960a72428dee/referenceDetails (accessed 10.20.19).

von Mulert, U., Luzhetskyy, A., Hofmann, C., Mayer, A., Bechthold, A., 2004. Expression Of The Landomycin Biosynthetic Gene Cluster In A PKS Mutant Of *Streptomyces Fradiae* Is Dependent On The Coexpression Of A Putative Transcriptional Activator Gene. FEMS Microbiol. Lett. 230, 91–97. https://doi.org/10.1016/S0378-1097(03)00861-9

Wagner, B.K., 2016. The Resurgence Of Phenotypic Screening In Drug Discovery And Development. Expert Opin. Drug Discov. 11, 121–125. https://doi.org/10.1517/17460441.2016.1122589

Walmsley, T.A., Matcher, G.F., Zhang, F., Hill, R.T., Davies-Coleman, M.T., Dorrington, R.A., 2012. Diversity Of Bacterial Communities Associated With The Indian Ocean Sponge *Tsitsikamma favus* that Contains The Bioactive Pyrroloiminoquinones, Tsitsikammamine

A And B. Mar. Biotechnol. N. Y. N 14, 681–691. https://doi.org/10.1007/s10126-012-9430-y

Walsh, C.T., 2004. Polyketide And Nonribosomal Peptide Antibiotics: Modularity And Versatility. Science 303, 1805–1810.

Wang, G., 2012. Post-Translational Modifications Of Natural Anti-microbial Peptides And Strategies For Peptide Engineering. Curr. Biotechnol. 1, 72.

Wang, H., 2014. Genome-Based Natural Product Biosynthetic Gene Cluster Discovery: From Sequencing To Mining. Univ. Hels. Fac. Agric. For. Dep. Food Environ. Sci. Div. Microbiol. Biotechnol. Doctoral dissertation (article-based).

Waterworth, S.C., Jiwaji, M., Kalinski, J.-C.J., Parker-Nance, S., Dorrington, R.A., 2017. A Place To Call Home: An Analysis Of The Bacterial Communities In Two *Tethya Rubra* Samaai And Gibbons 2005 Populations In Algoa Bay, South Africa. Mar. Drugs 15. https://doi.org/10.3390/md15040095

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., Lee, S.Y., Fischbach, M.A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M.H., 2015. Antismash 3.0—A Comprehensive Resource For The Genome Mining Of Biosynthetic Gene Clusters. Nucleic Acids Res. 43, W237–W243. https://doi.org/10.1093/nar/gkv437

Weber, T., Kim, H.U., 2016. The Secondary Metabolite Bioinformatics Portal: Computational Tools To Facilitate Synthetic Biology Of Secondary Metabolite Production. Synth. Syst. Biotechnol., Special Issue On "Bioinformatic Tools And Approaches For Synthetic Biology Of Natural Products" 1, 69–79. https://doi.org/10.1016/j.synbio.2015.12.002

Weber, T., Rausch, C., Lopez, P., Hoof, I., Gaykova, V., Huson, D.H., Wohlleben, W., 2009. CLUSEAN: A Computer-Based Framework For The Automated Analysis Of Bacterial Secondary Metabolite Biosynthetic Gene Clusters. J. Biotechnol. 140, 13–17.

Wenzel, S., Muller, R., 2005. Recent Developments Towards The Heterologous Expression Of Complex Bacterial Natural Product Biosynthetic Pathways. Curr. Opin. Biotechnol. 16, 594–606. https://doi.org/10.1016/j.copbio.2005.10.001

Wenzel, S.C., Gross, F., Zhang, Y., Fu, J., Stewart, A.F., Müller, R., 2005. Heterologous Expression Of A Myxobacterial Natural Products Assembly Line In *Pseudomonads* Via red/ET Recombineering. Chem. Biol. 12, 349–356. https://doi.org/10.1016/j.chembiol.2004.12.012

WHO, 2020a. Anti-microbial Resistance [WWW Document]. Antimicrob. Resist. URL https://www.who.int/news-room/fact-sheets/detail/anti-microbial-resistance (accessed 6.19.21).

WHO, 2020b. The Top 10 Causes Of Death [WWW Document]. URL https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed 3.2.22).

WHO, 2018. GLASS | Global Anti-microbial Resistance Surveillance System (GLASS) Report [WWW Document]. WHO. URL http://www.who.int/glass/resources/publications/early-implementation-report/en/ (accessed 3.4.19).

WHO, 2017. WHO Publishes List Of Bacteria For Which New Antibiotics Are Urgently Needed [WWW Document]. URL https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed (accessed 6.20.21).

Wietz, M., Månsson, M., Gram, L., 2011. Chitin Stimulates Production Of The Antibiotic Andrimid In A *Vibrio coralliilyticus* strain. Environ. Microbiol. Rep. 3, 559–564. https://doi.org/10.1111/j.1758-2229.2011.00259.x

Wilkinson, B., Micklefield, J., 2007. Mining And Engineering Natural-Product Biosynthetic Pathways. Nat. Chem. Biol. 3, 379–386. https://doi.org/10.1038/nchembio.2007.7

Wink, M., 2015. Modes Of Action Of Herbal Medicines And Plant Secondary Metabolites. Medicines 2, 251–286. https://doi.org/10.3390/medicines2030251

Winkelmann, G., Schmid, D.G., Nicholson, G., Jung, G., Colquhoun, D.J., 2002. Bisucaberin--A Dihydroxamate Siderophore Isolated From *Vibrio salmonicida*, An Important Pathogen Of Farmed Atlantic Salmon (Salmo Salar). Biometals Int. J. Role Met. Ions Biol. Biochem. Med. 15, 153–160. https://doi.org/10.1023/a:1015206419613

Winn, M., Fyans, J.K., Zhuo, Y., Micklefield, J., 2016. Recent Advances In Engineering Nonribosomal Peptide Assembly Lines. Nat. Prod. Rep. 33, 317–347. https://doi.org/10.1039/C5NP00099H

Wittgens, A., Santiago-Schuebel, B., Henkel, M., Tiso, T., Blank, L.M., Hausmann, R., Hofmann, D., Wilhelm, S., Jaeger, K.-E., Rosenau, F., 2018. Heterologous Production Of Long-Chain Rhamnolipids From *Burkholderia glumae* In *Pseudomonas putida*-A Step Forward To Tailor-Made Rhamnolipids. Appl. Microbiol. Biotechnol. 102, 1229–1239. https://doi.org/10.1007/s00253-017-8702-x

Wolf, T., Shelest, V., Nath, N., Shelest, E., 2016. CASSIS And SMIPS: Promoter-Based Prediction Of Secondary Metabolite Gene Clusters In Eukaryotic Genomes. Bioinforma. Oxf. Engl. 32, 1138–1143. https://doi.org/10.1093/bioinformatics/btv713

Xi, L., Ruan, J., Huang, Y., 2012. Diversity And Biosynthetic Potential Of Culturable Actinomycetes Associated With Marine Sponges In The China Seas. Int. J. Mol. Sci. 13, 5917–5932. https://doi.org/10.3390/ijms13055917

Xu, M., Wang, Y., Zhao, Z., Gao, G., Huang, S.-X., Kang, Q., He, X., Lin, S., Pang, X., Deng, Z., Tao, M., 2016. Functional Genome Mining For Metabolites Encoded By Large Gene Clusters Using Heterologous Expression Of A Whole Genomic BAC Library In *Streptomyces*. Appl. Environ. Microbiol. AEM.01383-16. https://doi.org/10.1128/AEM.01383-16

Yamada, Y., Cane, D.E., Ikeda, H., 2012. Chapter Seven - Diversity And Analysis Of Bacterial Terpene Synthases, In: Hopwood, D.A. (Ed.), Methods In Enzymology, Natural Product Biosynthesis By Microorganisms And Plants, Part A. Academic Press, pp. 123–162.

Yamada, Y., Kuzuyama, T., Komatsu, M., Shin-ya, K., Omura, S., Cane, D.E., Ikeda, H., 2015. Terpene Synthases Are Widely Distributed In Bacteria. Proc. Natl. Acad. Sci. 112, 857–862. https://doi.org/10.1073/pnas.1422108112

Yamamoto, S., Fujita, Y., Okujo, N., Minami, C., Matsuura, S., Shinoda, S., 1992. Isolation And Partial Characterization Of A Compound With Siderophore Activity From *Vibrio parahaemolyticus*. FEMS Microbiol. Lett. 94, 181–186. https://doi.org/10.1111/j.1574-6968.1992.tb05311.x

Yamamoto, S., Okujo, N., Yoshida, T., Matsuura, S., Shinoda, S., 1994. Structure And Iron Transport Activity Of Vibrioferrin, A New Siderophore Of *Vibrio parahaemolyticus*. J. Biochem. (Tokyo) 115, 868–874. https://doi.org/10.1093/oxfordjournals.jbchem.a124432

Yamanaka, K., Reynolds, K.A., Kersten, R.D., Ryan, K.S., Gonzalez, D.J., Nizet, V., Dorrestein, P.C., Moore, B.S., 2014. Direct Cloning And Refactoring Of A Silent Lipopeptide

Biosynthetic Gene Cluster Yields The Antibiotic Taromycin A. Proc. Natl. Acad. Sci. 111, 1957–1962. https://doi.org/10.1073/pnas.1319584111

Yang, Y., Zhou, M., Hardwidge, P.R., Cui, H., Zhu, G., 2018. Isolation And Characterization Of N-Acyl Homoserine Lactone-Producing Bacteria From Cattle Rumen And Swine Intestines. Front. Cell. Infect. Microbiol. 8, 155. https://doi.org/10.3389/fcimb.2018.00155

Yu, J., Han, H., Zhang, X., Ma, C., Sun, C., Che, Q., Gu, Q., Zhu, T., Zhang, G., Li, D., 2019. Discovery Of Two New Sorbicillinoids By Overexpression Of The Global Regulator Laea In A Marine-Derived Fungus *Penicillium dipodomyis* YJ-11. Mar. Drugs 17. https://doi.org/10.3390/md17080446

Zani, C.L., Carroll, A.R., 2017. Database For Rapid Dereplication Of Known Natural Products Using Data From MS And Fast NMR Experiments. J. Nat. Prod. 80, 1758–1766. https://doi.org/10.1021/acs.jnatprod.6b01093

Zarins-Tutt, J.S., Barberi, T.T., Gao, H., Mearns-Spragg, A., Zhang, L., Newman, D.J., Goss, R.J.M., 2016. Prospecting For New Bacterial Metabolites: A Glossary Of Approaches For Inducing, Activating And Upregulating The Biosynthesis Of Bacterial Cryptic Or Silent Natural Products. Nat. Prod. Rep. 33, 54–72. https://doi.org/10.1039/C5NP00111K

Zerrifi, S.E.A., El Khalloufi, F., Oudra, B., Vasconcelos, V., 2018. Seaweed Bioactive Compounds Against Pathogens And Microalgae: Potential Uses On Pharmacology And Harmful Algae Bloom Control. Mar. Drugs 16, 55. https://doi.org/10.3390/md16020055

Zgurskaya, H.I., Löpez, C.A., Gnanakaran, S., 2015. Permeability Barrier Of Gram-Negative Cell Envelopes And Approaches To Bypass It. ACS Infect. Dis. 1, 512–522. https://doi.org/10.1021/acsinfecdis.5b00097

Zhang, W., Li, Z., Miao, X., Zhang, F., 2009. The Screening Of Anti-microbial Bacteria With Diverse Novel Nonribosomal Peptide Synthetase (NRPS) Genes From South China Sea Sponges. Mar. Biotechnol. 11, 346–355. https://doi.org/10.1007/s10126-008-9148-z

Zhao, L.-X., Huang, S.-X., Tang, S.-K., Jiang, C.-L., Duan, Y., Beutler, J.A., Henrich, C.J., McMahon, J.B., Schmid, T., Blees, J.S., Colburn, N.H., Rajski, S.R., Shen, B., 2011. Actinopolysporins A–C And Tubercidin As A Pdcd4 Stabilizer From The Halophilic Actinomycete *Actinopolyspora erythraea* YIM 90600. J. Nat. Prod. 74, 1990–1995. https://doi.org/10.1021/np200603g

Zhao, X.-Q., 2011. Genome-Based Studies Of Marine Microorganisms To Maximize The Diversity Of Natural Products Discovery For Medical Treatments. Evid. Based Complement. Alternat. Med. 2011, 1–11. https://doi.org/10.1155/2011/384572

Zhu, S., Hegemann, J.D., Fage, C.D., Zimmermann, M., Xie, X., Linne, U., Marahiel, M.A., 2016. Insights Into The Unique Phosphorylation Of The Lasso Peptide Paeninodin. J. Biol. Chem. 291, 13662–13678. https://doi.org/10.1074/jbc.M116.722108

Ziemert, N., Podell, S., Penn, K., Badger, J.H., Allen, E., Jensen, P.R., 2012. The Natural Product Domain Seeker Napdos: A Phylogeny Based Bioinformatic Tool To Classify Secondary Metabolite Gene Diversity. PLoS ONE 7, e34064. https://doi.org/10.1371/journal.pone.0034064

Zink, D., Dufresne, C., Liesch, J., Martin, J., 2005. Identification/Dereplication Of Natural Products By LC-UV-MS, In: Proceedings Of The Spectral Search Parameters Presented At The Small Molecule Science Conference (COSMOS), Bristol, RI, USA. pp. 8–11.

Zivanovic, A., 2012. Marine Natural Products Isolation, Screening And Analogue Synthesis. Univ. Wollongong Thesis Collect. 1954-2016.

Zuppa, A., Costantini, S., Costantini, M., 2014. Comparative Sequence Analysis Of Bacterial Symbionts From The Marine Sponges *Geodia cydonium* and *Ircinia muscarum*. Bioinformation 10, 196.

7. Appendixes

7.1. Appendix I: media

List of media used for growth, fermentation, and bioassay during the study. Agar was only added for solid media and the pH was adjusted using diluted hydrochloric acid and 4M sodium hydroxide (NaOH).

Table 1: List of media used for both bacteria maintenance, experiment, and fermentation.

Media	Components	Amount (g / L) in Water
Luria broth / Agar (LB / A)	Tryptone	10
	Yeast extract	5
	Sodium chloride	10
	Agar	15
Soft agar		
	Tryptone	10
	Yeast extract	5
	Sodium chloride	10
	Agar	7.5
Glucose yeast and malt extract (GYM) –pH 7.2		
	Yeast extract	4
	Malt extract	10
	Glucose	4
	Calcium carbonate	2
	Magnesium chloride	5.3
	Potassium chloride	0.7
	Calcium chloride	0.1
	Sodium chloride	24
	Agar (optional)	15
Tryptic soya broth (TSB) – pH		
	Tryptic soya broth	3
	Sodium chloride	18

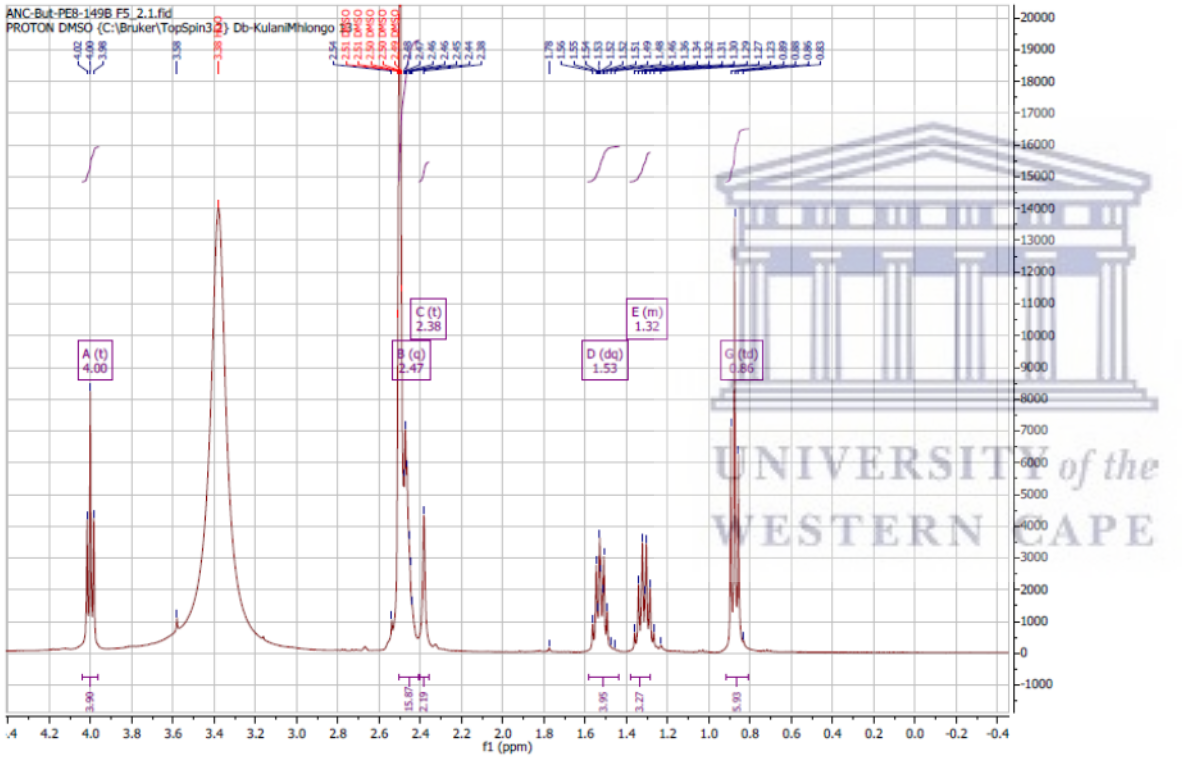
	Magnesium chloride	2
	Calcium chloride	0.075
	Agar (optional)	15
Activated charcoal medium (ACM) – pH 7.00		
	Sodium pyruvate	3
	Yeast extract	0.1
	Soyabean peptone	3
	Sodium nitrate	0.3
	Monopotassium phosphate	0.1
	Magnesium Sulfate Heptahydrate	0.15
	Activated charcoal	3
	Hepes	2.38
	Agar (optional)	
Zobell broth (ZBB) –pH 7.5		
	Peptone	3.75
	Yeast extract	1.25
	Sodium chloride	18
	Magnesium chloride	2 g,
	Potassium chloride	0.525
	Calcium chloride	0.075 g
	Agar (optional)	

7.2. Appendix II: **List of antibiotics to which *E. coli* 1699 is resistant and their Minimum inhibitory concentrations.**

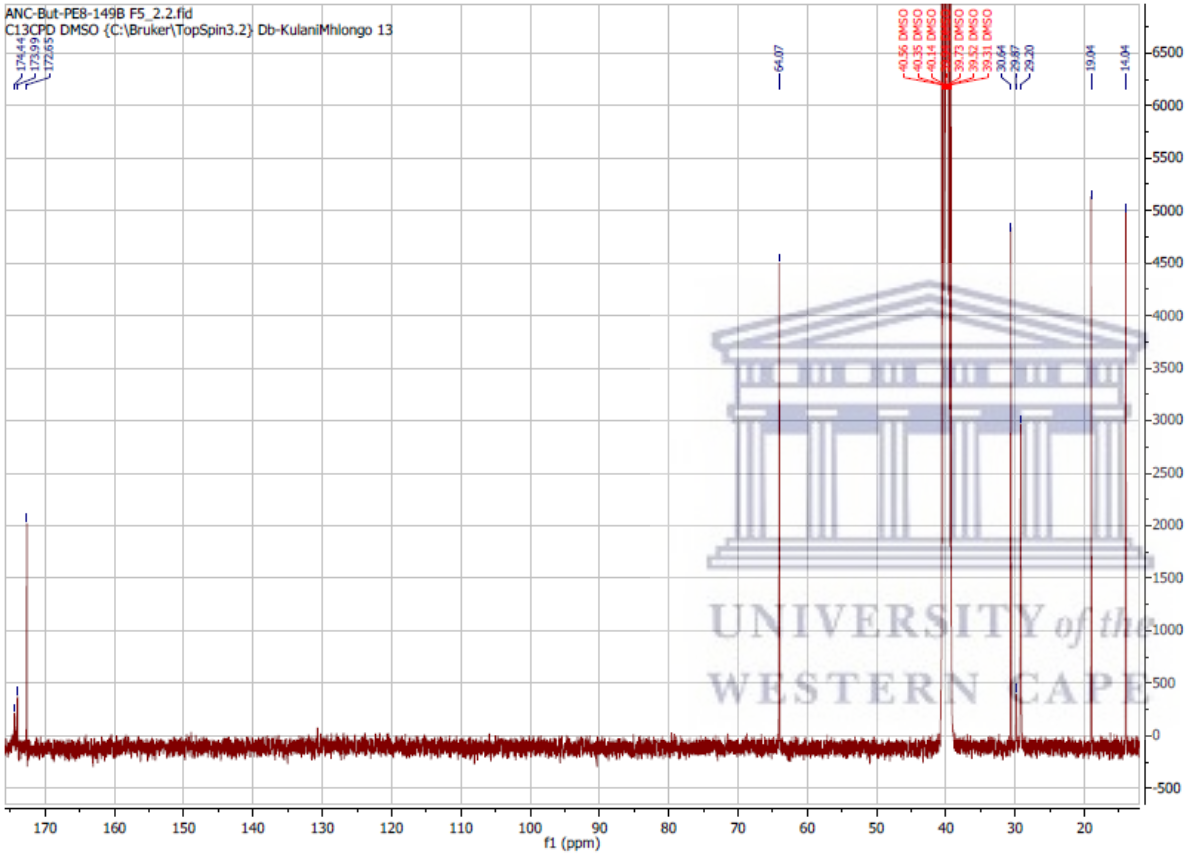
Compound	MIC (µg / mL)	Target	Antibiotic classification
A54145CB-181234	> 512	membrane	-
Calcimycin (A23187)	64	membrane	ionophore
Daptomycin	> 512	membrane	lipopeptide
Gramicidin	128	membrane	polypeptide
Polymyxin B	1	membrane	polypeptide (cationic)
Ampicillin	> 256	cell wall	aminopenicillin
Aztreonam	≤ 0.03	cell wall	monocyclic beta-lactam
Cephalosporin C	64	cell wall	cephalosporin
Penicillin G	> 256	cell wall	beta-lactam
Ristocetin	> 512	cell wall	aminoglycoside
Teicoplanin	> 512	cell wall	glycopeptide
Vancomycin	512	cell wall	glycopeptide
Aclacinomycin A	> 512	DNA interaction	anthracycline
Actinomycin A	> 256	DNA interaction	polypeptide (toxic)
Actinomycin D	256	DNA interaction	polypeptide (toxic)
Bleomycin A2	> 64	DNA interaction	glycopeptide
Coumermycin A1	64	DNA interaction	aminocoumarin
Daunorubicin	> 256	DNA interaction	anthracycline
Gliotoxin	32	DNA interaction	mycotoxin
Mitomycin C	1	DNA interaction	aziridine-containing
Streptonigrin	2	DNA interaction	quinone-containing
Streptozotocin	> 64	DNA interaction	glucosamine
Chromomycin A3	> 512	gyrase	glycoside
Nalidixic Acid	> 256	gyrase	naphthyridone
Novobicin	> 256	gyrase	aminocoumarin
Apramycin	64	protein synthesis	aminoglycoside
Dibekacin	> 256	protein synthesis	aminoglycoside
Gentamycin	128	protein synthesis	aminoglycoside
Kanamycin	> 256	protein synthesis	aminoglycoside
Kasugamycin	> 64	protein synthesis	aminoglycoside
Neomycin	16	protein synthesis	aminoglycoside
Netilmycin	128	protein synthesis	aminoglycoside
Streptomycin	> 256	protein synthesis	aminoglycoside
Streptothricin	256	protein synthesis	aminoglycoside
Tobramycin	> 256	protein synthesis	aminoglycoside
Puromycin	> 64	protein synthesis	aminonucleoside
Spectinomycin	> 256	protein synthesis	aminocyclitol
Chloramphenicol	-	protein synthesis	acetamide
Chlortetracycline	-	protein synthesis	tetracyclide
Erythromycin	64	protein synthesis	macrolide
Lincomycin	512	protein synthesis	lincosamide
Spiramycin	256	protein synthesis	macrolide
Tetracycline	> 256	protein synthesis	tetracycline
Thiostrepton	> 512	protein synthesis	oligopeptide
Tylosin	512	protein synthesis	macrolide
Virginiamycin	64	protein synthesis	streptogramin
Rifampin	> 256	RNA polymerase	rifamycin
Rifamycin SV	> 64	RNA polymerase	rifamycin
Albomycin	high	iron metabolism	cyclic polypeptide
Trimethoprim	> 400	DHFR	diaminopyrimidine

7.3. Appendix III: NMR spectra

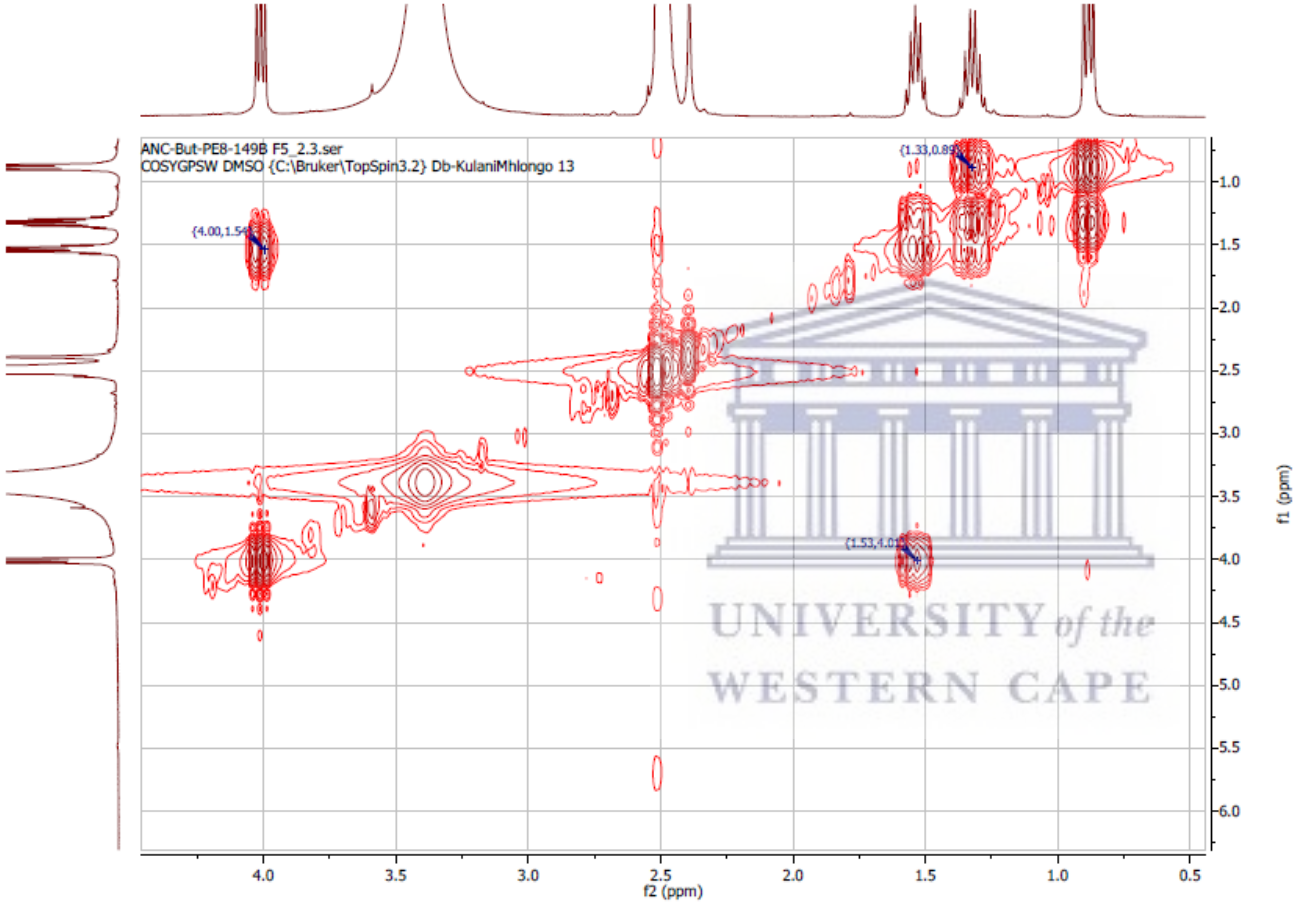
Algoa BW 1



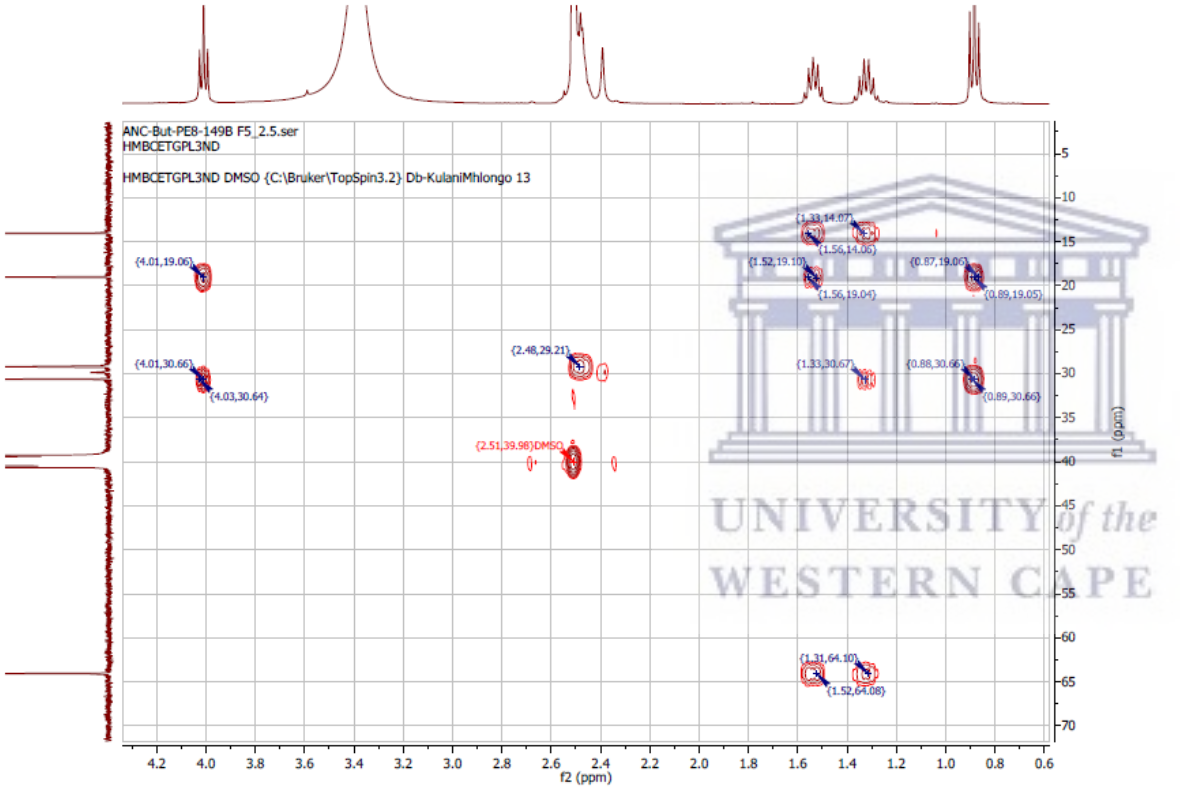
^{1}H NMR spectrum



^{13}C NMR spectrum

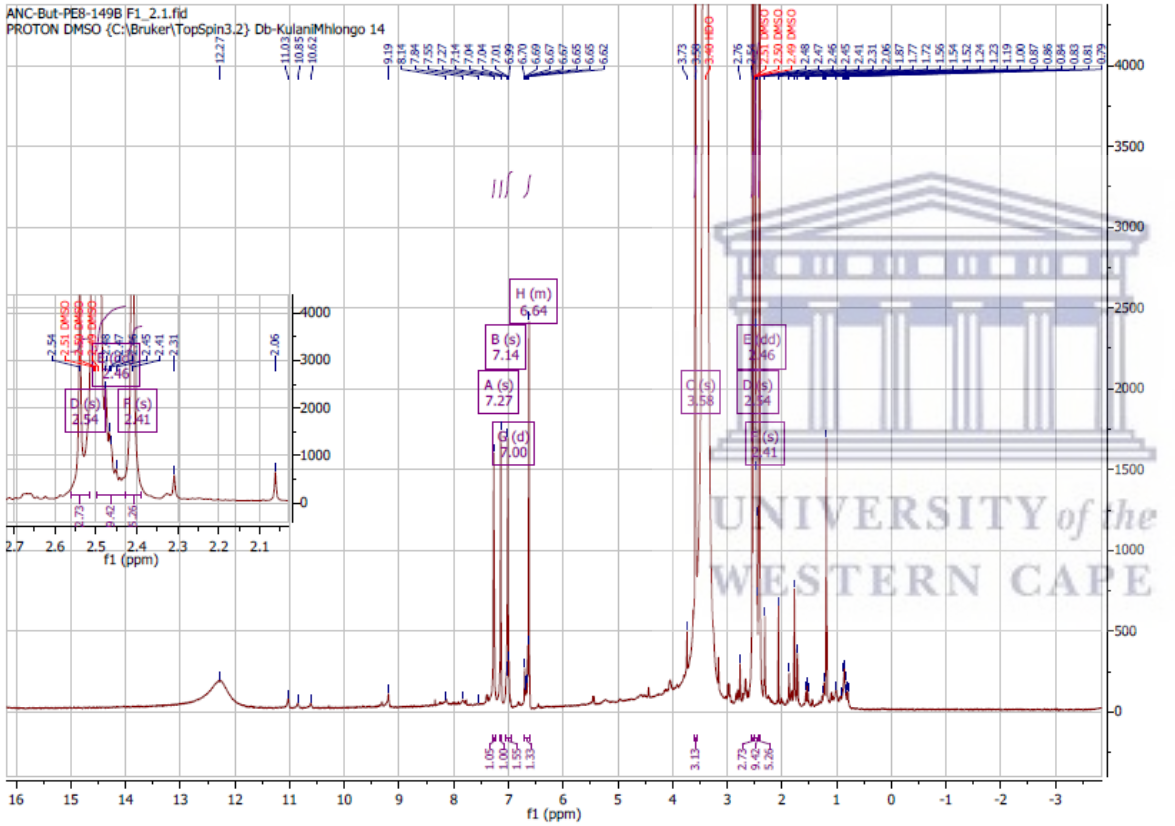


COSY NMR spectrum

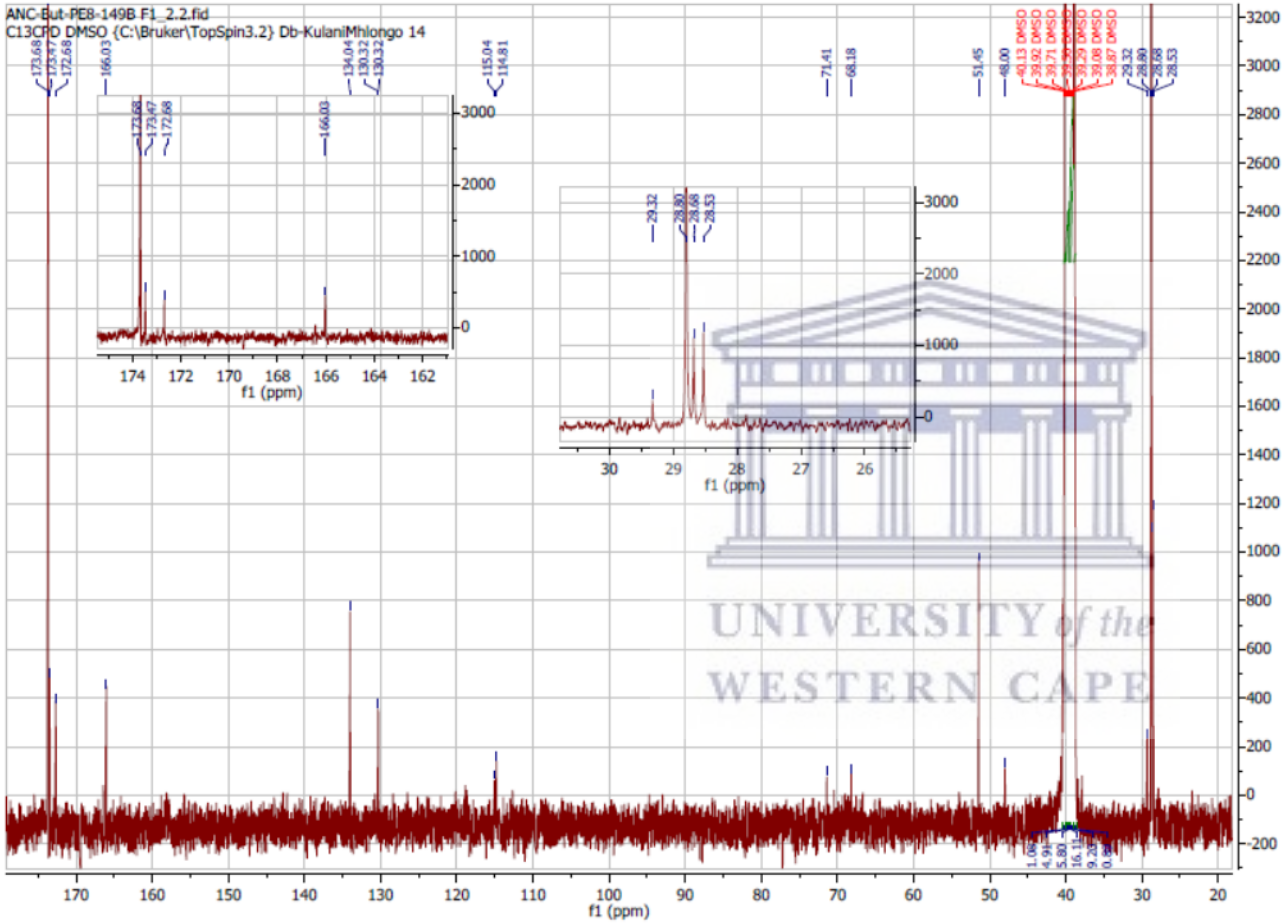


HMBC Spectrum

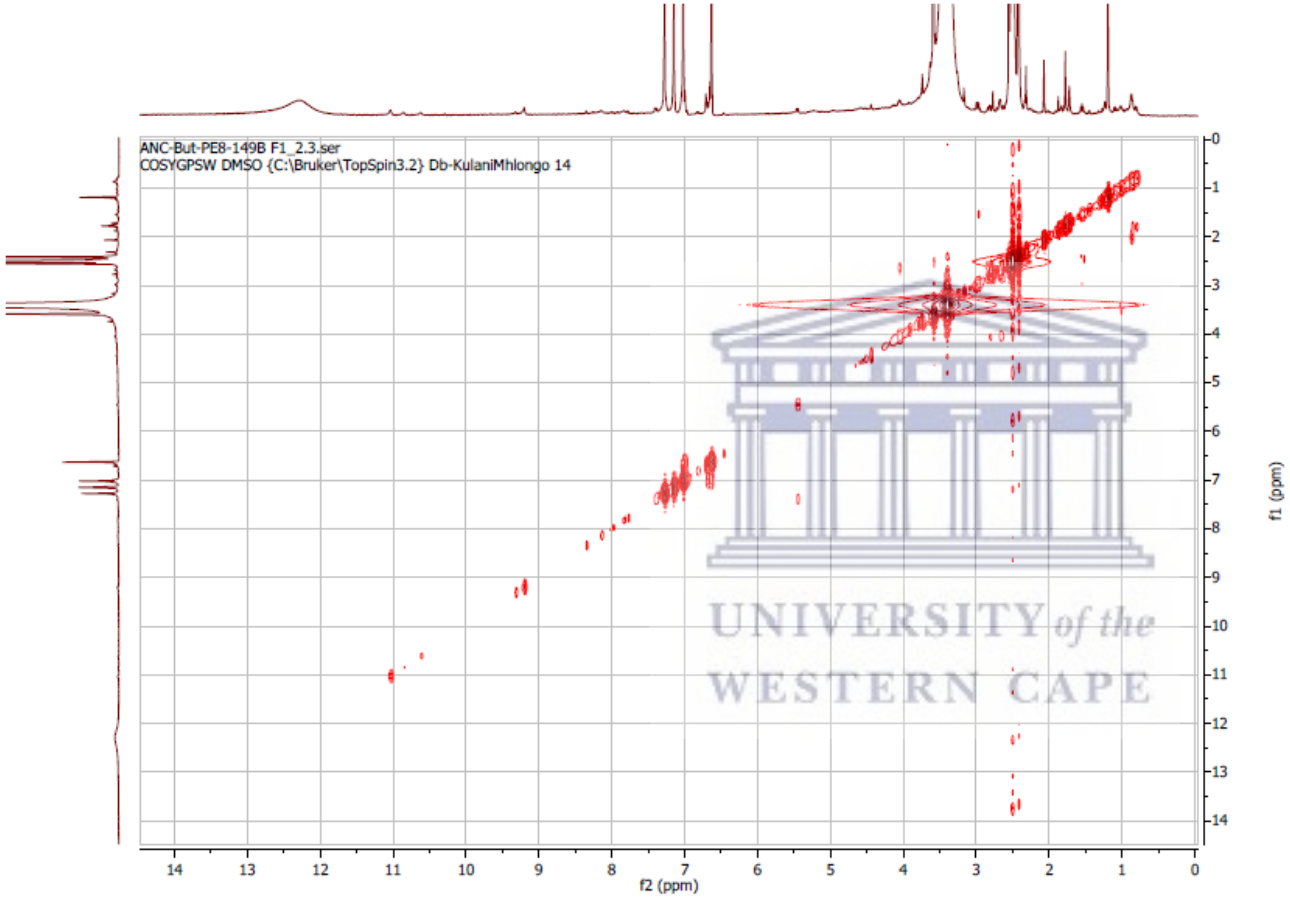
Algoa Bay BW 2



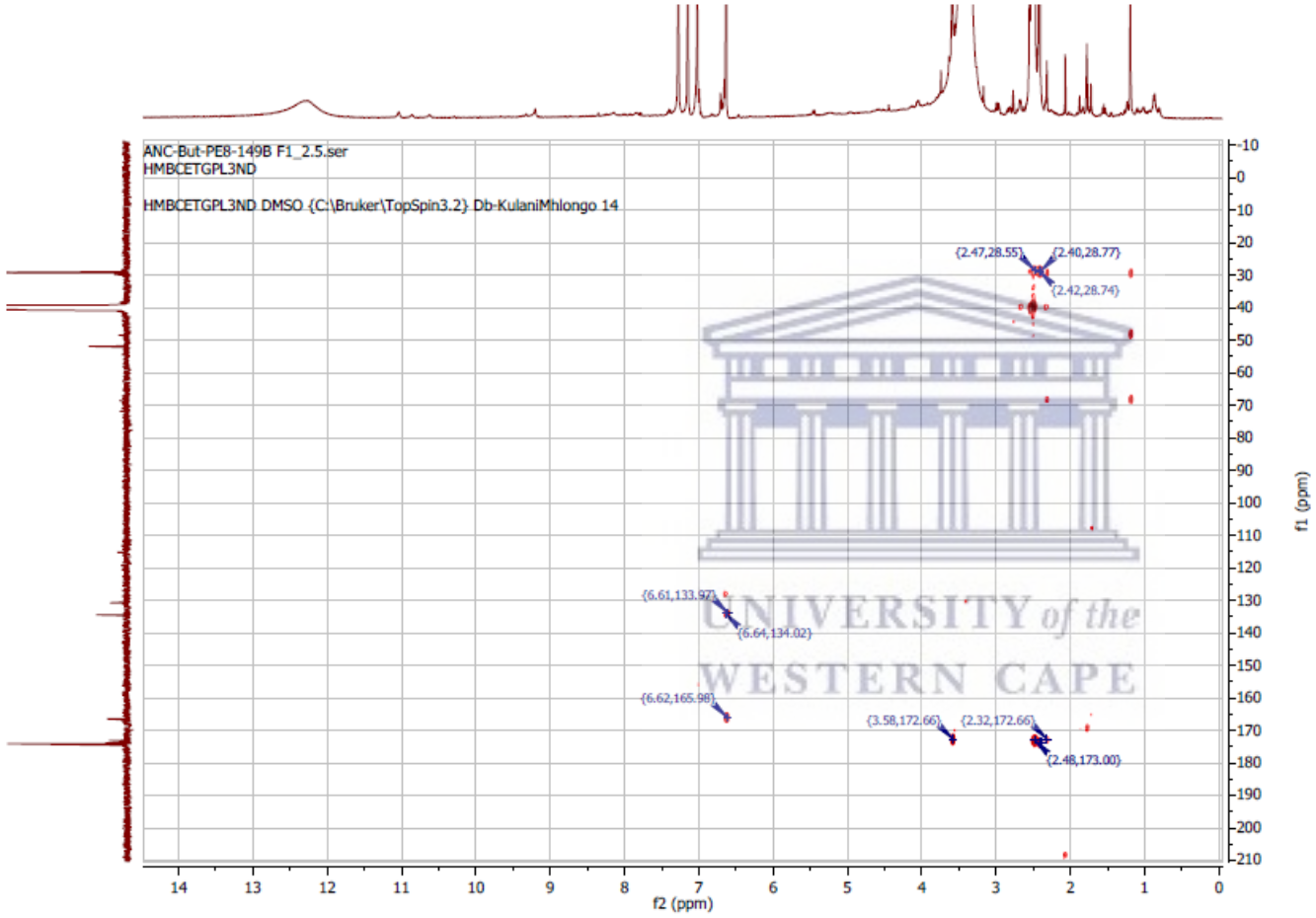
[1]H spectrum



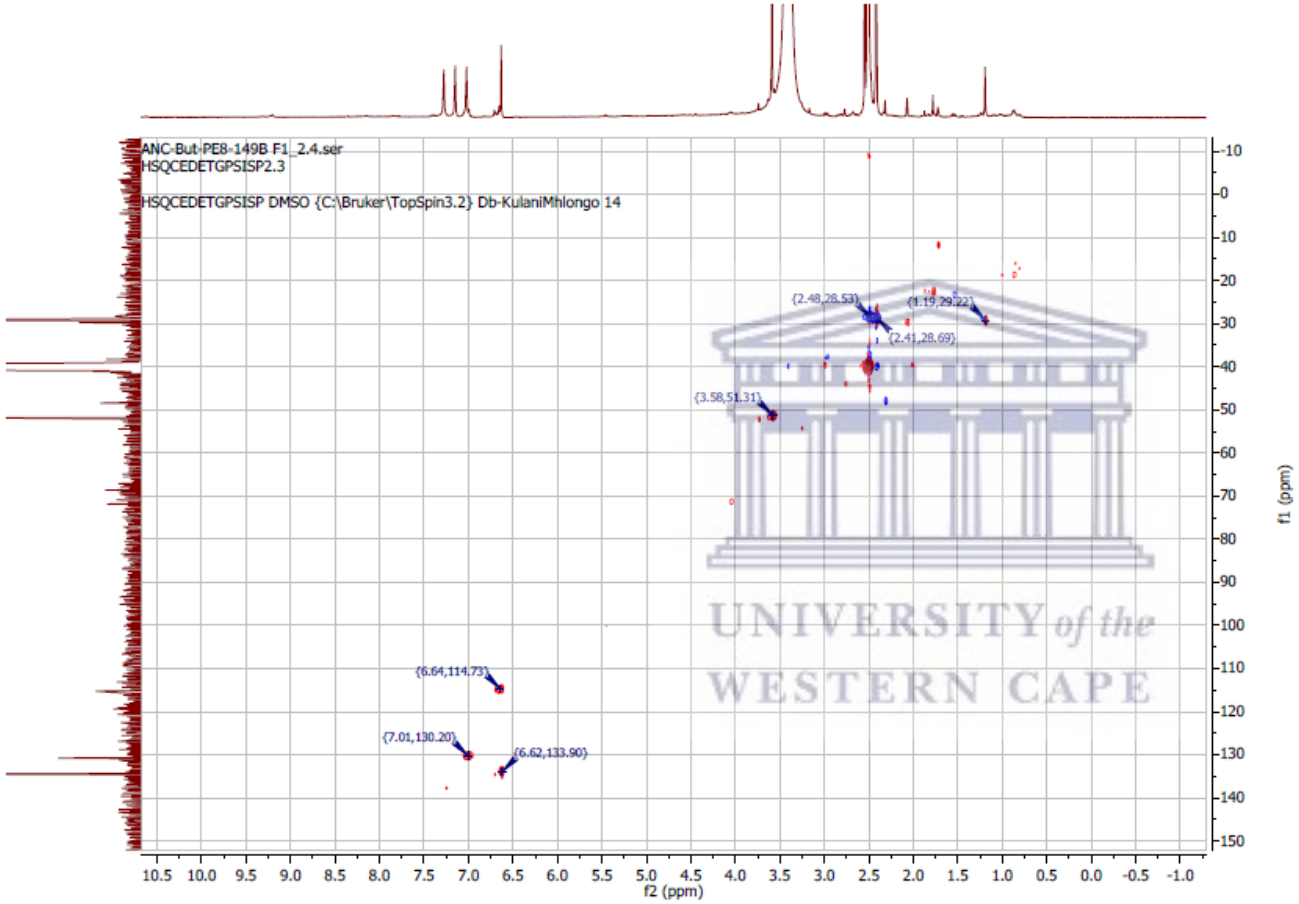
^{13}C spectrum



COSY spectrum

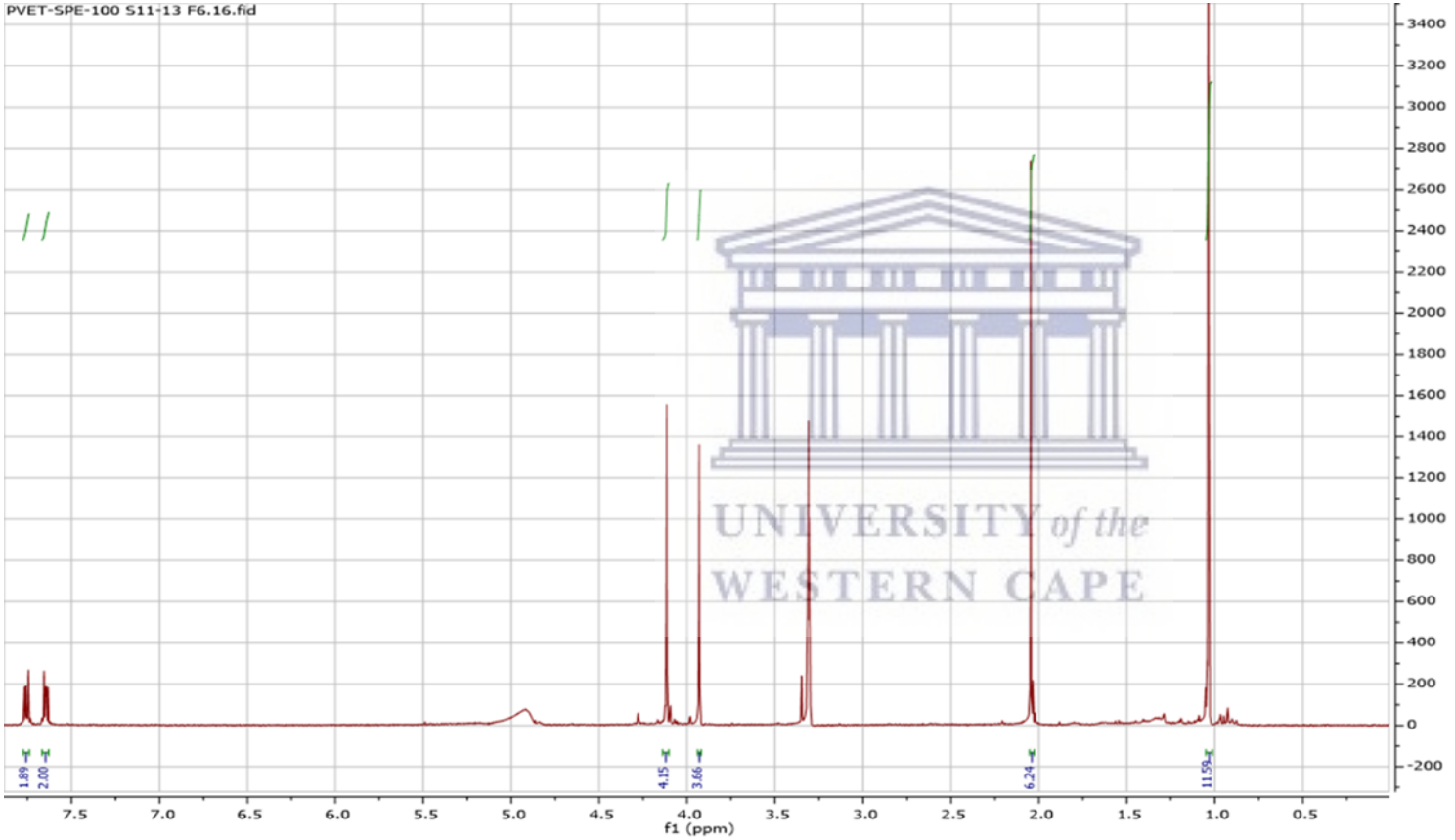


HMBC NMR spectrum

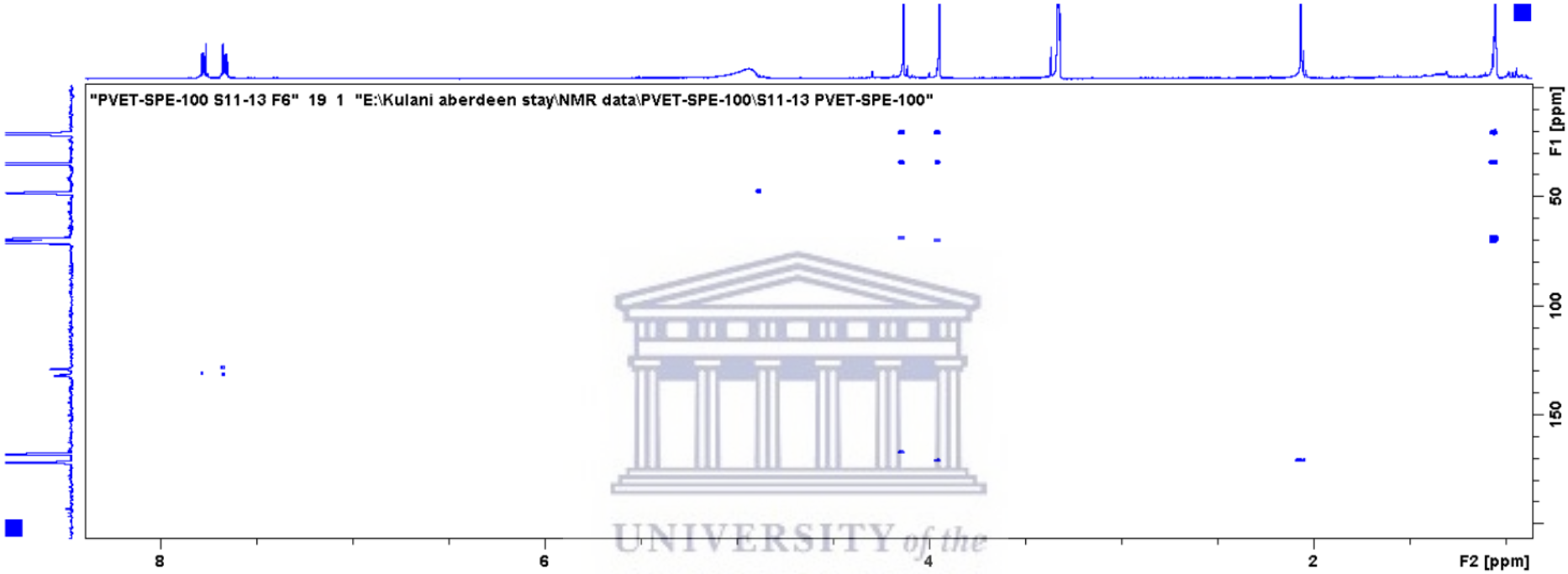


HSQC spectrum

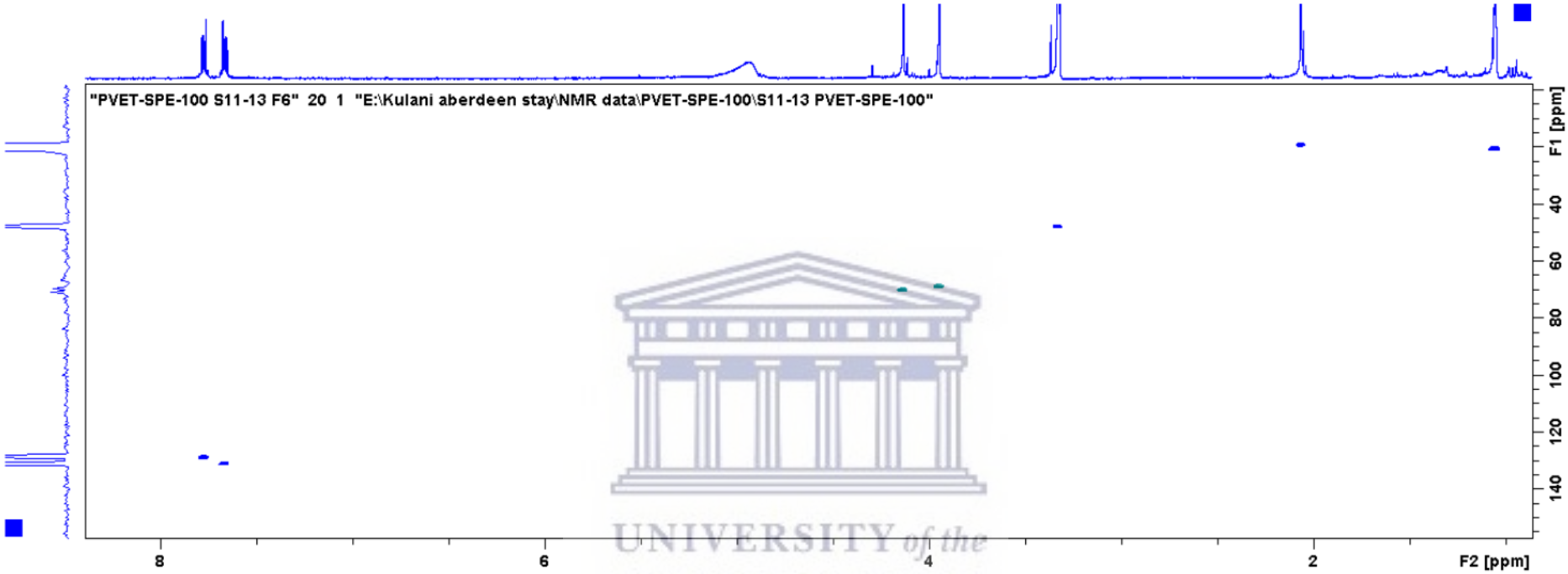
AlgoaPV1



^{1}H spectrum

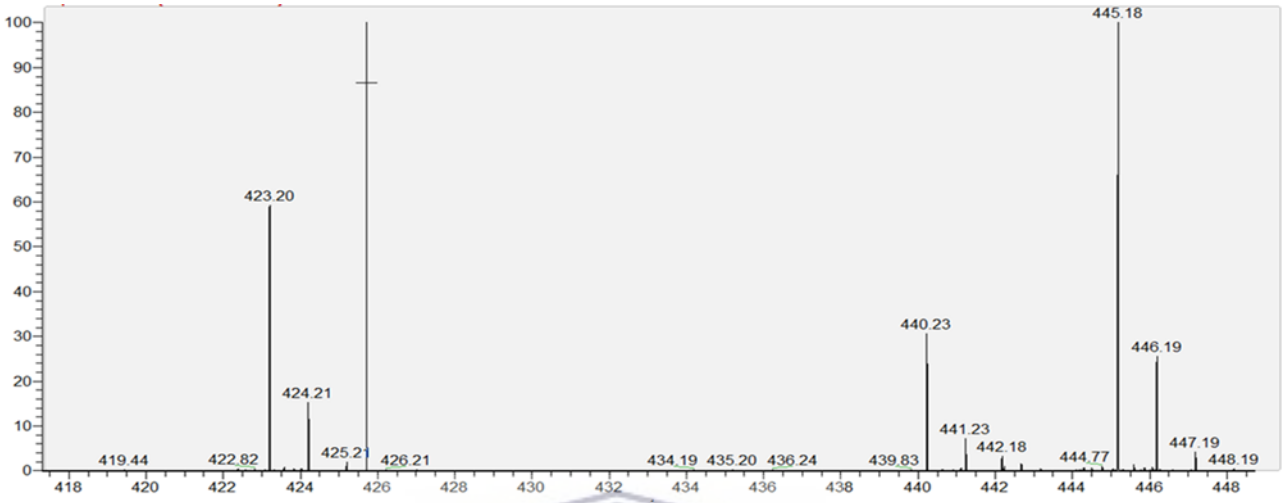


HMBC spectrum

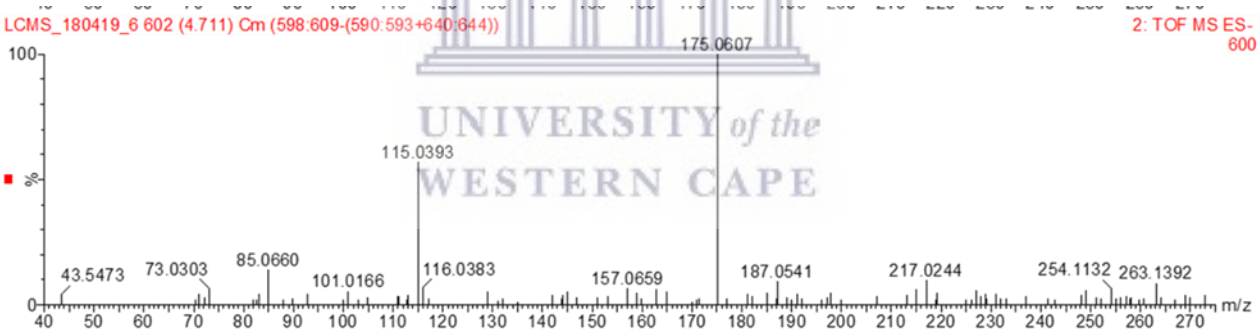


HSQC spectrum

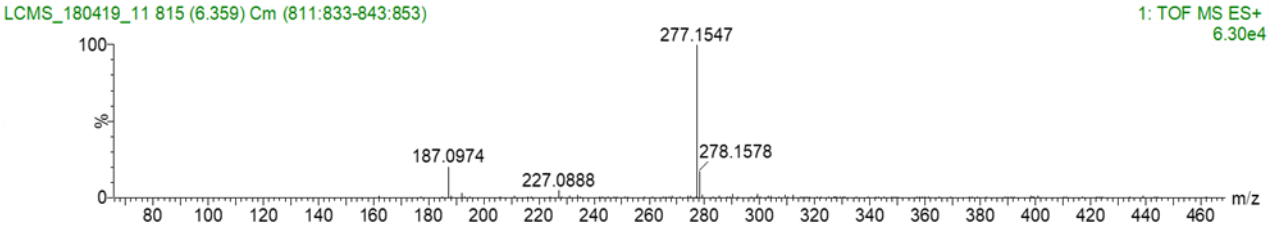
7.4. Appendix IV: LC-MS



Algoa PV1 LC-MS spectrum



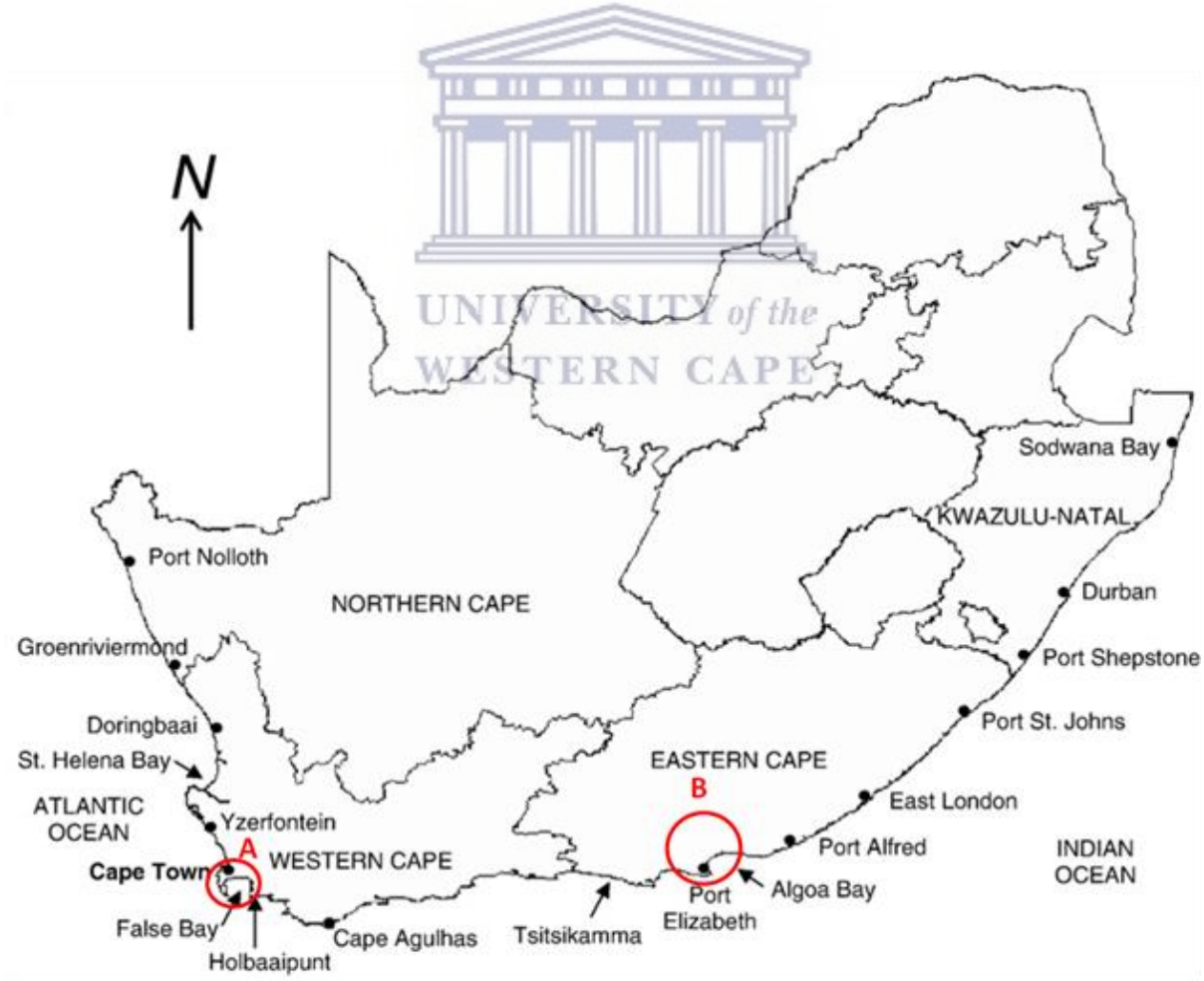
Algoa BW1 LC-MS spectrum



Algoa BW 2 LC-MS spectrum

7.5. Appendix V: Marine invertebrates collection and identification

The marine invertebrates from Algoa Bay, Port Elizabeth, South Africa (longitude 34°00.366S and latitude 25°43.209E) were collected by Dr Shirley Parker-Nance ((Just Blue, South Africa) on the 23rd of January and 21st of April 2013 with scuba diving at the depth of 23-25 m. The samples were stored and transported in sterile iced seawater. The samples were collected using SCUBA. The marine invertebrates from Kalka Bay were collected and stored at 4 for transportation to the laboratory for processing. Kalk Bay (coordinate: 34°07′40″S; 18°26′54″E), is situated along the False Bay, Cape Town, South Africa coastline with a diverse marine invertebrate. Kalk Bay invertebrates were hand collected from a tidal Pool. Taxonomic identification of the invertebrate specimens was undertaken by Dr Shirley Parker-Nance and the results are summarised in the Table below.



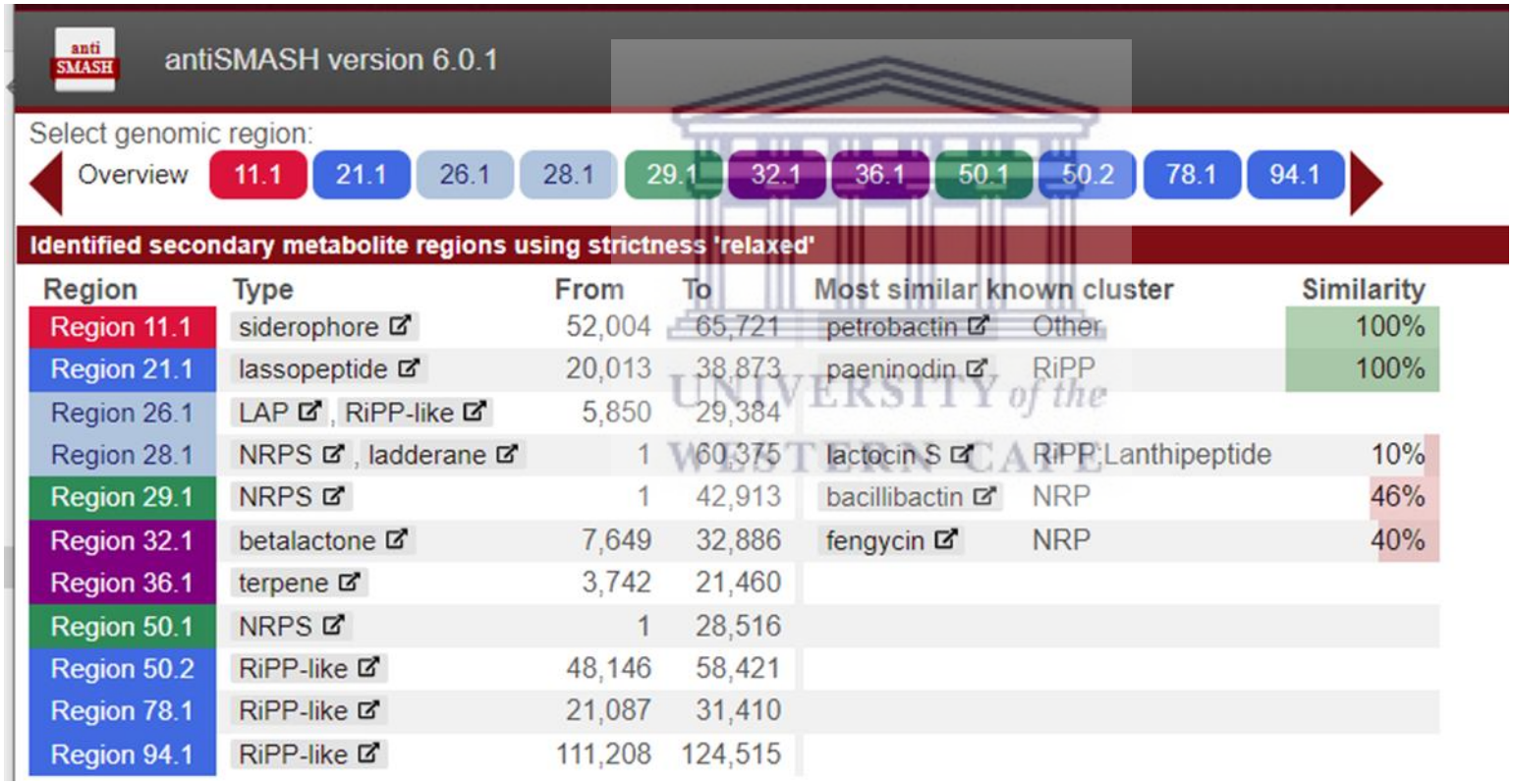
Locations of the sampling site on the map; A: Kalk Bay, Cape Town and B: Algoa Bay, Port Elizabeth. Adapted from, (Maneveldt et al., 2008)

Table 1. Sources of the marine bacteria used in the study.

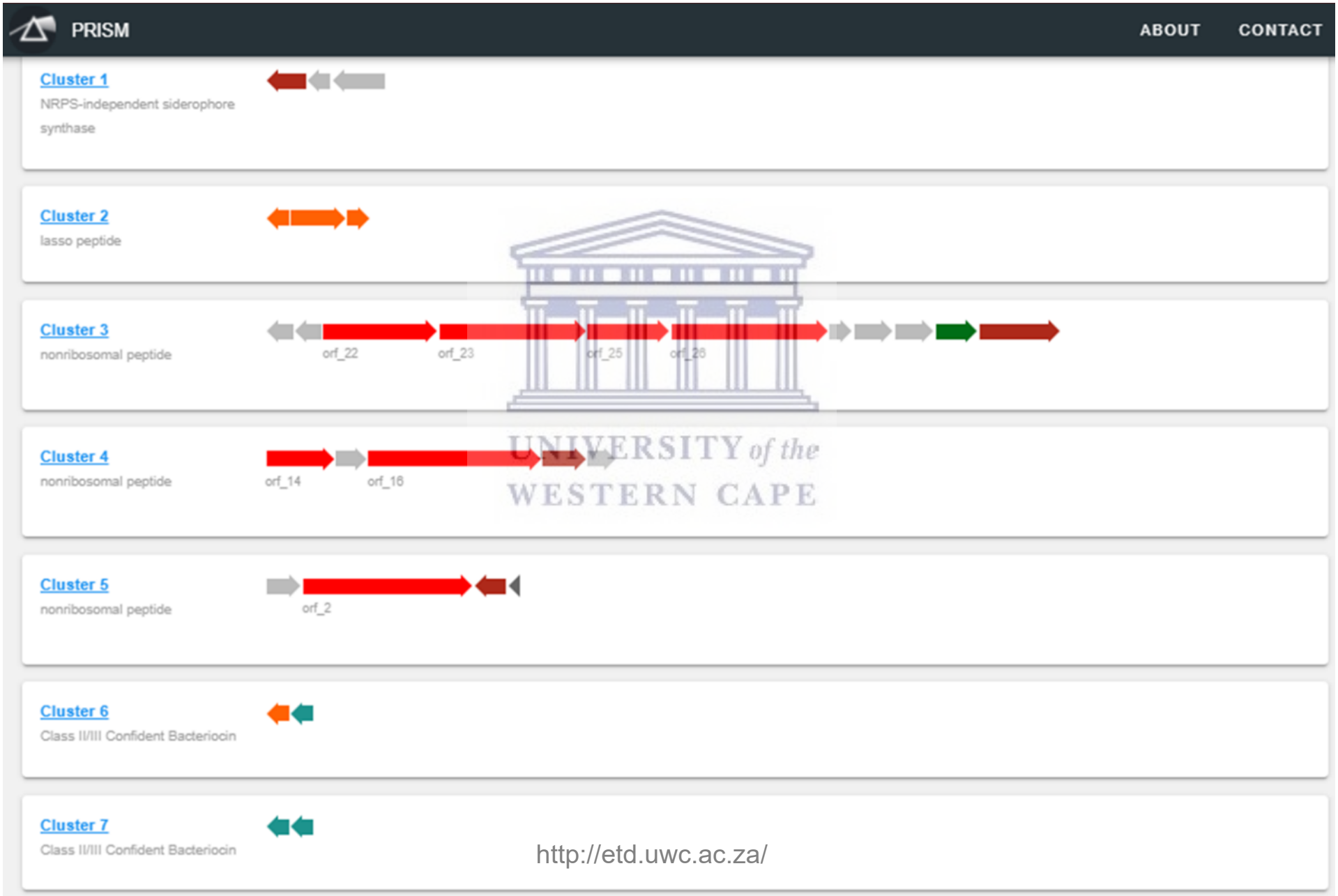
Sample ID	identification	Sample Date
PE06	*Distaplia skoogi Michaelsen*, 1924	23 January 2013
PE08	*Spongia (Spongia)* sp.001RSASPN	23 January 2013
PE12	*Eudistoma* sp. 011RSASPN	4 April 2013
PE13	*Pseudodistoma africanum* Millar, 1954	4 April 2013
PE14	*Hamacantha (Vomerula) esperioides* Ridley and Dendy, 1886	4 April 2013
KB07	*Halianthella annularis*	22-Novemeber-2012
KB08	*Leptophytum foveatum*	22-Novemeber-2012
KB11	*Hypnea spicifera*	22-Novemeber-2012
KB18	*Sabellastarte longa*	22-Novemeber-2012

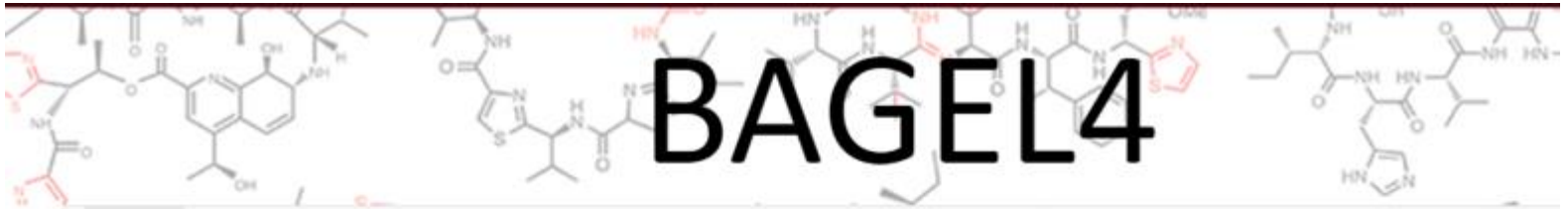
7.6. Appendix VI: Bioinformatic analysis output: PRISM and Begal screenshots

a. PE08-149B



Region	Type	From	To	Most similar known cluster		Similarity
Region 11.1	siderophore	52,004	65,721	petrobactin	Other	100%
Region 21.1	lassopeptide	20,013	38,873	paeninodin	RiPP	100%
Region 26.1	LAP, RiPP-like	5,850	29,384			
Region 28.1	NRPS, ladderane	1	60,375	lactocin S	RiPP:Lanthipeptide	10%
Region 29.1	NRPS	1	42,913	bacillibactin	NRP	46%
Region 32.1	betalactone	7,649	32,886	fengycin	NRP	40%
Region 36.1	terpene	3,742	21,460			
Region 50.1	NRPS	1	28,516			
Region 50.2	RiPP-like	48,146	58,421			
Region 78.1	RiPP-like	21,087	31,410			
Region 94.1	RiPP-like	111,208	124,515			





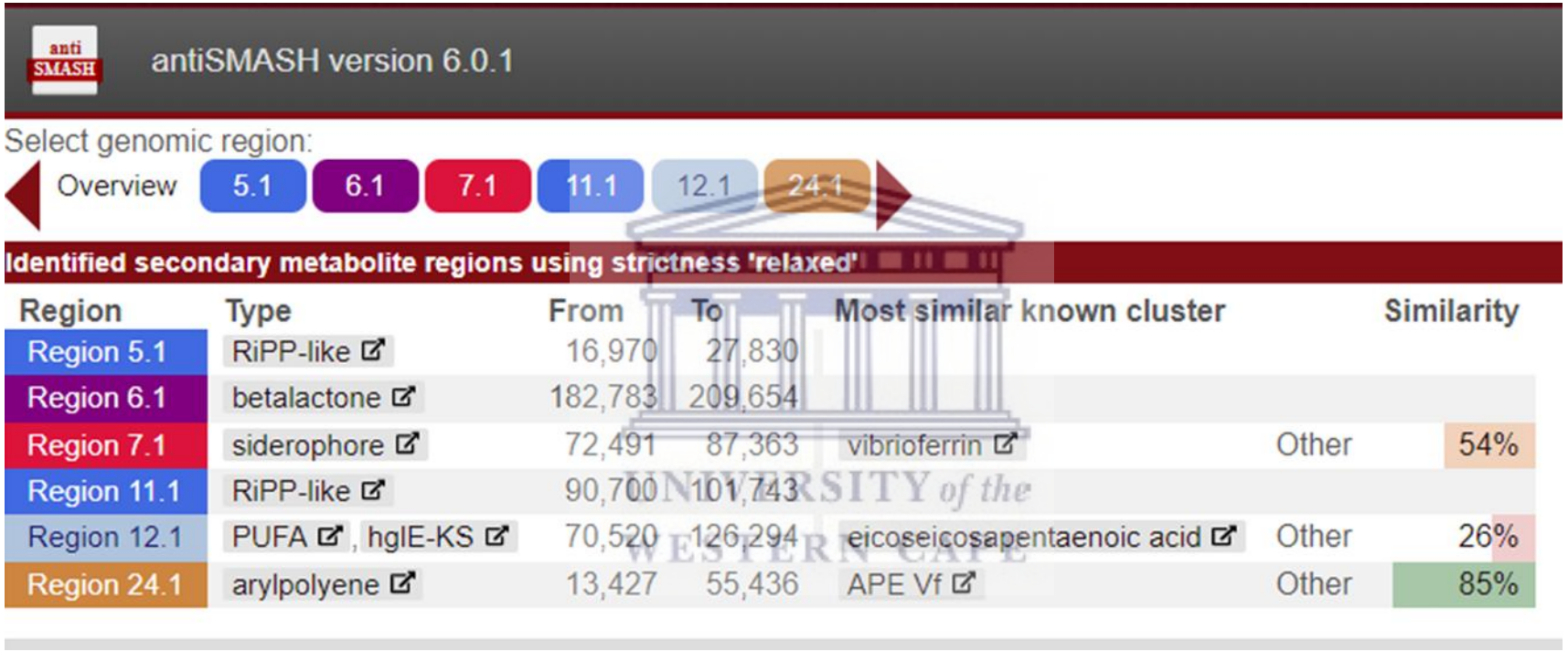
Home | Core Peptide Blast | Databases | Help | Links

Stop session and start a new run

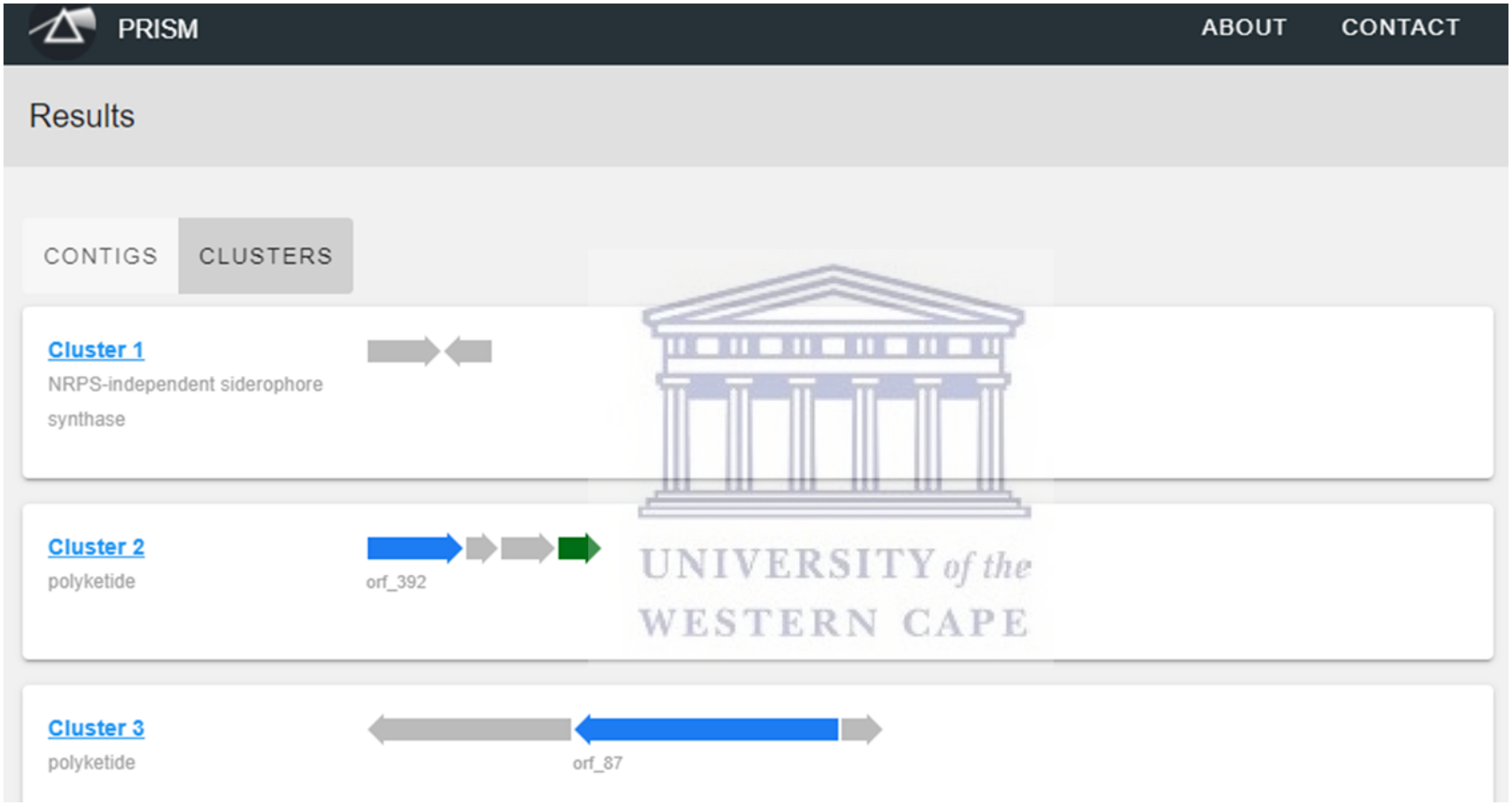
Run summary	
Number of files analyzed	2
Number of DNA fragments analyzed	113
Total bases in all DNA	5904342
Number of AOI's Areas Of Interest)	3

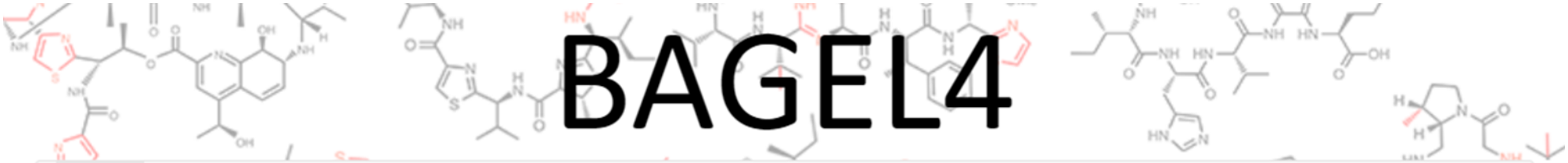
AOI	start	end	Class	Fasta header
NZ_LRPH010000441.106.AOI_01	14237	36366	490.1;Pumilarin	NZ_LRPH01000044.1 Bacillus mycoides strain PE8-15 PE8-15_contig3, whole genome shotgun sequence
NZ_LRPH010000961.12.AOI_01	9074	29074	LAPs	NZ_LRPH01000096.1 Bacillus mycoides strain PE8-15 PE8-15_contig77, whole genome shotgun sequence
NZ_LRPH010001011.87.AOI_01	23597	43597	Lasso_peptide	NZ_LRPH01000101.1 Bacillus mycoides strain PE8-15 PE8-15_contig81, whole genome shotgun sequence

b. PE14-63



Region	Type	From	To	Most similar known cluster		Similarity
Region 5.1	RiPP-like	16,970	27,830			
Region 6.1	betalactone	182,783	209,654			
Region 7.1	siderophore	72,491	87,363	vibrioferrin	Other	54%
Region 11.1	RiPP-like	90,700	101,743			
Region 12.1	PUFA, hglE-KS	70,520	126,294	eicoseicosapentaenoic acid	Other	26%
Region 24.1	arylpolyene	13,427	55,436	APE Vf	Other	85%





Home | Core Peptide Blast | Databases | Help | Links

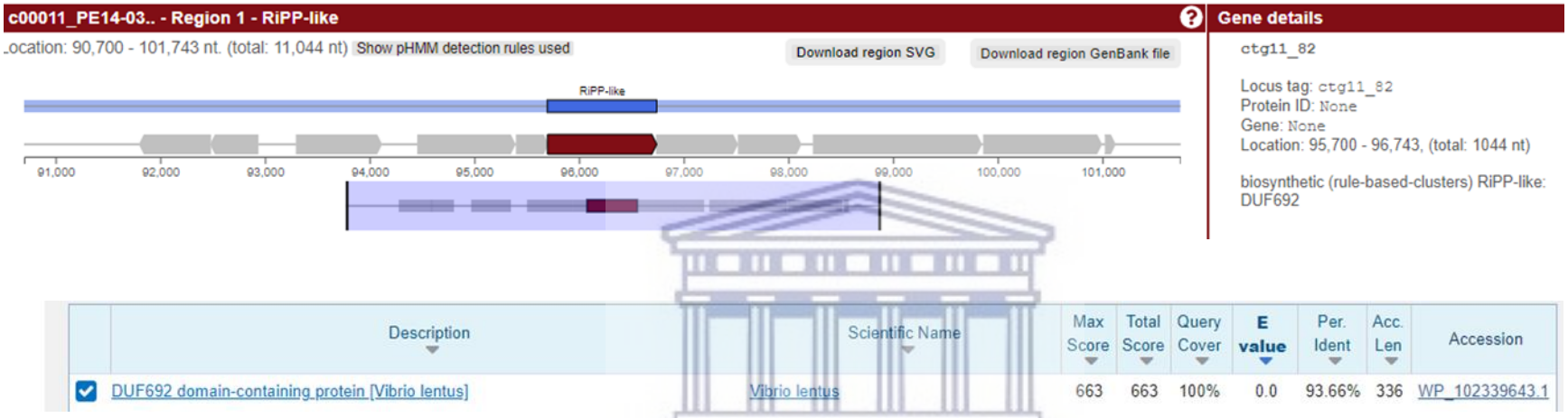
.....

Stop session and start a new run

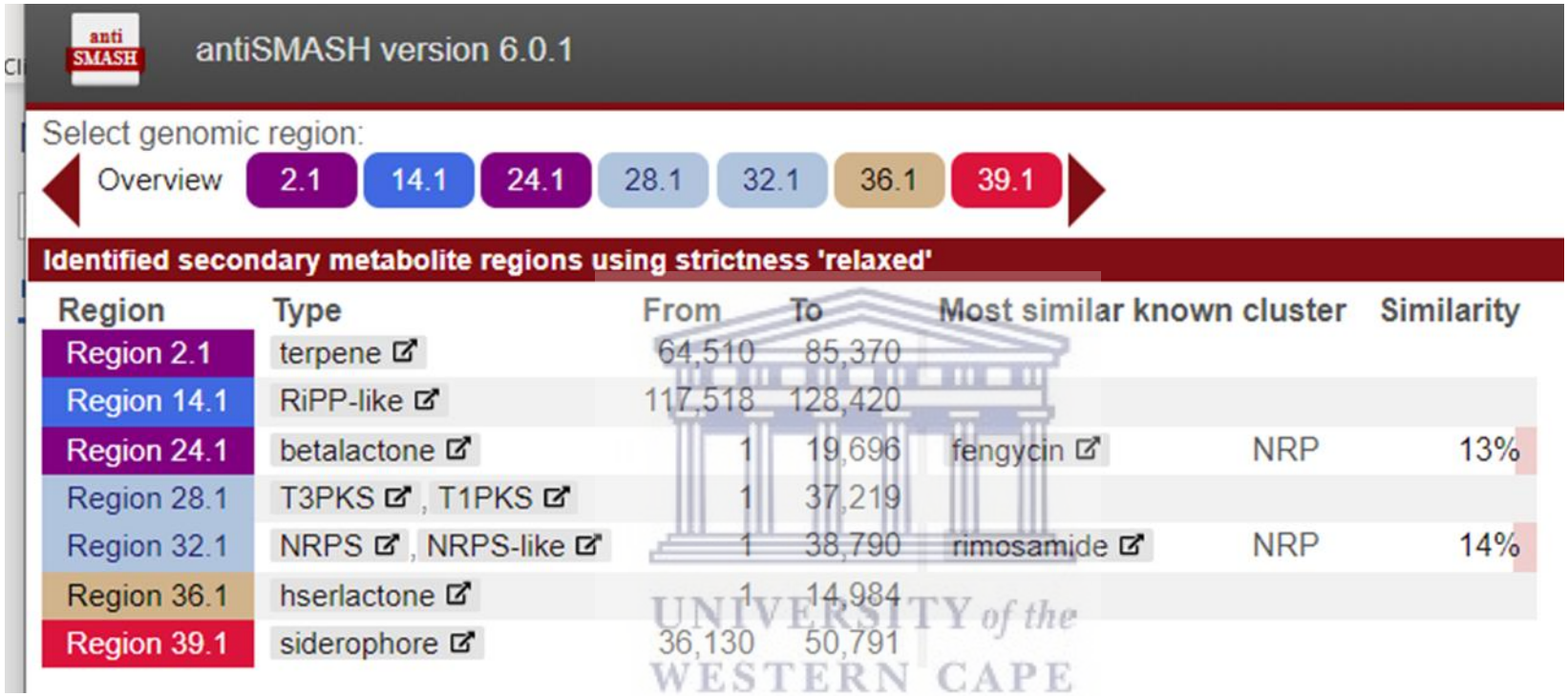
Run summary	
Number of files analyzed	2
Number of DNA fragments analyzed	24
Total bases in all DNA	5539638
Number of AOI's **A**reas **O**f Interest)	0

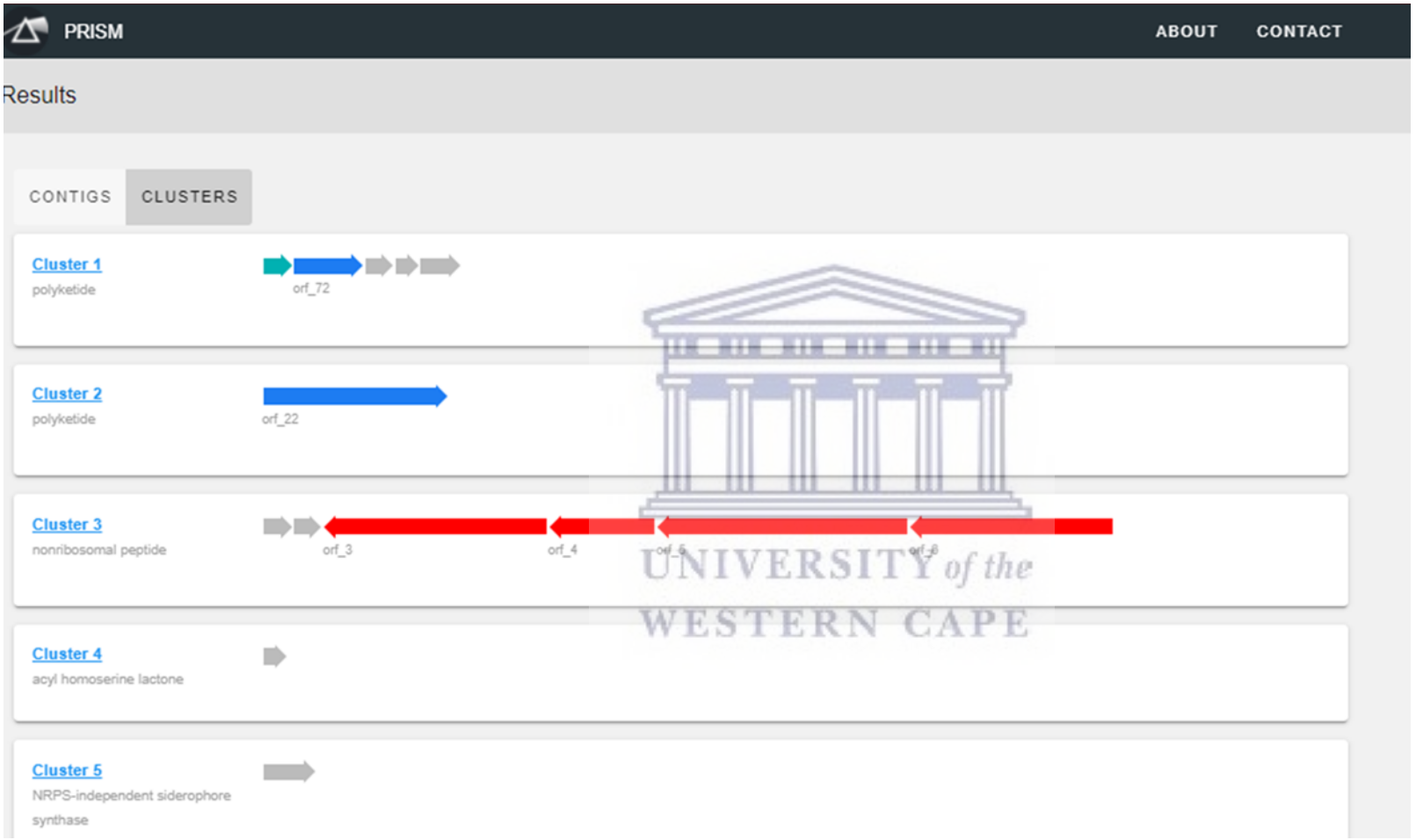
AOI	start	end	Class	Fasta header

Bacteriocins DUF region screenshots

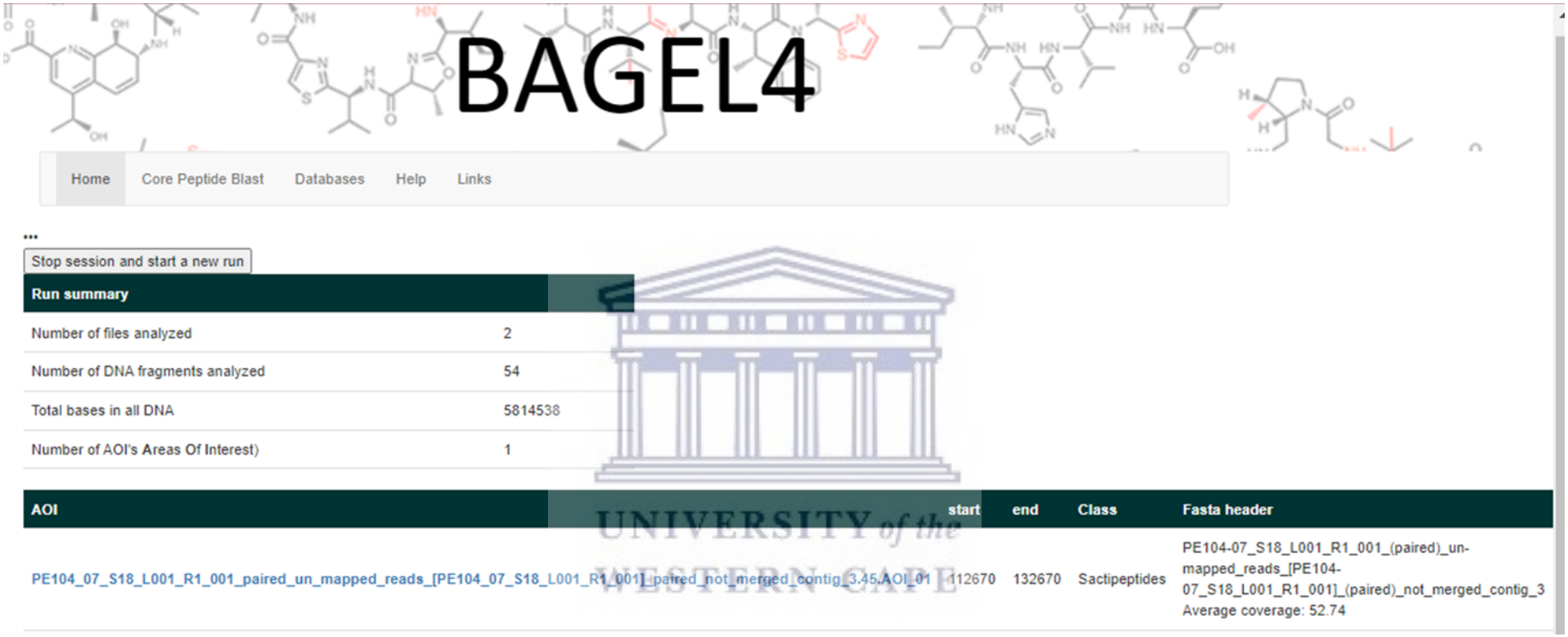


c. PE14-07

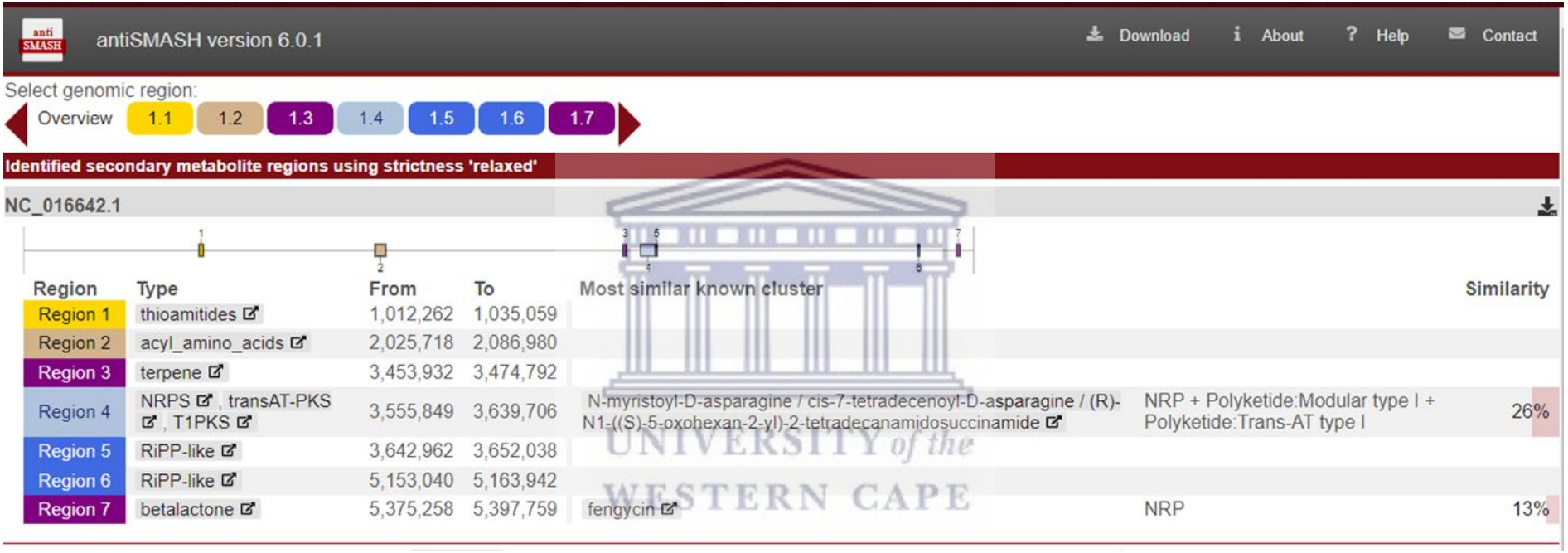




PRISM output for PE14-07



d. *Pseudovibrio* sp. FO-BEG1



Region	Type	From	To	Most similar known cluster		Similarity
Region 1	thioamitides	1,012,262	1,035,059			
Region 2	acyl_amino_acids	2,025,718	2,086,980			
Region 3	terpene	3,453,932	3,474,792			
Region 4	NRPS, transAT-PKS, T1PKS	3,555,849	3,639,706	N-myristoyl-D-asparagine / cis-7-tetradecenoyl-D-asparagine / (R)-N1-((S)-5-oxohexan-2-yl)-2-tetradecanamidosuccinamide	NRP + Polyketide:Modular type I + Polyketide:Trans-AT type I	26%
Region 5	RiPP-like	3,642,962	3,652,038			
Region 6	RiPP-like	5,153,040	5,163,942			
Region 7	betalactone	5,375,258	5,397,759	fengycin	NRP	13%

7.7. Appendix VII: TDA analysis

Query gene / contig from PE14-07 sequenced genome.

>TDA genes

CTTTTCATAGGCCACTCCACATGAGAACAACAACATGAACATAAACACCAAGCCTCCTTGAACCCCACACAAGTGTAAGCCCTGCCCTGCTGAACTCCCCGCAGAACGA
GAAAGGGCTGGGAAGGACCCAGAGACGCCCAGACAATTTTTAAAGGAAGATGCTACATTTGTTTAGACAAACGAGAGACGCTAGGCTTTCACTAGTCTCTCAGCTTATC
GTTCATGCATGAACTCTGCCAAGCTACCGGACTGCTCTAGTACAGTCTGAGCTTGTTTTCGGAGTGATCGGGTTGAGCCGGGGGTTTGCTTGTATTGCTTCTTGTAGTAGC
GCGTGAGGTGCGAGAGATCTGAGAAGCCACAACCGAGAGCAATATGGTCTAAAGGTTGATCTGAGTTTTCGAGTAACCATCGAGCATGGCGAAGGCGTAGGCTTCTATA
GTAGGCACTTGGAGACATGCCGGTATGCTCTAAGAATAAACGACTGAGCTGGCGTTGACTAAGGCTTAGGGTGTTGGCCAATTGCGTGATCTCAAGGACCTCTTCGAGGT
TCTGCTCCATCAAGAACACAGCTTTGCGTAGTCTCACATCGCTAAACTGAGGGATATCCAGATCAAAGGGCGTTTGAGATTGCCGAGCTTCTCTGGCGGAAATTAGGATC
AGGTGCCGCAGGCACTTGATGACTTTCCCCTTACCCAAGTGTTTCTCCACCAGGTATAAGGCCAGATCTGAGCTGGCTGCTCCTCCTGCACATGTAATGATGCCACGATCT
TCTATGAAGAGCTCGTCTGTAACTGCTATTGATGATGGAAAGTGATTCTTAAACTCCTGATAATGGTACCAATGCAGCGCCGTTCGTACATTTTTCA

TGATACCCGCGCGGGCAAGAGCGAATGTACCGGTACACATGCCGATCATTTTATTGCCTTGCTGACGGCATCTTTTGATGAACTGGATATAGGCATCAGGATATAAGCAA
TCCTTGGGAAGCTTTCCTCCGATAATGACAATGTAGTCATAACCTGTAAGATCGCATTCAAGTGTTTTGTTGGGTGTGACCGTTATTCCGGAACTGGACATGATTGGGTTA
GTCGAAGCGGATAACAGATCCCACGAGCAGTTGATCTGCCTGCTTTCATCACCATCGTCTGCCGCATGGCGTAAGACCTCAGTCAGACAGCTTAGTGCAAGAAGTGTGAA
ATTGGGAAGCAGCACCATGGCAACTTTCATCTTACACTCTGCTCTCCGGCTTAATGATTGAGGTCTCATCTTCTTCACTCTTAAAGAAAGTCACGTACTTTTCATCGTGAG
TAACTCAATAAGCACACCGATCTTAGGTCAATCTCAAACCCAAACGATGCGCTTATGAGATGTCAGAAGAGACTTACGTTTACACTGAACACAAACCTCTTCTAAACTTC
AGTTTTTTCTCAGCTGACGGTTTCTGTTTTAACTGAAACGCTAGACTTCAGTTGTTTAGCGATCATAAAGCCGAGAAGTACAATGCCAACATATCCATTCACCACATAAAC
AACGTTCAGAAGCTGATCGAACGGAATGAAAAGACCCAGCGTTGCACCAACCAATGTAAGACCGGCTGTGATCAACTTGGACTTTGGCGTTCCCTCTGGAGTAATCCGA
GCCGATACATTCCAAAGCAGAGGTGCTGTTGTTGTGTAAATGGCTGCGATGATGATGAATGAAAATGCGCTTGCAATGACCGGGCTTACATCAGCTGCCAACACAAGCA
TCGGGATCTGTGTTCCTGCAACATTATTCAACTGCCCCATCAAACCCAAAGCGACTACGATAATAGCGCCTGAGAAAAGCATCGCACCCAAAATGCCACCACCTTTCACT
TCAGAGTCACTTGATGTTTTTGAGCCAATTGCGGCCATGAAGGCACCAAGCAGCAACATGCAAAACCCAACATAGTTGAGAGCTGCAAAGAACCAGTTGCTCGAGGCTT
GCAACACAACACCTTCCTCAACTTTGCTTTCAATAAGCGGAGTGACATCAGTTGGTATCAAATTAGCAAGGTTACTGAAGATAGAAACAAGCCCAAGAAGAATTGCAAC
ACCGGCAATGACGGGTCCAAGTTTCCCGATAATGTCAACGATACGATTCAGGCCCATAATCACTGTGACAAACGCTAGGATTGCAATTAACGAGCCACCAAAGATCGTT
GGCAATCCGAAGTGTTGGTTCAAAGCAGCACCTGCGCCAGCAACCATTACTGAAAATGTCAGGAAGCAGAACAGTGTTGAGAAGTAATCATAGAACATCCCGACGTGTT
TACCGCAATAATACTGATAAACGTCATTTGCTTTATTGAGCTTTGATGTCGTCCAACAGACATGAAGCAAATGGATACGTAAGCAAATAATATAAGGGTAAGAGCAGC
GCCCAAGATGCCCCAATATCCATATGATACGAAATACTGAAGGATCTCTTGTCCCGTCGCAAATCCCGATCCAATCTGGTAGGCGGTAAATGCTCCCGCAATAGTACCAA
TCTTCATCAGCGAGTTTGATTTCGACTGCATGACCATTCTCCCGTTGGTTGTGTTGTCGATTAGTTGTTCGGTGTCAGAGCTTTGACGTCTGGGCCCCAGGTGTCAGACATC
GTGTAGCCAAGAGGGAAAGGATCGGTAGGATGGACGCCATGTTGCTCAATTGAGTAAATCCAACAGCGGCCAGTCAGGCGCGGTATAATCGCGTCCCTATCACCAACCT
TGCATGTGGCAGCAATCTCGCCTTCAAACCGACCACCAATCGTGGACTCCATGAACACCTTGTCGCCAATCTTCAGCTCGCCACGGGCATGTAGGCATGCCATTCGGCTT
GAAGAGCCCGTGCCACATGGAGAGCGATCGACACGACCGGGAAAGATGATTGTTCCGTTCTTGTAGACAGCTGTTTCGGTTGGGTGCACTCCGGTGAACATGATATACTC
AACATGGTTGAACTGCTCGATCTCAGGGTGTCGCACCCTCACCTGCTGATCGACGGCCTTTTTGATTGCGTGCGCTGTATCAACCATCTGCCGTGCGTTCTCTGGTGAAAT
CTCAAAGCCAAGGCCTTCAGCATCCACCATGGCAAAGTAACAGCCACCAAATGCCACATCGACTTTGATTTCACCGAGAACATCTACCTGCACCGGATGATCAAGGTGCT
CTACAAAGCTTGGGCAAATATCGATTGTGACTTTTTGAACTTTGTGGTCTTTGCAATGTGCAACAACCGGAATGAGGCCTGCTGGTGTATCAATTGTCAGAGTATTCTCAC
CGTCTTGCATTGGCAGCATACCAGATTCCAGAACAGCAGTTGTAACGCACATTGTGTTTGACCCGGATAGTGCATGGGCGGCATCCGGTTGCATGACGATAAAGCCCACA
TCAGCTTTCGGATTTACTGGTGGAAGAAGCAGGTTCACTGACATCTGCGCACTTCCGTGGGGCTCATGAAGGGCAAAGCGGCGAATGCTGTCATCCACGTTGTTGATGTG
GAGCATTTTGTCCAACATCGTTTCTCCCGGAATGTTGATCACTCCACCCGTAATGACGCGGCCGACTTCACCCTCGGCATGCACGCCGAGCATGGTAATACTTCTATTCCA
TTGCATTGGATTTTCTCCAGTTGTCGAAGAACTAATTTTTTGCTTTTTGTTTCAGGAGTTCCCGCGCAGTTGCAGTTACTGCGCATCCACCTCTGGCAGTGCAGAGCATGTC
CGAATGTGTTTTGCTGGCCACTTGCATCATGGACCCGATGAGCTCATCCATCGGAAGGACAGCTTCAAAGCCACCAAGGACTGCTGTGACCGAACTGATGGCATGAGAA
ACCCCGCCAATGTTTCTTGCGTGACACGGGATTTCCACTTCACCGTTTACGGGATCACACACCAGCCCCATGGTGTTCATCATGCAAATTGCAGCAGTGTCGAATGCTTGT
TGTGGTGTTGCACCCAGCACGTGTGCAAGGCCAGCTGCTGCCATTGCAGCTGACACACCTGTTTCTGCATTACAACCTGCAACTTCTGCTGCGAATGTAGCTTGCTGTGCA
ACAATCGCTCCCACCAAAGCCGCCACTTCCAAAAGGTTGACGCAGTCATCTTTTGTGAAACCATTTTTCAGAAGAGCGTATGCGATACCCGGAACTATCCCAGCAGAACC
GGCAGTTGGTGCGGCTACTACAATGCCTCGACTAGTACCGGACTCCATTGCAGCTAATGCCGCAGCAATAGCACAGCTCAAGACTGGGCTGGCGATGGTAGAAGCGTGG
ATAGATGCCATCAACTGAGATGCTTTGGGAGAAAGAAACTTCATTCCCTTTATGTCTCGTTCTAACGCTTCTTCTATGGAAGTGAGCATCGTTTCCAATCTTTGCAGAAAA
AGCTCGCGTACTTCTGCTTTTGTCAGACCAAGTTGCTTGGCCTCATATTGCTCTGCAAGCTCATGGAGAGACACCTGATTTGCTGCATTTAGAATATCGCTGCAGCTCTCA
AACAGCGGTGCTGCATCCTTGATGATGTACTGTGCTGGTTTGCATTGTCTTACGCGAAGAACATCATCTAATCGCTGTAGGTTCGCGGTGACATCGAGACTAAAAGGCTC
CAGAGTTTGAATTTCGATGGTGTTGCCCGAGCAAAAGGTCGAGCAAACTCCAAAATCAGCAAGAAGATTATTTTGTCTAATATGGGCAAGTGTTTGCTCAACTACCCGCT
GACTGCTGGCCTCGACCAGCGTGGTATAGAATGCACCACTGATCTGTAAGTCAGCTCCATTAAGGGACGTAAGCTGGAACATGCCGCCACCAATAGAGACAGAACGCCA
GTTATCGTGTCTGATTGAGCCATCTTGCTGCTTTATCTGTAGTTCAAACTCAGCATCGTTGGGATGGCTTACCGCTTGCAAAGGCTTGATATTAAAGCTGAGGTCGAACGG
CCTATCCTTCGTAATATCATCTTGGGCATGAGCGTAATCTGGATGAGTCATTGCCCGACCCAACAGTCCTGCCGCAATACCTTGGTCGGAACCCTGGTTTTTATAAACACG
CGCATAGGATCCAGAAGGATCGAAGTTGATACGTGCTGAGAGTAATCGCTCACCATCAACAAGACTGAGCTGCCGAATACTATCTGCAATAAAAAATGGCCCTGCGGTA

TGAGAACTCGATGGCCCTCGCATGACTGGACCCAACACTGAGTTAAGCAAGCTGTAGCTCATAGTTAACCTACGAATCCCGATAGTGGGATCTTCATTCAGTACTGATCA
GACAAACTAGAAGGTTCAGGAAGAACGCAAATGTACTGAGTCACTTGCCCTTTACTTACCTAAGATCAGTTGATTATGTGCTATTTAAAATACTTCTCGAAATCACCCTTC
GATAGGGTCATATGAGTTTTGTATTTCTTGCTTAAATTTATAAATATACTTGATTTTCGATAATAGGACGCACTGGACATCATGCAGAAATGATAGGACGTCTAGCCCTCG
TGCAAATACTGAGTAACTCAGTGGCCAAACAAAAAAGGCACTCCCATGTATAGAGTACCTTCTTGCCTGAAGCCTAGAAGCAGCTTTTGCTTCTTTGCAATGGATAGTCC
GCTAAATCAGAGCGATTGCTTCCTTTGAACATTTTGAGGCATCCCCAAGAGCTGCACTCGAAAGTTCGCTTATGATCTGCTCTTGTGACATCAAGCCCATCTCAACAATGA
CATCGACAAATCGTCTTCCATTTACATCTGCTTCCTTCACAATCTTGGAGACCTCAGCATATCCAAGTATAGGCACGAATACGGTAGCAAGAGCAGATGAGTTGAGAAGG
TTTTGAAGTGCTTTATCTTCCAGTGCCTCTATTCCGATCACGCATTTTTTTGCAAACAGGTCCACGCCATGTGTGAGCAAATCTATACTGTCAAACAGCCGCGAAGCGATG
GCTGGTTCGAAGTGGTTGATCTCTAATTGTCCGCTTAGAGCGGCAAGTGAAATCGCCGTGTCATTACCAATCACAGCAAATGCCAGCTGTTGAACCATCATAGGCATCAC
CGGGTTCACTTTACCTGGCATGATCGATGAGCCTGGCTGAACCGCGGGTAACTTGATTTCACCGATGCCGGATTGTTTGTCGGAAGACATCATAATTAAATCGAGGCTGA
TCTTGCCCAAAATCTCAGCAGTAACTCGCAGTTCAGCAGAAAGTCTTGAAAACCCATCTGAGTTCTGCATCGCATCAAACAGGTTGTCAGATTGCTGTACAGGATATGGC
AATTGTTGAGCTAAGTACCGATAAACAGCTTCCTGATAACCTGCACGAGCACCGAGCTCTGTTCCGATTGCAGTTGCTCCCAACGGCAGAACTTGCAGGGCTTCAGTAGT
ATCCATGATTTTGTCACGGGCTCGAAGAATAGCAGCTGAATATCCGCTGAACTCCTGCCCCATGGTCATTGGTTGCGCACCTTGCATACATGTACGGCCAATTTTGAGTAC
TTCGGCAAACTCATCTGCTTTTAGCTGAAATGCCTCTCCCAGCCTTGTCAGAGCCTTAGTGAGCGCTGGAACTTTCTCCAGCACAGCCAGCTTGACAGCAGTAGGAACCA
CGTCATTGGTTGATTGTCCCATGTTTACATGGTCGTTAGGCTTCAAGTATTCGTAAGCTCCAGCCTCATATCCCATAATCTGCAAGGCACGATTAGCGATGACTTCGTTGG
CATTCATGTTGATAGACGTGCCACCAGAACCTTCCAACATGTTGATGATGAAGTGGTCATGATGCTTGCCACCCATAATCTCATCACACGCCGCGGAGATCGCATTAGTC
TGCTTGACGGTTAGACCTCCCAATACACCATTAGCTTGGGCAGCTGACTTCTTGATGATCGCAAGAGCCCGGACCAGAGCTTGGTCGTCGCCGATCATTGCAGAAGAGAT
CAAAAAGTTTTCTGCACCGCGAAGCGTGTTGATGCCGAACAATACATCATCAGGTAGTTCGCGAGACCCGATGCTATCAGTTTCAATTCTTGGCATCCGTCATGCCTCCGT
AAATGTCAGGTTAGAATGATCTTTCCGTGCTCCAGGCAACGTCAGAACGCGGTTATGGAAGGCGATATAGATGCCGGGCTGTAAGAGCCGCGCGCACACCAATGCCTCA
ATCATGTTCTGAGAAGCATCAGACCCTTCAATGACATAGGGAACCATCGCACCTGTAAAAATCACCGGCACACGAGGTGCTTTGACCGCACAACATAATGCCTCACCACT
CAGCGCCATGCTGTCTGTACCATGCGTGATAATCACCGCATCATGAGTTGCAGCAGCGCCTGCAACTGAACGGGCAATTTGAGCACGTTCCTCATCACCAATTTCAAGAG
AGTCCAGATGCATCAAATCAATGAACGTGAACTCCAGATCTGGTAAGCTTAATGAGGTAATGATCTCTTTGACTTTGAGTTCAAAATTATAAAGTCGGGCTGAGTTAGGA
TCATAGCTTTTCGCTATTGTCCCCCCCGTCATGACAACAGCAACCTTCAGCAGCCCACTCTTATAATTCATCGTTGAGTTGGAGGCATCAGTCATGAGCAGCTCTCTTGTT
ATTGGGACGTACCACCAGCACAGAGCAATTTGCATGACGAACGATCTTTGCTGCATTTGGACCCAGCAAGCAATCTCCGAAATCAGGAACGTGAGACATCAGGATAATG
AGATCTGCTGATAACTCATCAGCGGTTTGCAGGACCTCCGAATAGATGTTGCCAGATCGAGCAAACCGATTGATGATTCCGAACTCCTCTGTTTCGGTTATCTCGCCAAG
CCGGGACTGCGCGACCTCGAGGGCTTTTTCCTTCAGCTCGCCAGGCACAAACGATGCAAAAGGTGCAGCCTGAGCGGCCTGCGCAATCACTGTGCCCAGATCAGGAACA
ACTGTCATGCAGGTCAACTCAACTGAGTTGCCTTTTGCATGAAGCCTTAGCTCTGGCAAAATCTCGGTCCATGAGCGTTTTTCTTCCACATCAATGGGAATGAGAATTCGA
AGTGTCATTTACCTGCCTCCTACCTGATGCCAGTTACGGCCATCGCCAGCCACATGGATAGTATTGGGACAAAGGTGGTGAGCAGAACCACTGGCATATATGCGAGGAA
AACAAACACCAGAGTAGGTTTTAGAACCTTCGCAAGCGACGTCTGCCCTATCAAAGTACCAAAATAGAGGATTGGCGCTGTGGGCGGCGTAATAAGACCCATGCCCAGA
TTGGTTCCAATAATTGCAGCAAAGTGCACCGGATGTACGCCAGCCTCGGTCACCACAGGCAGCAACAGCGGAGCTGCAAGCAAGATGCCAGAAACATCCTCCATAAACA
TCCCAATAAGCAACAGGAAGCCGTTTACAAGAAGCAGAAGTACGATCGGATCATCAGAAATGGAGAGCATCGTCTCCAGTATCTGTTGAGGCACTTCTTCAAGCGTCAA
CACACGCGCTAGCATGTTTGCTGTAAAGATCACAAGCATCAGTACCGAAGTCACTTGAGCTGCTTTCCAAACCACATCGTAAAACGATTTGAATGTCAGCTCCCGATAAA
CGAGAAACGCGACTGGCAATGCGTAGACCACAGCCACAGCAGCAGCTTCAGTAGGTGTGGTGATACCTCCATAGATAAACCCCAGGATGATCAAAGGCATTGCCAACCC
GAAGCCTGCTTCTCGTGTGCGTTGAAAAACTGTTTTGATCTGCACATTCAACGGTTGTGCTTCAGGAACAGTAATCTCCATCTTACGAGTCAGGACCATGTTCCAGAACGA
AAACAGGCCAACCAGCAACAGTGCAGGAACTAAAGGGGCAAGGAAGCTTGCTGCAATAGAAGTACCTGTTACCCAGCCATAAAGGATCATGGTCGCGCTGGGAGGGAT
GAGCAGTGCCAGAACTGCTGAGCAACTGACCAGAGCAGTCGCATAGCCACGGTCATATCCTCCGCTCTACCATTCTTGGAATTGCAATAGAACCGATAGCGGCAACAGCT
GAGGATGCCATCCCGGAGATTGCGCCAAAAACAGCTGTTGTAATGACGGTTACAACACCAAGGCCTGAGCGGTATCGACCAAAGAAGCTGTCTGCAACCTCAAGCAAAC
GTTTGGCAATTCCACCTTGTGTCATGATACCACCAGCCAGAATGTACAGTGGGATCGCAAGTAATACGACTGAAGATAGCTTAGAAAAGCCTGCGGTTATCAAGAAAGT
TGTCGAGTTATATTCGCCAAAGATGATCAGGAATGCGGCTGCGGCTACAAATGTCAATGCAACAGGCACACCAAGGATCATCAACGTCACAAGAAGCGCAACGTCCATC
AAAATTAGGGAGCTCATGATTTAGGATTCCGTATTTGGTGCGTGCCAGATTGTTGCGCGGAAGCTTTCAATAAACTCGAGCAGGAAATAGAAAGCCGTCAGAGCCAGTC
CCACGAACATCGTCCCATAGATCCAGATCATTGGGAGACGTAGCTCCGTTGATACGGTGTTCATACGCTCTGTGAAGACGAGCAGCCCCCAGAGTGAGTTGAGGATATA
AACCACCAAAATTGCAGACAAGAGGCGTGTGAAGGTATGGACGCCCGCAGTGATGCGCCCTTCTCCGAACATTGTATCTACCAGACTGGCCGAGATGTGACTGCGTTCA
TGCGTCCCATACGCTGCCCCGATAAAGTACATAATGACACCAGCAAAACTGGCGAACTCTTCAACTCCAAAAAGCGAATAGTTGAACAGATATCGCATTACGACCTGAA
GAAAGACTGCTGCGGTGAATGCGATACTTGTGATGAACATAATGAACCGCAACAGCTTCAGCTTAAGTTGCCAGATTGGGTTCTTAATTTGTAATTGGCGCTCTTGAAGC
ATACCGCCCTCTTTATTTTGTCCGAAGACCGGAGCAGATCTCCGGCCCTCGAAACCCATTAGTTTGAGTAGCGGCTTTCAAGCAGCTCAAAGAGCTCATCGCCGATCTG
ATCTTTCATCTTCGGCCATGCATCAGCGCGCACGGCTTTTGCCATGTTTTCGATCTCATCACGTGAATAGGTAACTACTGTCACCCCGTTTTTGCGAAGCTCTTCCAGATAG
CGATTGTCAGCTTCTTCTACGGCGTTAAAGCGCTCCGTAGCCATCTGGGCAGCTATTTCGTTCAATACAACGGCATCTTCTTCTGGTAACTGAGCCATGGAGAAAGCAGA
GGAAAGCAACCAGTTGATTTCGAAATGATCGTTGATCTGCAACCATGTATCGACAACACCATTCCAATCACGAACTGCATTTTCAGGAACACCACCAACCATTCCATCTA
TTACACCGGTCTGCACCCCAGTGTAGAGTTCCGCCCAAGGCATTGTGGTTGTTGAATATCCCAGACGCTCCAGCAGCGGCTTGTGCGTACCCACACCTGATGGCCAGACA
CGAATCTTGAGTTCACTATGGTCCGCATTCGGATCTTTCGGATCATCGACGGGCTTGCCAACAGCCAAACCACCCAAGCCGTTTGCATAAGGCGCCAAAACCTGAATTTG
CTGCGGTTTTACGAGGTCATTCACGATATTGAAGAGAAAACCACCATTGTTCCAGGCTTCTTTTGCGCTGTTGTAATCAACCGCTGTATATGGGAACCAAGTCAGAGAGA
GACGTGAGTCAAACTTCGTTGAGATTGACTGAATTGCCAGATCAACTGCCCCTTGGCTCAGCTGCTCGTAAACTTCTTCCCAATCACCCAACTGACTTGCAGGAAAAATA
CGCACATTCACGCGCCCATCAGAAGCTTCCGAAACACGGCTAGCGAAGTCCTTAGCCCAGTTATCCATGTCACTGCCTGGAGGTGATAAGTGAGATAGTCTAAGCGTTGT
CTTTGCTTCAACCGCAAACGTTGAAAGCCCGATTGTCATTGCTGTCGCTAGCAAAACCGCAGCTGTTTTCCGTAACTTAAAAACCAAGTTCAATCCTCCCAAGTTAAATCC
CCTCAACAATTCGATTGTGAGATCACCCTGAGAATAAGAGACTGACCGCATCCTATAAAAGAACACATAAGTCATTTTGGAGAACGCAGCAGACAAGATTAAGGGGAAT
TACATTCAGTAATTCCCCTCATCCGAGACTTTACGTTTGAGTTGTTCACCCGTTAGTCTAGCGCACCTGGAATGACATGGTCGATGCCAGTTCCCCAGGTGTCTGCCATGG
TGTAACCCAACTGGTATGGATCTGTTGGATCAGCTCCCAACTGCATGGTTGCGTAGATCCAGGCCCTCCCAGCGATTCGAGGAATAACGGCAGGTAGCCCAGCCACTTTC
GTCTCAGAAAGAATAGAGGCCTCAAAACGAGAGTTGATCGTTGATCGCATTTCGAGCTGCTGATTAGGTTTAATCAATCCCTTGGCATGCATGGCAGCCAGACGAGCAG
CGGTGCCGGTTCCACATGGCGAACGATCGACACGTCCTGGCGGCATAATTGTAGCATTACGATATGAACTCTGCTTTCCATCAATCATATCAGTAAACATCAGAAACTCG
ATGCTTTTAAAATTCTCAATTTCTGGATGCTGAACAGTGAGCTGCTCCTGAGCTGCACTTTTGATCTTGCCTCCAAGCACAACCATGTCTTGCGCGACATCTGGTCCAATA
CGAAACCCAAGCTGATTCACATCCGGCAGTACGTAATAAACACCACCAAATGCGACATCCACCGTGAGCGTTCCCAGCCCTTCTACCTCAAGCTGAGCATTCAGATGCTC
AGCAAACGCCGGGAAGAAGTCCAAGGACACTCGCTCGACACGCCCTTCTTTACAGCGCGCAACTGCTTTTACCAGTCCAGCCGCAGTGTCCAACACAACTGTCGTTTCAG
GTTCATGCATTGGCAAGATGCCAGTCTCCAGCAACACCGTCGTCACACACATTGCATTAGAGCCAGACATGGCATGCGTCCCATCGGCCTGCATGGGTATAAACGCAGCA
TGTGCCTCCGGATGTGACGGAGGCAGCAAGAGGTTCGTCGTCATCTGAGCACAACCGCGAGGTTCGCGATTTACAAACTGACGAAGCTCATCATTGGCCTTCAGGTGCTC

CATCTTCTCAAGAATACCAGCTCCGGGAACGCTCAGCACTCCACCTGTCAAAACACGCCCAACTTCCCCTTCAGCGTGCGCACCAATTACAGTCACAGATTTATTCCAGC
GCATGTACCCTCCTTGTTACTTATGGAGAAATGCAGTACGTAATAGGCTAATAATAGAACAAAAAAGACGTTTACTTGCAGATCTGGGAGTATCAGCTTGCTAAGCCCTC
ACATCACAGATGAACAGATTGAACCAGACGTTCGTTTTGCTTTTATTCTCGCTCCGGCGTTTACAATCCTGCCGCTAGCAGGATTTATCGACGCAGTGCGGCACTCCTCAG
ATGAAGCAGATAGAAGCAGGCAAGTGTTCTGCCACTGGGATATTTTAAGTTATAATTTGGAACCCATAGCATCAAGTTCTGGTCTTGCAGTTATGCCATGGAAAACGTTT
GAGGAAGCAGGCACCTATGATTATGTCGTCATCGTCGGCGGCCAACTTGAGCAAATGGAACAGATACCAGAAGAAACCATTGCATTCTTGCGAGACGTTCACACCAATG
GGACAGAACTGATCGGACTTTGCACTGGCAGCTTTATCATAGCCAAGACAGGCCTCCTTGACGGTAGAAAAGCCGCGATTCACGCGCACCATCGCCCACTCTTTTCGCAA
CAATTTCCAGAGGTTGAGCCCATAGAGAATGAGCTATATGTCCGCGATGGAAGTATCCTGACCTGCCCTGGTGGAACCGCTGCGATTGACCTTGCTGTAGAAATTTTGAT
TAAGCACTGTGGGCGCTCCCGTGGCATGAAAGGTTTGACGGCGCTCGTCGTGGACGAGCATCGATCAGCACATGAGATTGGTCGGCTTCCCTATCAGGATCTGGAGGAA
TGCGGAAACTGGCGCGTGGAGCAAGCTATTAAAATCATGCGTCAGAAGCTAAGAGATCCAGACACAACAGAAAAGCTTGCCGAAATGCTGGGCTCCACCGTTAGACAG
TTGAATAGAGCTTTTACAGAACATGCAAAAGCCACACCGCAAGAAGTCTGGCGTGACATGCGCCTCCAGCACGCACGCTGGCGCCTCATGAACTCAAAACGAACCGTCA
CCCAAATCGCTCACGAATGCGGTTTCTCAGACAGCAGCCACTTCTCCCGCTGGTTCAAAGCCAAATTCAACCAAACACCTAAACTCTATAGAACACAACGCCTGGTGGGG
CAGAGATAAGATCACAGTCTAGGTTTAGCACTTAACAAAAAGGCAATACTCTGACTATCCGATGGCTTGCTTCAAATTAATCCAAAACGGACTATGAGCATGGCAGTAG
CCCCGGTTGAGAATAACCATCAGCTGCCCTTCGCATTGATCTGAATGCGATCCCAACTCGAGCTAGATTTCAATCCCGAGGCAAACTTATCCAACTCATTGCTAAGGGCA
TGAAAGCGGGTGCATTGCAATCGCTATCCACGCTTTGCCGCCCAATTATCCCTTTACCGCTCCATCGCTGAGGCCTGCTACAAGGTACTTCTGCGCAATCATAAAAACTAC
TGTAATGGGGAGGCCTGATAAGATGGCAGCGGCGGCGAAGTCTCCCCAGAGGTATTTGAATTCATTAAGATAAAGGCGTGATCCCACAGCGAGTGTCAGGTTCTCTTCT
GAGCGCAGCAGAATAGATGCCAAGGGGTAGTCGTTGATTATGCCAATGAATGACAGGACGAAGACGACAGCGATGATTGGGACTGCCAATGGCAGGAAGATGAAGCGG
AAGGTTTGCCATGGTGTAGCGCCGTCAACCTGCGCTGCCTTGTCCATGGCTGGGTCCAGATTATCGAAATAACCCTTAATGGTCCAAACGTGTAGTGTGATGCCTGATAG
ATACACTAAACCCAACGACCAATGACTATCGATACCGAGCACAGGCGATACTTCTCCGAGGCCATCAAAGATTGCGAAGAGAGCTACCAAAGCCAGTGTGGTTGGAAAC
ATCTGAATGATCAGCATGGAATCCAGCATTACTCCCGCCCCCCTAAACCGAATGCGGCTGAACGCGTAAGCTGAGATCGTGGAAAGAGTGACCACACCGACAGATGCGA
TCAAACCGATTTTGATCGAGTTCCACATCCATAACAAGACAGGAAAAGCAGGCTCAACAACTGTTCCATCTGCTCGCGTATAATCCAACCCCAGCGCCAGTGCCCAGTGT
TCGAGCGTCAAACTCTCCGGAACAATTGAACCTGTTGTAAAGTTTCCCTCCCGAAAGCTAATTGATATGACCATTATGAAGGGAAAAAGGATAAGGGCCAGAAAACTCA
AAAGGAATGCATGAGCAAAGGCCTTCTGCAACTGAAGTCTTCTTGGGTGCTCTACGATCATGACGCTGCCCTCTTAGCTGTTTGGCGACGCATAGCAACGAAGTTGGCGT
AAGCCAATGCTGCAACGATAACAAAGATGATAAGTGTGATTGCACCTGCTAAGCCGTATTCCTGACCAGCATTTTGGAAAGCTATACGGTAAGTAAACGATCCAAGAAT
GTCTGTCTGACCTGCAGGAATGATTGTGCCCGGAATATCAGGTAGTCCGCGCGTCAGGAGCAGAACCAGTACGATATTGTTAAAGTTAAAGGCAAAGCTTGCGACTAAA
AGCGGAAGGAACGGCGGTAGAATTTGCGGAAATGTAATCTTCGGAAAATTAGAAAGGCACTTGCTCCTTCAAGAGTTGCTGCCTTTTTGTGATCTTCGGGAACTGACTG
CAAAAAGCCCATCGCGAGCAGCATCATATAAGGATAGCCGAGCCAGACATTGACGATCACGAGCATCGTTCTCGCCATGCCTCCATCTGTCATCCAGCTTGGACGAACA
CCAAAGAGCGCATTCAGGATCATGTTGATCTCACCAAAGTTCTCGTTGAACAGACCCGCGAAAATCAGGATCGAAATGAAGGCTGGAACTGCATAGGGCAGTATTAAAA
GAATGCGATAAACTGCTTTGCCGCGAAGATGGGGCCATTGTAAAGTAACAGCCAGTGCCAAGCCTAAAATAAACGTCAACCCGACTGAAAGCAGTGCGAAAAGGCAGG
TCCAGAGAAAAATAGAGAACATGGGGGCACGAATGCCCTCAGACTGAAAGATCCGCTTAAAGTTGGCAAGCCCGACGTTAACTCGCCAGCCTGGAGAGACGCGCTCCC
CAGTCATCTCGTTAGTATAAAATCCAGTCTCAGAGTCTGGAAATAGTCGTGAACCATCCAAATTTGAGATGAGTACTCCTTCGGGCGTCTCACGGAACTCCGGCATCTGG
GCCGCGAACTTTCTAAGGCCTGAGTTTTTGAGAATAACACCGTCTGGCAGTTGAACAGAAATCTCGCCAAGAGCTTTGCGGAGCTTGATAACATCACGCATGCTCAACTT
CTCGGCGCCCGGTAGGGGAGCCGCTTCTGCTTTTATGTTGGGCTGCGTTACGCTTTTCTCGAACGGTTCACTCAGGTAGCCGTTTTCTTTACCGGTAATGTAAAGTCGGTA
CCTATCGCCTTGCTGCCAAACACTAAAACCTCGCGAGGTTTCTTGATCGACTATTTTCTTGGCAAGGTGGATTTCACGGACCCGTTCATGGCTGAGCAAGTTGAACGAACT
GAAGTTCGTGAAGCCGAGGTAAACTGTGTATCCAACCGGGAAAGCGATAAAGATTAGCACCGCAGCAATTCCAGGAAGAATGAAACGATGCGCATAAAGCCGCTTGGT
TGAGAACACTATCCCAAAACATACGGCAACAATCAAAACGAGAAGGCCGACGACAGGCTGTTTTCCAAGAAATAAAAGGAATGCCGAATAGAACAGCACTGTGAAAAC
AGCGATCAATAGACCAACAGTCAAAAGACGTTGAAGTTTGAAGCCTGATATTTGGCCGGTGTTTTCGGAGCTAAGAGAGGTCATCAGAGGCAACTTTCTTTGCAATCAA
ACAACATCGCACCAGGAGGGTATCCCGGTGCGACACAGTCTTGGTTATTCTCCAAGAATTCGGGAAGCTGGGTCTTTAAGAGCTTCTTCTGGAGTTTGTTTGCCTGTTGTG
ATGTTACCAAGCGCCGGTGCCATGGCCGCCCAAAATGCACCCATTTCAGGGTTTGACGGCATAGGGACACCGGTTGAAGCGTTCTTCAAAGTGGCTGAGATTTTGTCACT
GTGCAATGCTTCACTCGCTGAGATATCCGCGAGTGCACCAAGAGCTCGGTTCTTGTTCCACTCAGCAAGACCTTCGTCAGTCAGCATGTAGTTTTCAATAAACTCACTTGC
CAGGTCAGCATTTGGGGAAGCGCTATTGAGAGCTACAGAATAGACACCAATGAATGGAGGTGAAACATTTCCACTGACTTGCGGAATAGGTGCGACACCGATATTCAGT
CCTGCCTCTTCATAGGCTTTCCAGCTCCACGGTCCATTGATGACCATTGCCACTTCACCACGGGCCATGGCAGCATCCATCACACCGTAATCAACACCGTCCGGCATGATG
CCATCAGTAAAGAGCTTCGCCAGCACTTCAGCGCCTGCAATCGCGCCCTGATTGCTCACACCTGTCTCTTTACCGTTATAAACCCCATCAACTTTTCTAAAGGAGTAGCCT
CCATTTGCCATTAGCATTGGCATCGTGAAGTAGGTGTTGGTGTAGTCCCAGAGGATAGGTTTTACACCTTCAGGCAGTTCCAAATGTGCGAACTCAGAAAAGTCTTTAGG
AGGTTCTGAAATCAGGTCTTTATTGTAGATCAGACCTATCGCTTCAATCGCAATCGGATATCCCCACACCTTACCATTGAAACTCACAGCCTCCCATGCGGTTGGAAGAA
CACCTGCTTTGAACTCCGATGTTGGTTCAACTGGCTGTAGTAGACCACTGGCGGCCCATTCACCAAAGCGATCATGGGCCCACATCACTAGGTCTGGCCCATCACCAGTT
GAAGCCGCACGTTGGAACTTGTCTGGCAACTCCGGGTCAACCAGCTCTACCTTTGTAGGAACTCCCAACTCCTCGGTAAATCTGGCAGCAACCGCCTGCAACTGCTTTAT
ATCCTTGTCAGCTCCGATCCAGATAAGTAGCTCGCCTTCCACAGCGGCCATGGCAGCAGAAGATGTTGTAGCAAGCATAGCTGCAACAAGTGTAGTTCTTAGTATTTTCA
TAGCGTCCTCCCAATTAGAAACGCGTTGTCGTGCCCTATGAAATCAGCTCGAGTCCTTCCAACTGATCAGAAATCAGGTAAGCAAATCCGTTTGGATTAAGCGTGAGTGT
TTGATTGAGCAGGGTTGCGCTGGATGGCCCAGATAAGCTCGCTGCGCCCTCTATTCTCAACTCTCTTACGGTCTCGGACAGATTAAAGACGCAAGTCAGCTTTTGGTTCTT
AGTCATTCTGTGAAAGGCCAGAATTGGTTCCGGCAGATCAAGAAACGCAGTGGCACCTGTCAAAAGCGCGTTGTTATTTTTTCTAAACTCAATAGTCCTGCGATAGGAAT
GTAAGACACTGTCATTGGCGGCCTCTGCAGCGTCAACAGCTCTTGCAAGTTGCGGATCCTTTACGGGTAGCCAGGGTTTCATGGAGGAGAACCCACCATGTTGCTTATCG
GCCTCCCAGACCATAGGTGTACGGCAACCGTCGCGACCTTTAACTGCGGGCCAAAAACGTAGGGCCTGTGCATCTGTCAGCTCATCGTAGTTAAGTTCAGTATCCTGTTG
CCCGAGTTCTTCGCCCTGGTAAATACCTATAGAGCCCTCAAAGGACATGAGCATTGCAATAGTCTGCTTTGCCAACTCTTTATTTGCGCTAGCCTTCTCTTCCCAGCGAGT
AACATGGCGCGGAACATCGTGGTTGGAAAAGCTCCAGCATGCTGTTCCATCCGGCGCAGCCAACCGATAAGACTCGAGCGCATTTCTGAAGTGTGAGGCAGTATAGTTG
GGTGATAGCAACTCGAAAGAGTAACAAAGGTGCAAACGATCTGAACCCTTGGTGTACTCAGCCATCAATTTAGTTGCCGCATCACCACCGTCACCGATTTCGCCGACCAT
CATGCGTCCCTCAAACTCGTCAGTGAGAGTGCGAAGCTTTTTCAAAAAGTCGATATTCTCAGGACGGTTCTTTGAGAAAATGTGCTCCTGCCTACCATAAACATCAGGCA
AGTCCTCGCGATCTGCACGGATGCTTGCAGGATTACTCCGCAATTGCTGATCGTGGAAATAGTAATTGACTGTATCAAGTCGGAACCCATCAACGCCGCGCTCTAACCAA
AACCGTAGAACATGAAGCAATGCTTCCTGCACTTCTTCGTTATGTAGATTCAGATCTGGTTGATCTGCGAGAAAGTTATGCAGGTAGTATTGCTTTCTGCGAGGTTCCCAC
TCCCATGCAGAACCTCCAAACACAGACAACCAATTGTTCGGCGGCATTCCATCTGGTTGAGCATCTGCCCAGACATACCAGTCAGCTTTGGAATTCGTCCGGCTAACTCT
GCTTTCTTCAAACCAGGGGTGTTTACTGGAACTATGAGATAGAACCTGATCAACAATAACTTTCAATCCCAGTTCATGACTGCGGGCTACAAAGTCATCAAAGTCTTGCA
GTGTTCCAAACAGAGGGTCGATGTCGTAATAGTCGCTGACATCATAACCCATATCCGCCATAGGAGACGTGAAGATTGGGCTAAGCCAGACAGCATCAACACCAAGCGA
TGCGATGTGCTCTAACCGAGAAGTAATCCCCTTCAGATCGCCAATGCCATCCCCATTTGTATCCTGAAAAGAGCGTGGATACACTTGATAGATAATGGCATTCCGCCACC
AGTTCTCGTTCATGACTACTCCGCGTTGCATGCGTATTAGCTGTTCACAGCATCAAAACCTAGCACATTTAAATACTGTGCAAACGTTTTGAGTCAAGAAGGTAGGTAGTT

TACCTGCATATAAGCTGAGGCATTTTGCTCCCTCTTCATTTTCTATTTTACAAATGCAAGATTCAGTTTAAAACGTTTTGAAAAGGATTCGAGACATGGCAACGATCAAAG
ACATTGCAAAGGAACTGGGAATATCGCCTGCGACCGTCTCGCGCGCTTTAAACGGCTTCCCAGAGGTGCGCGCAGAAACGCGAAAGCTGGTTCAGGAAACCGCAGCCAC
TCTTGGTTACAAACCAAACCAGATAGCTCAAAAACTAGTCTCAGGGCGTTCTGGCATGGTTGGGATGATATTATCTCTCGGACCAGCACAAACATCAGCGGCCGGTTTTT
ACGAGATCATGGCAGGGTTATCCGATCACCTCGCCGATCGTGATGTCGATCTTGTTTTTCATGCAGCAACGGCGACTGATCCTGTAGCGCCTTACCGCAGGTTCATTTCAA
AGAACACACTTGATGGGTTTATCTTGAATGCTCCAGAAGAGAACGACCCGCGCATCGCATTTTTGGAAGAGCAAAAAATCCCATTTGTCGTCCATGGGGCAACTTCTGAA
GCATCCTACGCCTTCTTCGACATCGACAACAAATCTGTTGCCGTTGATGGAGCTAATCTCTTAGCGGATTTAGGGCATCAGCGGATTGCATTTCTGAATGGGCCACAGAA
CCATGCTTATGCAAAAGCACGTGCTGAGGGCTATCTGGAGGTAATGGAACAGCGGGGCATTACACCACCGGAGTTTGCCTTACAATATGGCAGTGGCGACGAAAACTAT
GGCTACGCAGCAAGTTTGGCCTTGCTAACTGAGCGAACAGCTCCACGCCCAACTGCTATTTTGTGCGCAACAACACCGATTGCGGTTGGAGTGTATCGGGCAGCAAAGG
ACATCGGCCTGTCAATCCCGCAGGATGTCTCTGTACTTGCTCATGACGATGCTATCCCGCAGTTTCGCTCAATCAACTTCAACCCAACATTGTCGGTAACACGGGCTCCAT
TTCGTGATGCATGCGGCCCTCTTGCGCGCATACTCTGCGATCTTTTGGACGGCACGCCAATCAAAGACCTTCAAATTCGACAAAGAGCGGAACTGATCGTTCGGGATTCA
ACCAAGATGTTGCAAGAGTCTGAGAGGATACCATGGCAACCAAAGCACTAAACATCAAAAACTTGAATAAGTCCTTTGGTGAAACCCATGTGCTAAAAGATGTAAGCGC
CTCTATCAATAAGGGAGATTTTCTTGTCCTTGTCGGACCTTCAGGTTGCGGAAAATCTACCCTGTTAAACTGTATTGCAGGTCTGGAAGAAATTTCCTCCGGCATCTTAGA
AATTTCAGGCCGTGATGTCGTCGCGTTACCTCCGCGAGAGCGTGATATCGCAATGGTCTTTCAAAGCTATGCGCTTTACCCAACCATGGGTATTCGAGAGAATATCGAGT
TTGGGATGAAGATCCGCTCAGTTCCAAAAGAAGAAATTAATACTCAGGTCCTGCGCGTTGCTACTTTGCTTCAAATAGAACACCTTCTAGACCGTCGTCCAGGACAGCTT
TCTGGCGGACAACGTCAGCGCGTTGCGATGGGACGTGCTTTGGTTCGGGAGCCAGCACTCTTTTTGTTTGATGAACCGCTGTCCAACTTAGATGCAAAACTGAGGATGGA
GATGCGTGCAGAAATCAAGCAAATGCATGCAAAGATTGGCACAACAATTGTTTATGTCACTCACGATCAAATAGAGGCCATGACGATGGCAACCAAGATCGTCGTGATG
AAGGATGGAGTAATTCAACAGGTTGGGACACCTGATGAGATCTATCGCCATCCAGCTAACACGTTTGTTGCAGACTTCATGGGTGCACCACCAATGAACCTGATCGATGT
AAAGGTTGATGGCAATTCCATTTGCTTTCCAGATGATCAAGTTGTTGAGGGAATGCAAATACCATTGGAAGCCCAAAAAAATGGGAGAGCAAAAATTGGAATACGACCT
GAAGTTTTCTCACTTGGCCAACAACTAACTGTGAAACCTCGCATTGTTGAGCATACGGGTGCAGATCTTTTTCTGACCTTCGATATGGCAAATGAAGTTGTCACAGCTCGC
TTGCCTTGGTCAACGATCGCGCAGGATGAAATCGGACTGAGCATCGACCTTAATCAGCTCAGGTTCTTCTCTACGGTCGACGGGACAGCAGTCAACTAATTTAAAACAAT
AAAACAAGGAGTTACCAACCTGCGCGGTGCGCGGGACGGGAGAGGTGCGTCTGCATCACAAAAAAGCGTTTTAAACAACGCTGGAATAATTGGAGGAAGTAATGAAAA
ATGCTCTGTTGGGGACTGTAGCTGCGCTCTCAGTTCTTGCAGCCACTGCTGCGCATGCGGAATTTAAATACAATCAAACCGGCTATCTGCGTATTGGAGTTGGACAAACC
GAGGGCGACAAGTTCACTGCAATGCAACTAAACGGGGCCTGGAACAAGTACCGTCTTGGGAACGAGTCTGATTTCTATGGTGAGTATGGCGTAGGTGGTAAATATACCT
TCGAAAATGGAAGTGCTTTGATTGGCGGTGTCCGCGCGCATGTCGCCGGTGACAACAACGACTTGATCGTTGATGGCGAATTTGACGCCAATGTTGCAACTCGTGAGTTC
TGGTTGGGATACAAAGGCCTTGGTGAAGGCGCTCTCAGTGACAGTACAGTCTGGGTTGGTCGCCGGCTTCTATAAGCGTAAAGACATCCACATCACCGACAACTACTATGA
AAACTACTCTAACTGGGATGGCTTTGGCGGTGGTCTGGAAGATGTAGATCTCGGCTCCGGCAAGCTTTCCGTTGCTGCTTTTACAAATGGAGATCGCAATACCCATACTCT
GGACGCTCGATTTGAAGGCATTCATCTTGGAGAGGATTTGGTACGTGAAATCGGTGCGGCTTACTCCATGACCACTGGCGACGCTACTGGTGATGACGGCACCGCACTTC
GTCTACATTTGCAAAAGAACAACATTTTCGGCGGTTACATGAAGGCCAGCCTGATGCATGGCTTCAAGCGTGCCTCTGGTTTTTCTGCAGATGGCTGGAGTGACGCGGAA
GATCGCTCAATTACACGTGCTCAGATTCATGGTCTCGCATCCTTTACCGACGAAATTGAACTCTTCTACACCGCTTGGCATCAGCTTGATCGCAACAATGGGGACAACAC
GAACTGGTCTTCAATCGGCATTCGCCCGCAATACAACATCAATGAAGACTTCAACCTTCAACTCGAGTTGGGATTGGATCACGTTCAAACCGAAGATGAGAGCCGGACG
CTCGGCAAAACCACCTTCGCGGCGGCATACACCTTTGGTGAATCTGGTTTCTGGGCGAGACCACAGCTACGCGCGTTTGTAACATATGGCGCATGGAGCAAACCAGGCG
CTGTAGGTAACACGGAAATCTTCGAGGACAAGAAATCTGGTACAACCTACGGCGTTCAATTTGAAACCTGGTTTTGACAATTTGGGGCCTTAGTGGCCCCAATTACATTC
GCGGTATTGTTTATAAAGTCATCTCTTGCTGAAACGACAACAATCAACAAATATTACCAGCCCTATCGTATATTGCTGTTTCTTCAATAAACTACAGAGTTTAACGTCAAG
GTTTGACCATTTGCAGACATTCGTCCTATCTCCGGCTAACATAAATTAGAAGCCTCCCACAACTTTCAAGTGAGGTCATGGCCATTAAGCCTGAGATTAGAAAGGCAATT
GACCAAAGTTTCTGGATGATCAGGCTCACGCAAAGTTAGAAAAGTATCACCAGTCAAATTTTCAATTGATAGCGCATCCAGAAAGAGTGTTTGAGTAACTTGCCTTTCGA
TGAAGGGTAGCAGTTAGCTTGCGTTCAACAAGCTAAGCCGACTGGTTCACAAGATAATTTAACACTTTGAGTTCTTCCAAAACCATTTGTTTGATGAGTTCATTATGAATG
CTATTACCTTCCTGCCTGAGTTATTTCAGGAACAAGACTTAGCCTTTACCTTCCAAGCACCAGCTCTGTCTGACGGTGAATTGAGGGTCAGTGGCTTTTCTGCCAGTGAGC
AGATCAATGCTTTGACGAGCATCGAGGTCTCCCTTGTCAGCCTTGACGCGCATATTGACCTTCATGCTTTACTTGACACGCCTGCGACCCTGACGGTTCATCATAAATATC
TTGGGTTACGTCACTTTGCAGGGGTGATCTCCCAGATTGCTCGTGGTGACAGCGGTCATCACCGCACTTCATATCAGATTACCCTTCTACCAGCCTTGCACAGGCTCTCCC
ACACTACAGATTGCCGGATCTTCCAGAAGAAGAGCGTGCCCGATATTGTGAAAATCCTACTTAAGGAACACGGTATTGAAGATGTTAAATGGGATCTAAGTGGAGCCCA
TGAAGCCAGAGAGTATTGTTGTCAGTACAGAGAAAGCGTTCTGACGTTCCTTCAACGCATTGCTGCTGAAGAAGGGATCTGGTTTTATTTCACCTCAGGCAAAAACGGTC
AGCACACGCTCCACTTTATTGACAACCCAATGATTGTGCCGGAACTGGAGAACGGCAACCCGCTGTCCTATAACGCCACATCTGGCGGTATCGTGCGTGGAGCGTTCTGC
AGTAGCTTTATGTTAAGCGAACAATTGCGTTCCACCAGCAGCATGCAACGCGACTATTCTTTTAAAAATCCGCCCTACAATCAGCAGCACGGACACCAGCGCGGTGAGG
ACAATGGATCTGCCAGTGACTACGAGCTGTTTGAGTATCCGGGACGTTACAAAGCTGATGGGGTTGGCAAAGCATTCACACGTCATCGTCTGGAGGCGACACGTGTGGA
TGCGACCACAGGACAAGGGGCAACCAATTCCATCCATCTAATGCCAGGACACCTATTGGTTCTAAGTGATCACCCCCATGATGATTACAACATCAAGTATCATCTTCTTG
GTGTTTCTCATCAAGGCAGTCAGCCACAAGCACTCAGTGAGGAAGCTGGCAGCGGTACCACCTCCTACGGCGCAGGCTTTGAGGTGATGCCCGCACGTTTGCCTTACCGT
CCCAAGCTGATCAGGAAGCCGCTGGTAGACGGCCCACAGATTGCTCACGTGACGGGTCCAGAAGGTGAAGAGATTTACTGCGATGAACATGGCCGGGTGAAAGTCTGGT
TCCCATGGGATCGCCACGGCAAGAAAAACGACACCTCCTCCTGCTGGATACGAGTAGCTTCCAACTGGGCAGGTGCAAACTGGGGTCACATTGCGATTCCCCGTGTGGG
GCATGAGGTGGTCGTGGACTTTCTGGAAGGAGATCCAGATCAGCCTATTATTACAGGGCGGACGTATCACAACAATAACCCCAGTCCCTACCAGCTTCCGGATCACAAA
ACCAAAATGGTCATCCGCTCCGACAGCCACAAAAGCGAGGGTTTTAACGAGATCAGCTTTGAAGACGAAGCCGAGGGCGAAAAGATCTATCTGCATGCTGAAAAAGAT
CTTGACGTTCACGTCCAGAACAACCGGGCAAAGCGGGTTGACGTCACCGAAGTTGAATCAATCGGGAACAATAAATTCTCAGAAATCGGCCGAGATTTGAACACTCAGA
TTGGCGGGAATGTAGTGACTGCAATTGGCGTGGCAGACGTCCAGGTCAATACGTCCAAACCCATGAAAGCATTTGACCAAGGACCACTGCAATTGGCGTATCATGCATC
AAACTCTGAAGGCTTCGACGCTGCTCGTGGGCATTATAAGCTGATTGTTGCGCAGTCGGTTACGGAGCAGGTTGGTACCGAGTACATGTCTCGAATTGGCAATAGCTATG
GTCTGGCCGTTGGGACTACTTACCAACTCAATGTTGGTTCAAAGTCGAATATCATTGTTGGCACAGACTCTTATGAATCAGTCAAAGGTGGGCGCTATATGCATGTCGGC
CAAGAGATAGATTTGCGTTGTGGTGCATCTCGTTTCCTGATGAAACGCAACGGAGAGATTGAAATCACTGGCACGCATATAAAAATGCGTGCTGGGAAGATCAACCTCA
ATTAAAGGGGCCGAAGAATGAGGAAGTTATTGCTGACAGCAGTTGGTGGGGCAGTCCTTGCTTCTGCTATTACTGGAGTGCTCTATTTTCAATCTCACCAGCAGAATGAG
GCATGCGATCCTTCTACTGTTGGTCTGCCTGCTGATGCGCCTCCTGAGTTTGTTGCGGTCTTTAGTTACCAGTGCGCTAGCCACTACCCCAACATCTCTCGTGATGAGCGTG
CCAGTTTTGCCCTTAACTTTGGATATCCATCACTTCTTTTTGGTCTGAAGCCGGAGCAAGAGGTTGTAGATTTGATTGTTACAATAGCAAACGAAGGAGACAGGGGGTTCT
TTTCGCATGATCCAAACATTTTGGATGAACTTGTTGCCCGCTCTTTGGCTTATGAACTGGACGACTTGATCAGCTTCGTGCGCACTGCCTATGACAGCGAGAAAACCTTTG
AACAAGGAAAAGTGAGTAATGCATTCTTATCGGCACTACCTCTCTGGAACAATCCTAAAAATTCTGAGCGGGTCAAAAATTGGTGCTCCCGAAACAACTGTTCGATGACT
AGGACAGTATGTGATGAAAAAGTCGGCCTGGCCGATAAGCTTGCTCTCGCACGTGGTGAACTCAGCTATCAGCAGCTTTATGAGCAATGTCCAATAAGCTTGGACTGGAT
ACGAACGCTTCGATGGGCAGGTCTTGTGGACACGAGATGCGGCACATCCTTCCAATTTTCTACAGTACGAAGATTACCGGAAACGCTTGGCCGTGCGTGCTTCCCAGTGG
GGTTTTCTTTATCTAATCATTTTTCCTATGGGTCCTCCTCATTCCCCTCTTTCTGGCGTCAAAGACATTCTGAGAACGCCTCCCTTGGTGAGATGGCAGTTTCAGATATTCA

AAGGATGGGTAGTTATTTTGCACCGATTGAAAGCTTCCTGAATAAATATAAGCGCCTGAAAAGAATTGATGAAAAATGAAACTTACAAAAGACCAACTTGATGCGTTCA
TGTCCAAAACGCACTTTGACTTTTCAACGAAGGATGGCTGGGAAGAATTCTTGAAAGAACGGAGTGAGGACGGGACCGGACGCAGCGCCGATGGCATTCAACTGCGTGA
TATCAAGCATTACGCGGTGGGGACTGCATCTATCCATTCGATCGGGAATGCTGTCAACTCTCGTTTTACGATTAGAACATTGGGGACCGGCCTCGCCTTACCTGCTGGGGT
CGTCTACGTATATTATCCTGACGCAAATTGGGATTATCAAGTACATACTGAGGCAATAGTTCCTAGAGGAGACGTTTGGGAGTTCAATATTGATGTTGCCGGAAGTGGAA
CGCCAACACGGCCAGATTGGGGGGCACTTTACTACATCGTCCAAAAGGTCGGAGAGCTGCCCGGTAAAGGCATCGCGAAGATGGCGGGAAAAAAGTATTTGAAAGAGG
CCCTCAGTTCATCCAGTTGGCTGTTGCGACGTGGATTTAATATTCTGGCGAGCGGGGAAGCAGGGTACACGGTGTATGTCAAGGTTGGGTCTATGCCCGTAACAAACGAC
TACTTTCTGCCCCAAGAATACCTTTCTCCTTCTGATAAAAAAATGGCAAAAGACGTAATCGATGGCGTCAAGGATTACGCTAACGACCAGTTTCCCATGTCGAGTGTGAT
CAGTGATCCAACTCAAGGTTCCAGTGCAAGTCCTTTGGCCCAAGAGTCTTTTAGCCAAAGAGCGAGTAACTTTCAGAACACCGTTGATAAAGGTCTCAAAGACCGTTTCT
TTCCCAATATGCCCTTTAAGAACTAAAGCCGCACAAGATTTTTACTGTCCAGAAGCTTTGAAAGGCAGCTGGAACGCACAATCAATAAATGAGTATGTACATGCCAGCTT
CTTATCGTAAGGGAGACATAGCTTCTGGTCATGGGTGTTTTCCGCCCACCCCATCAGTTGCTGGAAGTCCCGATGTATTCGTCAATAAAATCGCTGTTGTTCGGGTTGGCG
ATGGGGTTGCTGCACATGGGTGCAAGAAGTGTCCATCACATGGAAGAGCTATGGCTGCGGGTTCGGGGACTGTCTTCGTTAATGGTCGCCCTGCATGCCGTATTGGAGAT
GGAGTAGGCTGCGGAGGTGCTGCTGCAGCAGGATCTGGAGACGTATTTTTTGATGGCTGAAGGCGCGAAGTCCAAAACCGGAGTTTTCGATCTTGGTGAAGACATGACG
CCGATCGATATCTGGCAAGGTTTGCATGCAAGTGAACCTTTATGGATCGCCTCTGCGGGCTTTGAAAGCGGTGAAGAAAACCAGAAACATATTGATGAAAGCGATCTGA
GCCTTCTGAAGAAACTGGAGAACTTTCCGGCAAAAAGATGGGCACAGATGTGTGATGGCATCGGCTGGACTCCGCTTGGCGCTGTTGCTCTGTCCTGGTGCCTGAATTCC
AACGAACAAGTCTTTAAAGCCGCGTGGTCAAATGCTGTTGAAGGTCAAGATCTTAGCGATAGTCAAAAGCGTGCATTAAAGATGGTCAGGGCAGCGGATTGATAATCAT
CGAGTTTACTGGATCACTAACAGGTAAACTACGAGGAGCTAGATATAGCTATGCCGCTTCGTGTTGAACATGCACAGGCTCCTTACGAAGGCATTCAAAAACCCTTTATG
CTGACAACAGCAACCCTTGATAGAACCAACTAAACGGCAGCTTGTGGCCCTGATTACAGGCAAGACGTCATTGAGCAGTTGCGGATGGGGTGGAGCAAAAAGGTAACG
CAGAACAGATAAGTTTGTCAGAAATCTTTGTAACTGACCGGAGTTAAAGTCATATTAGTCATTTATCAGCATGACGGCCCCAGAATTAGCTTAATAACGTCATCACTCAA
GAATGTGCTTGGATATTAAGTTGTTACACCAATAGGGCTTATATTTGAATGAGTGTAATAACACGATAAATTCCCAGATCTGAAGGTGCTCATGCTTCTCTTTCCATTTAA
AGTAAATTCGGCTCTCGCAATGATGGCAGTGCCGCCGGACGCATTTAATGGTGCTTGGCGCTCTGAAATGCAGATAGCTTGTAAAGCGCAGAGTCTAACTGCACAGGAA
ACTGCATTGGTGATTGTGGCTGAAGGGTTAGGTATAAATTGGCCAACCGATGTTGAACTCGCTATTTCAGTGTGGCGAAAAAAGGGAAAAGTTGACTTTTCAAAACCACA
AGTCACAGCTGCTTTAGAAAAAATGGGTTTTGCTGTCCACCTCCTGTAGGTCGAGGTGACCTAACTGATTTTAGGCGGCTGTAGCCACTTCGCGCCAGAGAAAGTGCTTT
CAACTCCAAGTATTCCTCCCTTAGCATGCGATTGAGAATAGAAGTTTGATTTTGAGATTGACTCGCAAAGCTATCGCCCTTCTACTCCTGTAGTCCTCTGGATAGTCATGC
ACTAGCATTCAAACCGGTTCACTCAAATGAGGTTGCTCAGAAAGCCCTTTACACCTAGAACGGATTGTTATTATTCGATATTCTGTATCATTTTCTGAACTCAAGGTAAGA
ATTATGGCGTTGATTATAGCTGGCTGGGCATTGGCCTCTTTCATTTACTTCAAAGTCAGCTTTGATGCTGCGTATTATGGCGATATTACCATTGTCGCATCGCTATTGATGG
CGATGGTATTGAGTGTTTTGGTATTCCTGATGTCCGGAATTTACGCTGGAAAGTTTTACAAATTTAGTGCTGCTTTTAATCAAGAACATGTTAGTTTATCCGCTTTATTAAA
GGCAAGCATCCTGCCTTTTTTAGGTTCAGTTGTTACATACGCAGTTATTTCTCGCGTAATTATTTATGGATTCTCCAAAAAACTTAATAAATATATTGATGTACTAAACTAT
AGAGGAGACATTAAAGAAGAATTTGTTTTGTTTTGTATCATTATAACTATACCAATTTTATTGCATTATTTGATGTGCATATTTGGAATACTTCGTAGTAATGATACCAATC
CCACATAAAACACCGTCAAAAGCCTGTTGACCTATCCAACATTGTGGATGAGTCACTTCCCTATGGTAAACCTCCTTCATTGGGGTGATCCAGCCGTATAGTTAGGAAGC
TGAAGGATTGGTCTTGGCCAGACCGTTGGAGCTATGGAGCGGCTCTCTTCTTCCAGCTGCTTCAGTTCGGAAGGCAACAAGCCGCAATACCTCTTACGCCAGCGAAAGCA
TGTCTGCTGGCAGATGCTGTATTGACCCAGCAATTCCCTGCTCAGCTAAAGTAGATCAACCCCTGGTGGAACTGATCCAATGTCAGCTAATGGCCCACAACGAGCTCTCA
ATTACTTTCGCCTCTGCAAAGCAGTGAATGCTTAGCTCGTCCTGTGTTTTGCAACTCACAGGTTGTTTGCAATCTAAGGATCGGGGGAACTTTATCAGTACCCGCGTTAGC
GCGCATATCCGAATAACTTCAACGCGCTCAAGAGCTAGGCACTTCAAGATTCCAATTGCCAACTTTGCTTAGGTATCCTTAAAGCCGACTGGAATAGTAAAACTTATAAA
TTAAGTACATGTAAGATAAATTCATTTTAACTATTGTAATATTTTACCATTTACAACTTCACGCACTATATGAATATTTAGTTTCCTGTTAATTTATTTAAAGAAAGCTTCG
ATTTGCGGTGATTTTATTTGTATCTCTTCTAGAATTATGACTGATGAATCGTAACGAATGGACAGTAAATGGTGTTCCTGCGCTTTCGACAGTCAGTTTTAGTTCTAGGATT
ATTGGTATGGTTTCATCCTGACGTATTTGCGCAGGCCGTGATTGATGATGATAATCCCTTGGGTCCAAACCAGACTGTAGGCGGGCCCGGTTCCACACATGGGGATACCT
GGAATGTTGGTGGCACACTCGTCGTTGGTGATAACGGTGCTGGCACTTTAACAGTTCAACAAGGTGGGGAGCTAACCAGTGGTTTAAGTACGTTAGGCACAAATGGCTCT
TCAGGAACTGTAACAGTTACTGGCAGTGGTTCTAATTGGACTGCAAGTAGCGGTATTCGTGTCGGGCGCGGAGGGACCGGTATCTTCAATATTGAAGATGGTGCCACTGT
TATCAGTACACAGGGCCGTGTTGGGGTGATTGGATCACAAGGCACAGCTACTATTACTGGAACTGGGTCTGCATGGACCATTAATACGGACTTGATTGTTGGGAATGGTG
GAACAGGTAATCTTATCGTTGAGCAAGGCGGAGCGCTTATTACTGGTGAAAGCTCTGTTGGCGAGAACAATGGTGGCAGTGGTACGGGCGTTGTAACCATTCGAGATGC
AGGATCAACTTGGGATCTCAGTGATCGCTTGCGGGTTGGTATAGCTGGCACTGGCACCGTAAATGTTGAAAATGGCGCTAGCGTTACAAGTACAGAGGGTCAACTCGGC
GTATCGAACAACTTAGGTACGGTAAATGTGACTGGAGCAGGGTCTACCTGGGACATTTCTGATGAGCTTGTTGTTGGAGCGACGGGGGCTGCTGGTGGCAACTTGAATAT
CTCTGATGGTGGCTCAGTAACCTCGCAGGGCTCCATTGTTTCTGAGTTTGCAACCAGCAATGGAGCGGTAACAATCACTGGTTCAGATTCAATTTGGACTGTTTCAGGAA
ACATAATCATTGGCCGTCAGGAAGATGCAGTTGTTACCGTTTCCGATGGTGGGACGCTTGATACCCGTGGGAGCAATACAATTGGCCTCTTTAACGGTACGGGTTCTGTT
GTTGTTACGGGTGCAGGTTCGAATTGGACGGCTGCAAACAATATTGTGGCAGGCCGTTCTGCCAGTGCGGTGATCACACTTTCTGAAGGTGGAGTTCTCGAGATTACTGG
CGGAAATCGAAGGTTGGTACTCGGTCAAGAGGCTGGGGCTCGTGGCACACTAAATATAGGTGCCGCTACTGGTGAAGTGGCAACGGCAGCTGGCACATTCATTGGTAGC
CAAGTGCGCCTGACCGCTGGCCGGGATGTGATCGTCTTCAATCATACAAGCACAGATTACACGTTTAGTGCTGGGCTCTTTGGGGATGGTGATGTTCTTATTGAGAACGG
CACAACCACTTTGACAGGGAATAACAGCTCTTATTCAGGAACCAACACGATCACGGGTGGCACGCTTCGCGTTACAGATATCGATCAACTCGATGATAGCTCTACTGTCC
TTAATGGTGGAAGGTTTTTATATGCCGACACATCAAATGGCTCATACTCTGCACTAACGATTGACACTAGCGGCGGAACAATTGAGGTCTCGGAGAGCTCTGCAACTTTT
ACAATCTCAAGCGGGCTGACAGGAACAGGAAGTATTACCAAGGACGGAGCTGGCATTCTTCGTTTAACTGGCACAAATGATTTTAGCGGCGGCACAACCATAACTGACG
GAAGTGTGCAGGTTGATGATATCAGCAGTTTGGGCACGGGCACGACGACTGTTAGTGGTGGAACGCTGAGCTACAGCGGAACAGCGAGTGCGACTTACTCCAATCTTGT
CCTGAGTGCTGCAGGCGGCGCCTTGGAAGTTACTGAGAGTTCAGGTGTTCTTGAGGTTAGCAGCAACCTTTCGGGGACGGGAAGCCTAACCAAAGATGGAAGCGGAACG
CTGGTGCTTTCTGGAACGAACACCTATAGCGGTGGAACGTTTATTACAGGTGGAGTACTTCAAGTCGATGATCTGGACCGCCTGGGAACCGGGGAAACCACACTCAACG
GTGGATCTCTTAATTATACAGGGACAGGTGCTGTCACGCTCTCTGATTTGGTTTTAGGTGGCGCTGACAACACGCTTGATGTCAGCAATAGTGCGGCGAGTATTACGGTTT
CGAGCGGTCTTTCGGGTACGGGCAGCTTAACCAAGAGTGGAGCTGGTACTCTCATCCTTACCGGGACCAACACTTACAGTGGTGGAACAACTGTTTCAGCAGGAACATTG
CAAGGTGATAGCTCTAACATTCAGGGAGATATGTTGATTGCCAGTGGGGCTACTGTGGCTCTTACTCAAGCGACAAACGGAACCTATAGCGGAGACCTGTTCGGTTCTGG
TGCTCTGACTAAGACCGGATCTGGTATTTTGACCCTATCAGGTAATAATTCCGGTTACTCTGGCGTGATCACCGTCGATAGCGGCGGGCTGTCGATCTCGGCTGCCAACCA
GTTGGGTACTGGAACCACTCTGCTTAATGGTACATCACTCAGCTATACTGCGGGCACCGACGCTACTCTAACCACCGTAACTTTGGGAGCTGGTGCCAGTACATTGGAGG
TGGCCAATGCTGCAGCGGCTGTGACGCTTTCAAGGGATCTATCGGGTGCGGGCAGTTTAACCAAGACCGGGGCGGGTACGCTTATCTTATCCGGCACCAATACCTATAGC
GGTGGGACGACTGTTTCTGCAGGGACATTGCAAGTTTCACAGGTTGCTAACTTGGGAACAGGAATCACCACTCTTAATGGCGCGACGCTGGCCTATAGTGGCGTAGGAA
CGGACAGCTATGGCTCGCTCGCAGTTGGAACGGCTGGAGCTACTGTCGATGTGCTAGATCGCAGTGGAGCTCTGCAGTTTACGAACACTGTGACCAGCACAGGAACCTTC
ACCAAGACTGGTGTCGGATCACTAGCCTTCAATAGCACTAGCAACTCATTTGTGGCAGGTCTTACAGTTCAGGAAGGAACGTTGCTGATTGGCGATGCGACGGACTTTAG
CACTGCATCTGTTGAGTCAACTTTTACAGTTGATACAGACGCTCGCCTTGGTGGGCACGGAACTGCTAGGGGGAGTGTTACACTCGACGGTGGCATCATCGCTCCCGGCG

CCTCCTTTGGTACCATCACGCTCGGCAGTCTGGATGCAACAGACGCTGGGGGGGCGCTGGAAATGGAGATCGGGTCAGATCCAGTATCTGGAACGGTAAATACAGATCA
GATCCAAATCCTCGCGGGCGGAAGCGTGGATCTGAATGGCATTGCGCTCAATCTTGTTAATGTCATGGGTCCTACACCGGTCAAGATTGGCGATGAGTTCGTTATTATTG
ACAATCTCTCCACTCCGGGGACAGAGCCTCCGACTGGCACATTCAGTTCAGTCAACCACGACATGTTCATGCTCGATGTCAACGTTATCTATGGCATTCCAGGAGACGCA
AATGATGTCGCGATCCAGTTCATCCGAAGCTCCAACAACTTCAATTTTCTGGCAGATACACGCAACGAGCATGCTGTCGCACAGGCATTAGACACCATCAATGAAAGCG
GCGACCTGTTCACACAGCTTGCATCGATGAATAACGCGCAAATTAAAGCTCTTCTTGATCCCCTAGCCGGAGAAATCAACGCCTCCACCCGGGCAATCCTTCTCGCTGAC
AGCCGCTTCCTACGGCAAGCTGTGAATGACCGGCTCTACAAGGCTCATAGTGCTTTTGCATCAAAAATGAATGACGACATGGTTTCCCGCACCGTCGGCGACTTTGAGAT
GTGGGGACAGGTCTTTGGCAGCTGGGGGGAATTCTTCTCAGACGGAAATGCCCGCACCACAGGCAGAAGCGTTGGCGGTGTTCTGGTTGGTGTTGATGCGACCTACTTTG
ACATGGTCCAGCTGGGTGTGGTCGCAGGTTTCGGTAGAACCTCCATCGACATCAGCTCCATCGATGCTTCATCTGACAGTGACAGCTACAGTGTCGGCGGCTACGCCAGT
GCCAATCCAATCGATAACGTCAAACTGCGCGGAGGTGTTGCCGGTACATGGTACAACATTGCAACTGATCGCGATGTAAGCCTCTTCCCGGAAGAAACACTGTCGTCTGA
CCAGTTCGCCGGTTCTGTGCAATTGTTTGGTGAGCTTGGATACACGTTAGAAAGCACCTACGGCAATCTGGAGCCTTTCGCAGGAGTTGCTTATGCTTACCTCTCCGGCCC
CGGCTTCAATGAAACCGGCGGCGCAGCAGCACTCAAAGCAGATGGGCAATCCAACTCCCTCACCTTCCTGACGCTGGGAACACGCGGTGAGTTTGATCTATCCAACTAC
CTTTATGGGGATACAAGCTTTGTTGGCTTAGCAGCATGGCAATACTCTTCTGCTGATACGCTCGATTACACCCAGAGCTTCATCGACAGCGGCAATCCCTTCACCGTCGCC
AGCGTCCCAATCGCCCGCAACGTGGCCCTGATCGAAACAGGTCTCCAGTCCAACCTCAGTGATCACATCCAACTCAACATCAACTACTACGGCCTCCTCGCCGACCAAGC
CGTAGACCACGGCGCTAACGTAAGGCTGTTCGTGCAGTTCTAAGTTGAGAAAAAAAGTACGCTCAGAAGTATTGCCACATATTTAAAACTGCCTGATCTGTTTGCCACTT
TGAGTTAAGCCGAAGTGGCGCCGTCAAGTTCCTTCATTTTGCTCAAAGTTCGCCTGCTCGGTATCAGCACGACGATTTGGCTGTGGGAAGAAAATCAATTATAGCTCAAT
TTTATAATGGCTTAGCTAATTGGTATTCAATCCATGATGGAACTGCTCGAATGTCCGCTTCTGGCCCTCTGCTAATACATCAAAGCGTTTAATCCATCTCTTAAGCTTCTTC
ATAACTCCCAAAACGCAATACACTCTCTAAACCAGTCAGTTAAAAGATTCATTTTCGTTTCGCTTGACCCATTTGGATAAAATGGTACTCTCGACCCAATTTAAGAAGTGA
GTCGAGCATGATTGGGAAGAGTGACACCTATAGCCGTACATATACGGCGTCCGAAGTTGCCAAGATCGTTGGGATGAATATCGAGACGTTGCGGGTCTGGAGGCGGCGG
AAGCAATTGGTGCTTGAGGGCGAGCGCCAAGGCTGGACCCGATATACCTTTAATGATCTGATGGCCGTTGCCACCTACCACAACATGGTACGCTCTCACTGCACAAATAC
GATTGCAGAGCTGATTGCCGGGACTGTCTCACAGCAGCTCGCCAACAAATGGCCTGGTATTCTTGGCACAGAAAGTAGTGTTTTCTGCCTTTGCAACAATGCAATGGATG
AGGAAAATATTCAGATAGAGATGCATCTCGGACTGGAAGAGCTTCTGGAGGCGATCACTCAGATTATGCGAAGCTCAAAACGTCAGCCTGCCTATCATATTGTCGATTGT
GGTGCTTTGTTTTTGAAAGTGCTGGTGTCTTGCACAAAGGTTCGTGAGGAGAGCGAGTAACCGCTTTGGCTCATTAAATCACAGGCCCATGCTGTCGGCATGAAAACATC
TATAGCCTGAAGATCGAGAGGGGATCGATATGGCGAGAGCGCAGTTGAACTTGAGTATACAGCCGTATCGCATGCTGAAGAGAGCTGATGCGGCTTCCTATTGCGGACT
AAGTGCAACCAGCTTCCTGGCCAATTGCCCGGTTCCCGCTGTTCTTCTTCCTGGAGGGCCCAAGGTCTGGGACGTCAAGGATCTGGATCGCTGGATTGAAGGCCTAAAAG
AAAACAGCCATCCATCTGATGATGATCTCCTCAATCAGCTGGGAGCTTGATATGACGATAATCCGGGTCAAAGGCTTCAAGATCTTCAAAGACCAGCATGGCAAGCTGC
GATGTTATCATCGCAAGACTGGAGAGGCGATCAATCTGGAAAAAGCGCCAATTGGCAGTGCTGCCTTTCTGGCGGAATGCGAGCGACTGACAAGCTTGATGTCAGTGGC
TGATACCAGCAGACCGGGCACGCTCGGCAACCTGATTGACAAGTATCGCCCCAGCTCTACCTTTCTTGAACTAGCCGATCGCACCAAGTCGGACTACCAGAAGGTCTTTG
ATTATCTGGCGCCAATCAACAACACACCCCTCACCCGCTTTACCTCACCCTTCATAGTCAAACTCAGAGATAAAGCAGAGGAGCAACGTAAGCGGCGTTTTGCTAACTAT
GTAAAGCAAGTGCTCTCTGTTCTGTTTTCATGGGGCAAAGAGCGTGGGTTCGTCAAAGAAAACCCTGCGCTTGGGGTTAAATCCATTCGCAACCCAAGAACGCTCCAATC
GCCAACAGGCCCTGGACGGATGACGAGCGCAGGATCGTTTTAAAACACGCCCGGCTCAACTGGTACCTGCCCTTACCTTGATGATGTATACGGGACTTGGTCCGCAAG
ACGCGCTTACGCTTTCCAAGCAGCAATACCAAAATGGTTTCATCTCAACACGGCGAGCCAAAACAGGGGCTCCGGTCTTTTGGCCTGTCATTGCACCTTTGAAAGCAGCT
CTTGATGAAGCCCCTACCCACTCGGCTCCAACACTTTGTGCAAACTCCAGAGGTCAACCATGGACTGTCTCGGGGTTCAGAGCATCGTGGAGAAAGTTGCGCATCCAATT
AGAAGACAAGAGCCTGATTGAGCCGGGCCTTACTCTTTATGGTCTGCCGCATACGGTGGCAACCATTTTGCCGGGAGATTGGGCTGAATGACAGAGACCTTGCCGCTGCG
CTGGGACATGAGACAGAAGCCATGGCCCGGCTTTATGCAAAAGGAGCTGATTTGCGCAAAAGCATGTCGGATATTGCTGTCCGATTTGAGAAGGAAATCACTCTGCGCG
GGGAAGAGTTCAGCAAGCAGTTGGAGCAAAAGTGTCAAACCATGCCGAAAAAAGTGTCAAACCTTGAGAAAGCAGAGCGAATACTCGGTTTATTATCTAGTAATTTCAA
GGAGATAAGTGGTGCCCGGAGGCGGATTCGAACCACCGACACGCGGATTTTCAATCCGCTGCTCTACCAACTGAGCTATCCGGGCACATTCTGCTAGGGCGGTTTGTTTC
CCGTCCTGCGAAGGTGAGCGTCTTATAGGGTGCCAGATTTCTCCTGTCCAGCGGGCAAATGCATTTTTTACAAAGAAATCTGTTTTCTCCACTGCCCGCTCAACATTTTCT
CTTAAGCGCCTTAAATCATGAGGGTTTTGTCCCAGTTCGAACCTCACCAAGGCTCCGAATATCTCAGATAGGCGCCCCTTTTTCCGAAATATCCCAAAACAGCGCCACCA
TATTCCGCATGGCATCGCGGGCAACCGGTTTGAGAATATGTTCATTGGGCGCGTGTTGTGAGCATCCAGCATAAGAGTGTGGAACCCACACAGTTGGCAGGCCCAGTGTT
TCTGCAAAGCATTCATTAGGCAGTGAGCCAGCGAGGTTTGGCAAAAGGTGGGGTGTCTTCCCGCTCGTTTCTTCCAGGGAAGCCATCACCATGCGCACCCAAGCGTTATC
TGGCGAAAGTCGCGTTGCATGAAAGAAGCCGCGATCTTCCTGTCTTATCTCGACATTATCAAATCCATTAGCATCCAGATGCGCGCGCAAAGATGGTAATAGACCCTGCA
CATCTGTTCCCACCACGTAGCGCAGCTGGCAAACTGCGAAGGCATAACCATCAATCGCGTTTACTGGTGCTTCTGGCACGCCGCTCTTCATTGCAAGCACGGCAAAGGAG
TTCCAGCCAAACACACGCTCGGCGGGGGTGAGGCCCTCCTCTCCCCAATTCTCGTTTATGGCCGGGCCTTGACTATCACCAACAGGAAGATCTCTCAGAGCTTCCCTAAT
GTCTGGTGTCAGGCTGGTTGGTCGCCAACTCGGAATCTGGACCTGCCCGCGTTTGTCTGCAATGCATGAGATAGCGTTGGCGAGAATAATGGCGGGATCGGCAAGCAGC
CCGCCCCAGTTGCCCGAGTGGTTTGCGCCATCGCGCAGGTTCACTTGTAGGGTGAACGTTTGCGCCCCACGTGCACCTGTGAAGACTGTTGGCGTATCCCATTGCAAGCG
AGGTCCATCAGAGGCGATGAGCACATCTGCGGCCAGCATTTTTTTGTGTTCATCGAAAAACTCTTGCAGACCAGCAGAGCCGGTTTCCTCAGAGGTCTCGATGATCACCT
TGCAATTGAAGCCAAGTTGCTTCTTCTCCTGCAAAAGACATCCCAGAGCCTGAAGGTTGATGAGGTGCTGACACTTATTGTCGGCTGTGCCGCGCCCGTAAATGCGTTCG
CCGTCTTCTTCCAGAACGAATGGATCTCTTCCACCCATCCAGCTACCGGCCTGTCCGCGCACGACATCTCCGTGGCCGTACGTCAGAACTGTCGGGAAACTGGGATCTTC
GATGCGTTCACCGAACAGGAAAGGACCACCAATCGGGTTCGGATTGTCGAAGATGTGTGTTTTGAAACCAAGACGTTCGAGAAGAGGTCGCATCTCCTCTTTCAGGTAA
CGCATCAATTGGTGTTTTTGGTAGGGTTCCTGGCTTTCCGTCGGAATTGCGATTAATCGGGCCAGGTCAGTGAGGAATGCGCCGCTGTCGAAGTATGTTTCCGCCCGTTGG
ATGGCTTCATTACGACTGGACATTCTCATTCTCCCTTTCAAACTAGGTGGGCGCAGTTAATCACTTAAGGGGGTTCCAATCCACTCTGTTTGTTTCAGCAGCTTTAATGCA
CTCTGAAGACCGGCACCCGTTGCAATGCACATGCCGGGCAGGATCATCCACTCCAACGCCCAAGCTGAGCCTGAGCGCTCTTGCTCATGCACTAAGGCGCTATGCATTGC
AGAGACTTGAGTGCTGTTGTAACGAGCAAGCGTCACCAACAGTTCAGCGTTGATCGGGTTGTGTTTGTGGGCCATGGCGGAGGAGCGACCGCCACCGGAAAGCTTGATT
TCCTCAAGGCCCTGTTGTGCCATCAAGGTGATGTCTGTACCAATTTTGCCAAGGCTGCCTGAGATCATAGAGAGAACGTTTGCGTAGTCCACCATGTTATCTCGCTGCGTG
TGCCATGAGTGCTCAGGAACCCTGAGGTCCAGCTCTGCAGCCATGTGCGCAGCAATAGCAACTGCGTGTTCGCCCTGAGCATGACTGGTTCCGACAGGTCCGCCAAACTG
GAGCACAGCCACCTCAGCCGCAACATCGTTTAGCCGCTGATAGTGCCGTTTGAGCGGCTCAATCCATGTAGAAATGCGGTCGCCTGCCGTTATTGGCAGCGCAGCTTGCA
TGCGTGTTCTGCCCATCAGCGGATAATCAGCATTGGCTGATTTCACGGTTTCAAGTGCAGCGATAACTGCAGACAAACGCTTTTGTAAAATCTGCGAGGTTTTTCGCAGA
GCTAATGCCAGTGCGGTGTCCATCACATCCTGACTGGTGGATGTTTTGTGGCACTCAGCAGCATATGGCTCGCCAACGCGCATACGCAGTTGCTTAATGAAAGCCGGAAT
TGGCAGGCCATCTTTTGCGACACCTGCCTTGATCCGCGTGATATCGGGCGCAAACCCTTTACAGCCACAAAGCACAGCGTCTTTTGCTTCTGCAGAAACAAGAGAATGTG
CAAACAAGGCCTTGGTGAGAGCCACTTCATACGCCAGATAGTGATCCAGAGTTGTTTCAAGAGAAAAGGCGCGGTGCACGTCCTCATCTCCAAACAGTCCATTCCACAT
AGGGTTGGTGAAGAGGCTCACACTCACATTAGTCCCTTTCCAGCGCGATCGCGATACCCTGACCGACACCAATACACATGGTGGCCAGTGCACGCTTTTTGCCTTGCAGG
CTCAGTTCCAACGCGGCCGTTCCGGTGATGCGTGCGCCAGACATGCCCAATGGGTGACCCAATGCAATCGCGCCGCCATTCGGGTTGATATGAGCTGCACCTGCATCCAA
CCCCAGTTCGCGCAGAACTGCAATCCCCTGACTTGCAAAGGCTTCATTGAGCTCAATGACGTCAAAATCAGAAGGTTGCAAGTTTAAGCGCGTGCAAAGTTTTTGCGCAG

CGGGCGCTGGACCAATGCCCATAATCCGCGGCGCAACACCAGCGGTGGCACCGCCAAGCACACGCGCAATCGGGGTCAGACCATATTTCTTGACCGCTTCAGCAGACGC
AACAATCAGCGCTGCAGCGCCGTCGTTCACACCTGAGGCATTGCCAGCAGTTACGCTGCCATTTTCACGGAATGGTGTCGGTAACTTTGCAAGTTTCTTAGCTGTCGTCCC
TGCTCGTGGATGCTCATCAGCGGACACAATGATCGGGTCGCCTTTGCGCTGCGGGATTACAACTTCAATGATCTCCTTGGCGAGGCGGCCATTTTCTTGCGCAGCAGCTG
CCTTTGCCTGTGAGCCCGCGGCAAAAGCATCCTGATCTGCCCGAGAAACGCCAAAGTCTTCAGCAACGTTTTCGCCAGTCTCAGGCATACTGTCGATACCGAACTGCTTC
TTCATCTTAGGGTTCACAAAGCGCCAGCCAATGGTGGTGTCAAACACTTCGTTAGCGCGAGAAAAAGGTGTGGTTGCCTTTGGCATCACAAAGGGCGCGCGGGACATGG
ATTCCACACCGCCAGCAATCACAAGGTCGCATTCACCTGTCTTGATGGCACGAGACGCCGCAATGACAGCATCCATACCAGAACCGCACAAACGGTTCAGCGTTGTCCCT
GGCACAGCCTCAGGCAAACCTGCCAGCAGCGCGACCATGCGAGCCACGTTGCGGTTGTCTTCACCAGCCTGATTGGCACAGCCGTAATACACTTCATCAACGGCTTCCCA
ATCAACTGAAGGGTTCCGCTCCATCAACGCTTTAATTGGCAATGCGCCAAGATCGTCCGCCCGAACAGAAGACAGCGCGCCCCCATAACGACCAATCGGTGTTCTAATAT
AATCGCAAATATATGCCTCAGACATCAGGCCGTCTCCACTCTGTTACCCGCATGCGCCGCCGCTGTTCGCGCTTTTAATGCGCGCAGAACAGACAGTTCCTCCTCAGTCGG
CGCAGGGGTCTCCTGCAAATCATCACTGAAACGTACAGCCCAAGACACGGTTTGTTGCACCCGTTCTTTGGTCGTACCCGGATGAAGCGAGACCACTGTCATCTCTTTGG
TAACAGGATCAGGTTTCCAGATAGCAAGGTCCGTAATCAGCAGGGTTGGGCCTTTTGTGGCGATGCCCATGCGAGTGCGATGGTCCTCTCCTTCTCCGTGTCCAAAGGAG
GTGAAAAAGTCCAGTTTCTCTACAAAACCACGCTTGCTTTGCTTCATGGTGATGTAAATTTCCTGAGAGGACGAACAGATCTCAGGCGCTCCACCACCACCCGGCAATCT
TACCGTTGGATTGTCATAATCTCCCACAACAGTGGTGTTGATGTTGGCGAAACGATCAATCTGCGCTGCCCCCAGAAAACCGACAGTGATCTTGCCGCCTTGTAGCCAAT
AGCGGAACATCTCGGTTACTGAGACTGTCGTTAGCGCTGTATCACAGAGTTCGCCGTCACCAATGGAAAGCGGCAGAACGTCCGGCGCAGTGCCAATCGTTCCACTCTCA
TAAATCAGTGTAATATCAGGCGCATGGGTCAATCGTGCTATGTTGCAGGCAGCTGATGGCGGACCAATGCCCACGAAGCAGACATCCGTGTTGCTTAAAGCACGGGATG
CGGCAATTGCCATCATCTCATCCGGGGTAAAATCTTGTGTTGTCATGATGTTACCTCCAACCCTTTGATGCGGTCAGCAAACACCTCTGGTCCCTGTTTCAGCACGTGTTC
TTCCATCCATGCCTGAAACGCCTCACGATCTGCTGCGATCGGATCCCAGTCGATATACGTCTTATTGTCTCTGCCGTAGTAGCCTTGCGCATAAGACGGATGTGCCCCGCC
GCGCACTTCGCAAATAGCGGAGACGGTCCAGTGTGGCAGGATGCACGCGTTCATATGCGTATCCAGCTCATCAACCACTTCCTCAACCGTCACGATGGAATGCTTGGCGG
CCAGTACAGCCTCCTTCTGAACACCCACAATTCCTTCGATCAGCACATTGCCTTTGCGGTCTGCTCTAAGAGCGTGGATGAAAGCCACGTCAGGCCGAATGGATGGAACA
GCGGCCAGTTCTTCCCCCGTAAAGGGACAATTGATTTTCTTGATATTCGGATTTACGCGCTGAAGCTCTGTGCCCATATAACCGCGAAACACAGCGCATGGTAGGCCAGC
AGCCCCTGCTTCATAAGCATTGGCAAGAGCGGCATGGGAATGCTCTTCAATCTCGATGGACCGCGGCCAGCCGTTTTCGATTGAATCACGAACGCGCCGCAACAACCCA
ACACCAGGATTGCCTGCGTAAGAGAAGATCAATTTCCTGACCAACCCCATGCCTATGAGCTGATCATAAATAAGATCAGGCGTCATGCGGATGACAGTGAGTTCATCAA
TGCCTTGCCGAATGACTTCATGGGCAGCGGCGTGGGGAATGAGATGCGTAAAACCCTCAAAAGCGGCAGTGTTGCCGCTTTTGAGGTATTTCTTGACCGCATCTTTTAAA
GACAGGAATTCAGCCATATTTATCAATCAGATATCAAAGAAGACGGTTTCGTTGTCACCCTGAATGTAAATGTCAAAGGTGTAGGTATCGCCGGATTGCTTCGCGATCAA
CGTTGAAATGCGCACGCTATGTTCAACACGACCAAGCACTGGATCAGCTGCATTGGCTTCTTCCTCTTCCGGGAAGTACATGCGCGTATTCAGACCAAGGTTGATCCCAC
GCGCAACGATCCAGAAGTTGATATGTGGGGCTTGCGGCCTGCCATCCGGAAACGGTACAGCGCCCGGCTTCACAGTTTCAAAGCAGAACTCGCCACTATCCAGATTGAC
CGGCTGGCGTCCCCATCCGGTGAAGTTGGGGTCAGCTTCGCCGCCCGTCTCCGATCGACTGTTGAACAGACCTTCGCTATCCGCCTGCCAGATTTCCAACATAGCGTCTTT
CAAAGGTGTTCCGGTTCCATCAAACACACGGCCTTTGATCGTGATGCGCTGGCCTTTGGTTTGCTCATTCACCATGCGTGCGCCAAGGTCTTCCTTGAAGATCCCGGTGTT
GCCGATGAAGTTCGGTGTACAGCCGATATGAACGTACGGACCAGCCGTTTGGCTTGGGCTTTCTTTCAGATACTCCAGCTTTGTCATATCAATTGCCCTCCATACGGTTCT
CAAACATGGTTGAGCGGCGGCCACGCAGGACGATATCAAAGCGGTAAGCGCGCGCGTCCATTGGCACAGAATGGTTCATGTCCAGTTTGGCAACCAGCATCTCAATCGC
AGCCGGATCATTGATTGTCCTCACAATCGGACACAGCGGAATGTGCGGATCACCCTCGAAATACATCTGTGTGATCAGACGCTGCGCAAAGCCGGAACCAAAGAGAGAA
AAATGAATGTGGGACGGTCGCCAGTCATTCACGCCATTTGGCCACGGATACGGGCCTGGTCTCACAGTTCGAAAAGCGTAATATCCATTCTCATCGGTTATCACTCGGCC
ACATCCACCAAAGTTCGGGTCAAGCGGAGCAAGATAGCCATCTTTCTTATGGCGGTAGCGCCCACCAGCGTTCGCCTGCCAGACTTCTACCAGCGCATTCGGAACTGGTT
TGGCTTCTTCATCCAGAACGTGGCCATGGACAATGATGCGCTCTCCCTGCGCACTTTCACCCGGTTTGGCATAGTTCTGCACCAAATCGTGATCCAGCTCACCCAGCGTGT
TGTGGCCGAAGACTGGCCCGGTGGTTTCGCTGATTGTATTCCGAGAGACAGAAGTGCCTTTTGCGGAGACCGCAACACACTAGTTTTATAGTTGGGCGACAATGCCGGT
GGATGCCAGTTCCGGTCTCGTGGAAAGAACGGACCAGCATTGGTCGTTGGCTTGCTCATCTCACTCCCCTGACGCATCAAGCTGCGCGAAAACATTTTTGACGACTTTGA
TGGCGCTGTTTGCGGCTGGTACACCGGCATAAATTGCGGTGTGCAGCAGAGCTTCGCTTATATCTTCCGCGTGTTGCCCCGGTGTTTCTTGCGGCCCGCACATGCATGGCCA
CCTCATCTAGGTGCCCTAACGCAGCGAGAAGCGCAATAGTTACAATGGACCGTTCTCGCAATGTCCAGTTGTGCCGCGACCACACGTGCCCCCAAGCCGCTTCAGTGATA
AGCTCCTGAAAAGGCGCGTCAAAACTGTTCTTTTGAACCTCAGCAGCATTCACGTGGGTGTCGCCCAACACCCTGCGCCGTGTCTGCATGCCTTGTTTGTAAATTTCGCTT
TGTCGCAGCTCTCGCTGATCAATCTTGTTCGGCCATCGAGAAATTGCCTGCCGAACCAGATGCGCCAGATACTCAGGCTGCTCCAAACTAGGCAGGTGTCCGCAGTTTCC
CACCCGCTCATATTTAGCGTTCGGAATGAGGTTCGCCATCTCCTCAACCAGTTCCGGCGGTGTGGACTGGTCGTCTTCACCAGCAACACAAAGAACAGGCACGTCGATCT
GCGCGGCCTGACAGGTCAGATCAGCATCCCGAATAGCTGCACATGTGCCGATGTATCCGTCGAAAGGTGTTCGGCACAAAAGATTGCGATATCCTGCAACCGCTTCTGG
AAAATCTGCCTGAAACCCCGCTGAGAACCAACGAGACATCACTGTATCAGCGATCCCCTTTATCCCGTTGGCCTCAATCGTGTTGATACGCTCTTGCCACGTCTGTGCGTC
ACCGATGCGCGGTGCGGTGTCACACAGGATCAGTCCACTCACCAGATCAGGCCGAGCGGCCTGCAGTGCCATTGCGATCATGCCCCCTACAGATACACCGCAAATCACT
GCATTCTTGATGGACTGGCTGTCCATGAGCGCTGCGAGGTCACGCATATGCAAGTCAATCGAATACCGCTCATTGCCAACACCAGACAGACCATGCCCACGTTTATCATA
GGTTAGAACTTCGCCTTGATGCCCGAGATGAATCAGAAATTCATCCCAAATGCGAAAGTCTGTTCCCAATGAATTAATCAGAACAAGCGCTGGTGCTGCTGACTCTGAGT
TGGTCTTACTGAAGTGAAGCGTAGTAGAGTTTAATTCTATGAAGTGCATGCGCATCCCCCTAAACGGATCAATTGCATGTCTGTTTCAAAATGAATAATGATATTATCCAG
ATTAATCATATTGAGATTGTTATGAATAACTTGCGCAGATTAAAGCTCAGGCATTTAGAAGCATTTGTTGAAGTGTCTCGCCAAAACAGCGTGAGCCGCGCTGCGGAGGT
GCTTCATCTCACACAGCCAGCTGTCACCCGTACTATTCGCGAATTGGAAGAAATATGCGGGCAGCCTCTGATTGAAAAAGACGGGCGCGGAATCCGCATCACCCATTAC
GGGGAGGTGTTCCTGCGTCATGCAGGTACCAGTCTGGCCGCTGCCCGTAACGGCATCAATGCATTATCTCAGCTCAACTATAGCGACGGCCCGATGGTCCGCATTGGTGC
GCTGCCAACAGTAGCTGCATCACTCATACCGCAAGCGGTAAAAGCCTTCCTTGATACGGGGCTACGAAATCGGCTGCACATCGTAACAGGTGAAAATCACGTCCTGCTTG
ATCAACTCAGAAACGGTGAACTGGATCTGGTCATGGGGCGTCTGCCTGCACCTGAAAATATGCAGGGATTGGTGTTCGAACCTCATTATCGCGAACGTGTCGTTTTTGCC
GTCGATAGAAACCATCCACTTGCTAGCAAACCCCAGATCAGCGTGAGCGAACTGGATGCTTATCCGGTGCTAATGCCTTCTGTGGCCTCAATTATTCGTCCGTTCGTCGAT
CGACTTTTTATTGAACACGGCATTCCTGAACCAGAGCAAGCAATTGAGACCGTCTCTGACTCGTTTGGCAGGGCCTTCACACAGGAATACAACGCAATCTGGATCATTTC
ACGGGGTGTCATTGCCCGCGAGATTGAATGCGGCCAGTTCGTGGAGTTGCCGATTGATACCCGTGCGACGGTCGGTTCTGTTGGCCTCAACATGCGCGCAGGCGAAACGC
TGGATACAGCCGCCCAGTGCTTTGCCGAGATCCTTATGGGACTGAGCGAGAAGGAAGAAGCACGCGCATCAGCACACTGAGCCTGCGCGAATTCGGGATGCATATTGAT
ACCCAATACGGCATGAAAAGCACGGTATTTTTCATTATTCATTTTGAAAAGTATATGTAATTCTCCGATTTCACAGTGGTGATTCCTAAGAATTCCCTTGGCGGATCATCC
AGATCGTGATTAAGAATTTGACGGCCTTAGGGAGATGAGACTGCCCCAAGTAGCAGGATCAGGAAGGCTTGTTTGGTTTCGACCATTCTTCAGGAGGGTCACATACTTAT
GTTTGGTAATTTCAAAGGCTTCTGCGCTTCTACAGCACTTTTGCTTGCCAGTGTTTCCATGGCAGGCGCAGCAGAATACACATTTAAGCTCCACCACTTTCTAAGCCCTAA
AGCGCCAGCGCAGACCAAGATGCTGGAACCATGGGCGAAGGCCGTTGAAGAAAATTCCGGCGGCCAGGTCAAAATCGAGATCTATCCATCCATGTCGCTTGGTGGTCGC
CCGCCAGAGCTGGTGAACCAGGTTCGTGATGGTGTGGTTGATCTTGTCTGGACCCTGAATGGCTACACACCAGGCCTGTTCCCGCGCACGGAAGTGATGGAACTTCCTTT
TGTTTTCCTTGATGATCCGCGCGCAGCAAACCTTGCACTTTACGACATGTTCAAGAGCGATCTTGAGGTTGAATACAAAGGCTTGGAAGTCATGTTCCTGCACGTACATGC
AGGTGCTGGCCTTCACACACGTGATAAGGAAGTCCGCAGTCCGGCTGATATGGCAGGCATGCGCATCCGTACACCATCACGCACAGGCGCATGGGTGATTGAAGCCCTT

GGCGCTTCTCCAGCGGCGATGCCAGTTCCAGAACTGCCGACAGCTCTGCAGAAGGGCGTGATTGATGCGGCGTTTATTCCTTGGGAAATCATTCCGCCGCTGAAAATTCA
GGAACAGACTGAATATCAGATCGAAGGTGCTGGCAAAGCGCGCTTTGGTACTTCTGTCTTCCAGGTTTCCATGAATAAAGAACGCTGGGAAGGTTTGCCTGAAAACATTC
AGAAAGCTTTCCGTGATGCATCTGGTCCTGAGTGGTGGGCAAAGGTTGGTGAAGTTTGGCGTGCGACAGATGACTTTGGCATCGGCCTCGCTGTAAAGTCAGGCAACAA
ACACATCACTCTTTCTGAAGAAGAAACTGCAGTCTTTAGAGAAGCTCTGGCTCCGGTCGTTGACAAATGGACCTCAGAAGTCTCCAAAAAAGGCATTGATGGCAACGCG
ATTGTTGCAAAAGCACAGGCATTGATTTCCAAGCACAGCGAGTAGCTATGGGCGCCTTTCAAAAACTCGGACGCACCGGCATTGCCGGTGCTTTCGAAAAACTTATTGTA
GCATGGGCACTTACCGGCGGCGTAGTTCTGCTGGCAGTGGTCGCGATAAACATCTACTCCACCCTTGGCAACATATGGGCTGCTCCGTTTCCCGGAGACTTCGAGCTCAC
TGAGATCGGCGTTGCTATTTCAGTCTTTTCCTTCCTGCCCTTCTGTCAGATCACTGATGCCAATGTCACAGCGGATATCTTCACCAGTAAGGCTGGCCCTCGTTTGAAAGC
GTTTCTGGCTCTGATCGCATCTGTGATCGCGTTTGGGTTTGGCTTGCTGCTGCTGTGGCGCATGTACCATGGTATGTCTGATCAGAAAGAATACGATTACACCACCGCAAT
TTTGCAGATCCCTGTTTGGTATGCGTATGTACCTGTTTTGATTTCGCTTTTGCTGCTGAGTATTGGCGCGGTTATTACGTTGAGTGAAAACACCAGAGATCTGGCTCGGAG
AACCTGATGGCGGCAAGTCTTTTCATGGGCTTTGGAGCCCTCGCCATTCTGGTGTTTCTCATTGTTATCCGTATCCCAATTGCTTATGCGATGATCCTTGTTGGCGGCGTTG
GAATTGCGCTCATCAACGGCCCGATGTTACTGCTTAGTCAGTTGAAAACATTGGCGTATGGCCAGTTCGCCGTTTACGACCTCTCCGTCATCCCGCTGTTTATTCTGATGG
GTGCGCTCGCCTCGAAAACCGGTTTGAGCCAATCTCTCTTCAGAGCAGCTAACGCATGGCTTGGCTGGATGAAGGGCGGCACAGCGATGGCCGCGATCGCCTCCTGCGCC
GGGTTTGGCGCGGTCTGTGGTTCATCGCTTGCGACCGCCTCAACCATGGGAAAAGTCGCCCTTCCGGAGTTGAAGCGTTACAATTATTCAGGTGCTTTGGCTACAGGAAC
GCTTGCGGCTGGCGGAGTTCTGGGCATTCTCATCCCCCCAAGCGTAGTATTGATCATCTATGCCGTGATCGTGGAAGCAAACATCGTTACTATGTTCGTTGCTGCGTTTAT
TCCGGGTATTTTGGCCGTTCTTTTGTTTCTCGCAGTCATCGCGCTGTATGTTTTCGTCAGTCCCTCATCTGGCCCTGCAGGTGGTGCTGGAACGCGAGAAGAGTTCCTCGCA
GCGTCCCTCGGCGTGCTGCCGGTGCTGACTGTTTTCGGACTTGTCATCGGCGGTATCTATTTCGGCTTTTACAATCCCACACCAGCTGCTGCAGTAGGTGTGTTTCTTGTGA
TGCTATTCGGTCTCCTCCAACGCAATCTTGGTTTGGGGGATATGATCGAAGCATTGAAGGAGACGGCTGCAACCTCTGGCATGATCTATCTGATCCTGCTTGGTGCAGAA
CTGCTCAAGATCTTCATGTCCCGCATTGGCCTTCCGCAGGAGACGGCGGCTTGGATTGCAGGCTCAGGTCTGGAACCAATGACCGTGATGCTGCTGATCCTGCTGGCTCT
GATCGCGCTTGGCTGCTTGATGGACAGTATGTCCATGATGTTGCTGGTCATCCCGTTCTTCTGGCCTGTACTTATGGAACTGAACGGCGGCATGTATCAGGGAGCAGATG
GTGCCGGGTATGGCATGAGCACCGAAGACCTGAAGATCTGGTTCGGTATTCTGGCGCTGATTGTCGTCGAGCTTGGGCTGATTACACCGCCAGTGGGCATGAATGTCTTC
GTGATTTCAGCTTTGGCAAAAGACACACCAATGATCCAGACCTTCAAGGGTGTTGCGCCATTCTTTGGCGCCGAGCTGGTGCGTGTCGCCATTCTGGTGAGCTTCCCAGC
GCTAACCCTTTGGCTACCCAGAGTGCTTTCCTAAACGCAAAACAACAGGTTGTCTTGGCAGGACTGATTGCTGAGACAACCAAACATTCCAAAGCGCTCACATAGCAATT
TTGCTATAATTTCTGCCTATAATTCAATTTTACATATCAATCATAAACAGCAACACTGAACACTATTCAAACGTGTATTCGCAATCCATTTTACTTCTGACTTAGCGCATAC
CCCAATTGCCGCGATCTGCCAGATGTTTTGAACCTAAGGAATTTTTCATGACCCACCATAGAGATCGTACTCAGGTTGCGATTATCGGCGGCGGTCCTGCTGGTTTATTAC
TCTCTCACATCCTGAGCGAAAACGGCGTGGACAGCGTTGTGCTGGAACAGCGCACCAAGGCTTATGTTCTCTCTCGCATCCGTGCTGGTCTTTTAGAAACCGGAACTGTG
CAGCTGCTGCGCGACTATGGCCTTGCAGAACGCATGGACAAGAACGGGAAGAGCAAAAAAGGCTCATGGATTACCGCGGCAGGGGTTTGAAAGTCACTTCATTGATACGC
ACAAATGGACTGGCAAGGAGATGATGGTCTACGGTCAGACCGACATCACCGAAGATCTGTATGACGCCCGCGAACGCGCTGGCGGCAACATCATCAACGAAGCGACTG
ATGTAACTCTGCATGATGTAACCTCTGACAAACCTTCTGTGACCTACGTGAAGGACGGGCAGACCTATCATCTGGACTGCGACTACATTGCTGGTTGTGATGGATTCCAT
GGTGTCAGCCGCAAGACTATTCCTGACAACATTCTGCGCACCTACGAGCGTGCCTATCCGTTTGGATGGCTTGGCATCATGGCAGAAGTGCCACCCCTGCCGGATTTGGT
CTATGCCTATCATGAGCGTGGGTTTGCACTTGCCTCACAGCGAAACCGCCTGCTGAGCCGGCTATTACATTCAGTGCCCGTTGACAGACAGTGTTATGGACTGGTCGGACG
AGCGGTTCTGGGAAGAGTTGATTAGTCGTTTTCCTGCGGATGTGGCGTCAGATATCGTCACCGGACCATCTATTGAAAAATCCATTGCACCGCTCCGCAGCTTTGTCGCTG
AGCCGATGCAGTACGGGCGCATGTTCCTTGCGGGCGATGCCGCTCACATTGTGCCGCCCACAGGAGCGAAGGGGCTCAACCTGGCTGCCTCCGATGTGTTCTATCTTTCA
CGCGGGTTAATTGAGCAGATCAAGAGCAATCACTCCGCTCATCTGGAGCCTTACTCCGAGCTGGCTTTGCGCCGAGTATGGAGCTCTGAGAACTTCTCCTGGCGGATGAC
GCAGCTGTTGCATGTGTTCCCGGAGATGAACGGGTTTGATGCGAAAATCCAGCAAAACAGCTACGAGCTGCTGCTGCAAAATGAGACCCAACAGCGGGCCTTGGCAGAA
GAATATGTTGGTTTGCCATTTGAAGAGCTGGATGCGCTGACCTTGTCAGCAGCTTAGTCCAGCAATTGGAGAGAGGCCCATCCCTCTCTCCCCATCTCTTGAATGAAGCTT
ACGATTCTGTTTTGTCAGCAATTTCAACAATTGCAGCGGCAAAGGCCTCAGGGCTTTCTTGCGGCAGGTTATGCCCTGCGCCCTCAACAACACGCGCCTCATAATGTCCGT
TGAAGTTCGTCTCATGCCCTTTGCTGTGTCGCGGTGGGCTGACAGCATGATTGGCACCAAGCAATGTAATTGCAGGCACATTGATTTTGGGCTGAGCTGTCAACAACACCT
CAATCTTCTCAAAAGCAGGATCTCCTTCCACCAACCCAAAACGGTGCCTGTAGGAGTGAATAACGGTCTCCACAAAGTCAGGGTTATCAAAGCTCTCGGCGGTTTCACGG
AAGGTGCGGCGGGAGAAGTCCCAGGTGGGTGACCATTGCTGCCAGAGCAGTCTGGTCAGGTCTTCCCGGTTCTCTGTCAAACCATTACGGCCACGCTCACTGTGGAAGTA
ATACTGATACCACCAGAGCTGCTCTTCCTCAGGAGAGGCAGGTCGGTAGGCTGCTGGGATATTCTGGATATTGTAGGCGACGCCGGCACTCACGAGCCCCCTCACCTTTT
CTGGCCAAAGAGCCGCGGCGATGCAAGCGGCGCGACCACCCCAATCGTAGCCCGCAACAATCGCCTGCGGCAACTCAAGAGCCACCATCAGAGACCGCAGATCAAAGG
CCAGCGCGCCTTGCTGACCGGAGCGCGGTGTTTGAGGAGACAGATATTGGGTTGGCCCGTAGCCGCGTAAATAAGGAATGATGCACCGAAAGCCGCGATTGGCGAGAAT
GACGGAAACCTCGTCATAGCTGTGCACGTCATAAGGAAAGCCATGAAGCAGAACCGCTGGAACACCTGACTTTGGCCCAACATCCATGTAGGCGATGTTCAGCACCCCG
GCTTCAATAAACTTCATATGTCCAGTCCCCTATATCGCGCTGTAACCTTTCAAACATTATAGACCAGAGTCACTGGGAGAAAGCCAATCCTTATAGTCATTCTTAAGGCGT
GCAACCTCTTCTCCGTTCAGCAAAAAGCAATCAAATCCACCATCGTAAGGAGCGACGATCCGGGAGGTTTCCCGGTTCATGAGCAGGAACCGCTCTTCATCCTGCGCTGC
GTCTTTGAGGGCCTCGCGACACACCTCCGGAAACCAGCGGACTTCATTCCCGTAAAACCCCAAGGTTTCCTGCTCCCCAAAGGGATCTTCCTTGTACTTCCACTTAAACCA
GTATTTCAGTCGAAACTGATCGAGGAGTTTGTGGCTGTACGGTAATGGATACTCATCAAAATTTGGAAAGTACTCCACGAAGAGGCATTTCTGACCCTCCTTCAATATCC
GGTTTCCCAAAATCCGCGCGCGATGAAAAATCACGGCTTCCTCTTCTGGTGTCTCTGCATATCGCTTCGCCTCTGGCAAAGCATGAAACCTGACCCAGCGCTCCGGCAGC
GCCTCACGGAGATAATGGCTATAGGGCGGTAAGTCGCCAAAATAGCCTCTCCATTGATCAATGAACCTTTTCACAATGATACAACCTTGTCTCAACCATACAGCGGAACC
ATTGAATAATCAGAGGATACTTAGTTGAGATTGGTTCGAATTTTCCTAGGTAAATAGCTGCAATGCTCTTTGAATTGAGAGGATTTACAGAGACACCTTCACATTCTCATA
GCTTTGCATTGCCGCACAGAAGGCTGCAAAAATGCGAGTGGAGACAGGGTTGCTCACCGCTTCAAACTCCGCATGCCACTGAACGCCAAGGAAGAACGCTTTGTCTGGG
TGGTGGAAGGCTTCAATCACGCCATCCTGAGAAAGGGCGTCAAGCACAAGCCCGTCACCAAGCCTCTCAACCGCTTGTTCGTGCAGAGAATTTACGTGATGGGCGCCTTC
GCCAAGGATGGAACTCAATAATCCAGTCTCTTCAAGGGTGAGTTGGTGAGCTGCCCCGTATCTTTGTTTTGGTGGCAGTTTGTAGTCGATGGGGTGGTGCAGCGAAAACT
CATGGTTCGCATATAAGTCGTTTCGGATTGTGCCGCCAAAAGCCACGTTCATTTCCTGCATGCCGCGGCAGATTCCCAGCAGCGGCAGGTCTTTTTCAAAAGCGGCACGG
ATCAACGCAAACGCGGCAGTGTCGCGTTTATCATCAAAAGGGCCATGGCTCTCAGAGGTGCCATCATTATTGTAATAGGCTGGCTGAATGTTGGAGGCATCGCCTGTCAG
AATAAGTCCGTCAACCAAAGACAGCACATGAGCCAGTTCGGCTTCCTTCAGAAGTGTCGGAATGATGAGAGGGATTGCTCCAACAGCCTGCAATGCAGCTGTCAGGTAC
TTATCATCAATGGTTTGCGTTCCAAGACCAGCTTCTTGCGAGGAATTCGCCAAAATTCCGACAATGGTTCTTTTATCAGCAGCCACGCTCTGCGGCGTGGTTTTGCCAGTA
TTGCTTGTCATAAGCGTATCCCTTCAGGGCACGAAAACTCGTGGCGCACCTATGCGCTCAAAATCAATGGAAGAATAATTATCGCTTCAGCGATTTTGCGTAATGTGTCTC
AATTCGCTGTTGCACAAATTCCAGCATAACGGACATGACCCAATAGATGATGGAGGCGGTGATCAGCATCTCCAGAATCTTGAAGTCGGACCGCCCCTGCGTTCGTGCGA
TGAACATGATGTCCCAGACGCCCACAACTGAAACCAGCGAGGAATCCTTGAGCATCGCGATGAACTGGTTGCCCGTTGGCGGCACGATAACCCGCATGGCTTGGGGCAG
GATGACTTTTCGAAACGTCAACAAACGCGGCAAGCCCAACGCATCCGCAGCTTCCCATTGCCCAAGTGAGATCGACTGGATGCCAGAGCGGAAAATCTCGGTCATATAC
GCCCCGTACACCAGCGACAGGGCTATGATTGCTGCAGGAACGGGCGAAATGACATAGCCTATTTGCGGTAGGCCGAGGTAGATCAAAAAGAGCTGCACCAACAGTGGC
ACACCGCGAAAGTACGAGGTGTAAAAGGCTCCAACACCTTCAACAACACCAATGCCCGAGAGCTTGGCCGTCGCCCCCAGCAAGGCAAAGATGAAGGCCAGTGTGATC

GCAATAGCAGAGACATAGAGCGTGGTGAACGCCCCTGTCGTTATCAGGAAGGGCAGTTTGCGACCGATAAACTCAAAACTCAGCTCAAAGCTCAAAAAGAACCCGAGA
GCCCCTGCAGCAAGCAAAAGCCACACGGTTATCAGCTGAGCCCGCCGGGACAGGCTGAGGACGAGAAACCCCAGTCCTGACACCAGCAGGAAAATCACCATGGATACA
GAAAAACGCCCCCACCAGCCGGCAAGTGTTTCCTCACTCATCGCAGGGCGAAGGAATGACCCGATCAGGCCTTGATTGAACGCGAGAAAGAACGCGAGCACGAAGGCA
AGTGAACCAAGTGTTGCTGCAGCCTTCAGGCGGGCTGGACCCTGCCTGAAAGCACCATTGGTGAGTGTGACCATGATCTTACTTAACGCTCACGAGATCCGCACCGTGCC
ATTTCACGGAGAGCTTGGTCAAAGTGCCGTCACTCTTCATGTCCTCGACAATTGAAGCCAGCTTGTCACTCAGTTCCTTGTCGCCTTTGTCAATCGCAATCGCCAGCGGCT
CATAGAAGATCGGATCGCCCAACTGCTTGATCGGCAGGCCTTTCTTCATCGCCTGCTGAATGACAGGAACGGACTGCGCCAATGCGTCAAGGCGTACACCATCGCCAAG
AGACAGCTCGTCAAACTCATTGATATCGCTATAAAGTTTGATCTCGCCCGGCGTTACAGCATATTCGAAAGCTGGTGCGCCAACCGCATCAATCTCAAGCTTGTGCTCCA
GATACTTGTGGTAGGTAGACGAACCCACAACGCCAATCACCTTGCCGTCCAGATCCTTAAGCGATGTTGCCTTGGAGTTTTTATGAACCGCAAAGGCCGCCGGGGTGTAG
TAGTACACGGCAGGGAAATCCAGCACCTGCGCCCGGGCTTTGGTTGGGGTCATAGAACCTACAACCATATGCCATCTGCCAGACCATCGGCCTGCAGTCATGATTTCCCA
GCCCGGAGTGACAAACTGCACATCGACGCCGAGACGCTTTGCGATCTCTTTGGAAACGTCAATATCAAAGCCCTGCAGCTGGTTGTCGTCTCCCATAAAGGCATTCGGGC
GATAATCTGAGTTAGAGCCCACGAGCAACGTCTTGCTCTTCATGATCGCATCCAGAGCCTCTCCAGCCTGTGCGCTCATGCCAACAACGTTTGCAGAGACAAACAATGCA
GCCCCTAAGATGGCTTTTCCGGTATTTTTAAACATCCCAACCCCCAAATTCGTCATTATGACGGAAATTTCTCAAAGTTTTGGGCAGTTTATAAGTGCGGGTCAAAGCAAG
GCAAATCAGTCTGATTGATCTATGTATAGAAAAACTCTATACCTCATTCATGACTATCTCTCGACGTTTCGTCCCCACCCTTCAGGAATTACGCGCGCTGGCAGCAGTCGC
ACGGCACAAATCCATCTCCGCTGCCGCGCGGGAACTGAACGTCTCACAGCCAACAGTATCCTATCATATCAAGCAGTTAGAAGAGCGTTGGGACGTCAAACTGTTCCGC
ATGAACGGACGAAATCTGGAGCAGACGGAGCTGGTTCAGGATATTCTTTCGGAGGTCTCTGCAATCAACAATGGTATAGAGAACCTCTCATACCTACTCGCCAGCCAGA
CCGCACGAAAACAGCTTACAATCGGCATTGCGCCGTCTCTTGCCTCCATTATGCTGGTCCCGCGTTTAGCTCAATTCTCCGCTTTGCACCCGCATGTGGCGATCCGGATCA
GTGCAGCCAATCGCTTCGTTGATTTTGCAGAAGAAAAGCTCGATGTGGCACTGCGTCTACTGCCCAAGCCTGAAGGCGACCTCGGCGCATCTCCGCTTATTCCCCTGCCG
CAGGAGCGGATGCGCGTGGTGTGCTCACCAAGCTATCTGGAGGCGATTGCGGGAGAGCAGCTGGCGCCGCATTCAGCTCCGGTTGAGCTCTTCGCCAAAGCTGATTTCAT
CCATGAAGACGAGGCCTTCCACTGGCAGAAGTACCTGGAAGCCAGTTGGCCAGACTTTTCCGGCCAGATGGTGCAGCAGCTCACGTTCAACAACACGGACCTTATCCTG
CAAAGCGCTATTGCCGGACGTGGCTTTGCCATTCTCAGAGATCTCTATGTCATTGATGCATTGAATGCAGGCACGCTCGTAGAGCCCTTCGCTGAGACACTCCCATGCCA
GCGCGCCTTCCAGTTTGTCTTGCCAGAAACATCTATCCCCTCAGCCCCTGTGTGGGACTTCATCAACTGGCTGAGCAACGAGATGACTCAGATGCAAATTCGCTAAGTCTC
TTGCATCAAAAACTATCTGGATAAGCTCCTCAGACAAACCCACATGCTCCCGACACATTGTCAGCTCGCCCTTCATGGTTCCCGTCAATGTGCCGAGGAAGTAATCACCT
TTCACCCCATCTGCCCAATACATCGTCAAGCAAACAAAATTCAATCTCTAAGTGATTATCAATAATTCACTTTAGCACTATACAATATTCATTGGTTATGTATTGTCATATT
AAATAAATATAATTCCTATATTCATAAATACAAATTCTCACAGCAATAAATACAACAAAATTAAAATCACTAAACCAAACTACGGAAACAAACTACTTATATGATGACG
AGCAATCAAATTATTCATCGAAACCTCAAAATAATTTACTTATCGCGGCGGAATCAAACATCGCTATCACAAGGAATTCAGGAGCTCAGCATCCGGCCCGGCAATAAGG
CCGGAAATGCATTTCGCAGAAGAACAGGCATTCAAAAGTAGGCTTGGTCTGACTGCAGAGTTATCGTAGGGTGATAGCAGGGCATACCCGTTAAGCGGGTGCCAGCACA
CTAGCCCAAAGCCAACGGCTGAGCATAGTTATTGTCCTTTTCAGAGCCGAATACGTTCAAGAAACTCTTTTCTTATTGGAAGGAAAAGGATGGACATTCATCAACTCAAA
ACCTTCGTGACGGTGGCGCGCGAAGGCAGCATAACCAAAGCTTCAGAATTGCTCTTTCGCAGCCAACCAGCTGTCAGTGCGCATATCAAGAACATGGAAGAGACATTGG
GATTAACGCTCTTCCAACGCACTCCGCGCGGAATGGCTGTTACACCCGATGGGCAACGCCTCCTGACGGAAGCACACCAATTGCTGGATCGGCACCAGAAGTTTCTGGA
GGAAGCATCCCGCTTACGCGGAAGCCTTGCTGGTAACTTGACCCTCGGCGCTGGTCGCAACTCAAGCACAGGCGTTGTAAGCAATCTGCTTGCTGCCATGTCTCAGGAAT
TTCCTGAGGTTGAAGTTGAGCTGCAGCATCGCCCCTCTGTGAGGGTACTGGAAGGCATTCGCAATGGTAGCCTGGACGCCGGATTCTATACGGAAACGGGCGAGGACGA
TGCCGGACTGTCCACGTTGGAGATCTCCAATTTCTGTGTTTATCTCGCAGCACCTGCGGACATGGTGCCCGTCGCTGCTGAAGTGAACTGGAAGGCGCTTGAAAACGTGC
CGTGGATCTATCCAAGTGGCAACTCCTGTTGTGGGCGTGTGGCAAGGGATCTTCTGGACACGCACGGCATCAGACCCAAACGGGTGATCAATGTGGATGACGAGGGCGT
TACACGGGCCTTAATTGCCAGCGGTGTTGGTATCGGGCTCGTTCATGCCAACTCAACCAACCCTCAAGAGAACCAGGGTGATATCGTCTTGCTCTCTCAGGTGCAAAAAT
CAGCCCGCGTGTTGTTTGCGCACCTGTCCTGTCGGATAGAAGACCCGGTGATTGCAGCAGTCGCGCTCCATTTTGGCGACAGACAACTGCACGAGCCTCTCGTTTGAAGAC
GACCGAACACTTGGGGCGGCGCTGATCGCCTCCTGACACTTTCAAAAGAACCAGATTTGCAAGACTCCAGCTCAGCCATGAGCTGGAGTCCGCTGCATCCTGCTTTCCAG
TTGGGCGGATGGCTGCTCTGTCCAGTTGGCGTTAAAGACAAAAACAATCCGGAATACTGAGGAATTTTTTAGCACAAACCCATTGTTTTATTCTCATCGGGAGAAGCCAT
ATGCAAAGACTTCTTTACAGCCTCAGCGGAGCCTCAGGGCAGCACTACTCCCCGCATGTCTGGAAAATTATGATGTCCCTGAAGCACAAAGGGCTCGATTTTGACCTGAA
GCCCGTTTCCTTTGGCGACATTTCTAAAATTGAAGATGGATCCTTCACCAGCGTTCCGGTTCTCAACGATCAAGGCACTCTTTTAGGCGACAGTTTTGAGATCGCCGAATA
TTTGGAAGAGACCTATCAGGATGCGCCTAGCCTGTTTGATGGCCCTGGAGGTCTCCAACTTTCCCACTTTGTGGAAAGCTACTGCAAAACCGTCCTCCACCCGACATTGG
CGGTGATTGCCGTGATGGATATGTACAACATCATGAGCGATGCCGACCGCGCCTACTTCCGGCCTGCCCGCGAAAAACGGTTCGGAAAATCGCTCGAAGATGTCTACGC
AGGGCGAGACGCCGAACTGAGCACATTGCCGGAGAAGTTGAACCCCGTGCGCGACCTGCTCACCAACCAGGATTGGTTCGGTGGTGAGAAAGCGATGCATGCAGATTAT
GTTCTGTTCGGGACCTTGCAATGGTGCAAAGTGTGCTGTCCAGATCCAATTCTTGCTCCGAGTGACCCCGTCAATTTGTGGTTCGAACGCATGTTGGATCTTCATGATTAT
ATGGGTCGCAAAGCTGTGAGCGAAGCTGTCGCTGCCTAACGCCACAGATATTTCACGCGGTTCCAAATAACTAACCCCTTCAGAAGGCATCGCAAATTTTTGCGGTGCCT
TCTTTTTATGCAAGCCCCTGTATTATCAATAATTTACAAGCAATAACTTCCGCTGAGGCATTCCTCATCTATCCTGAGGGAAGCCATCAGTTTTTTCAGCTTGTTCAAAGCA
AACCTTTCCGGCTCCATAGGGTCACACGAAAAAATATTACTGGACGCGAGGAACTAAAATTCACCTCAAACCCGGCAGGGCAAACGGAGATTCATTCGAGTCTGCGCGT
GCTACCGCTCTGCCATGCATAAGAACAACGACAAGCCACGCAGTTTGTCAGGGGGTGCGCCTCGCGCAAATCCTGCATCGGCATTTTTCAGTCTTTCTTCCCAAGTGAAG
GATCAGAGAATTTATGGATGCACATCTCGAAGCTTCGGACTATTTGGCCAAAGGAACAGCCCCACTGTCCTTTGATTGGGTCCGGCAAGTTCACACCTTGGGGCCGACTG
GAACGAACTGCGAAAGGGCAGCTCTGACCTGGGTCAAGCGCCATTGCCAGAACGCCGAGGTCTTTCTTCACCCAACCATGGAAATTGCAGCGGAAGCCGTTGCCAAGTG
CCGCAAATCCGTTCTGTTGAGTGTCGTCGCCTACCCACATCTTCATTCGATTATTTATGAGCACATCGGCCAACTGAAATTGATGGACGTGTTCATCATGAACACCGACAA
CATGGTGCTGGCCTCATCTACCGGCAAAACGCCCGCCCTTTGCGCAACACATCCGGCACCGGAAAAACTGATCCCGGCATCTATTGAGCGCAAATACGTCACCAGCAAT
GCGCAAGCCGCAGCTGTGTGCGCCAAAGGCGAGGCTGATGGCTGCATCACCACGCTGTGCTCCGCCAACGAGTTTTCACTGACTGTCCTGACAAAGCACGGCCCCGTGC
CTATGGGCTTCACCATTCACGGCCCATTGCAGGCCGCAACACATACCCAGCCTTCATTCACCAATGATTTTCAAAACCCCAACAGGAGGGAATTTCATGACACAGAAAGC
TTATGAAACACCAATCGATATCCGCTACCGCGATACAGATTCAATGGGCCATGTGAGCAGCCCGATCTACTATGATTACATGCAGTCCGCTTACCTGGAATACATGCATG
ACCTGCTCAAGCTGCCTAAATCTGAAAAGCTGCCGCACATCATGGTGAAGACCTCTTGTGACTACGTTTCTCAGGCTCTTTATGGCGACCATCTGGTTGTGCTGAGCCGCG
TGACAAAGTTTGGCGGCAAGAGCTTTGAGATGCAGCATGAAATGCGTCTTGGCGATGCAGATGGCCGTGTGGTTGCCAATGCAGCTTCCACTCATGTGATGTTTGATTAC
GAAAAGCAGGCGACCTATCCGGTGCCGGATGAGTTCAAGACTGCGGTTGCAGCTCTGCAAGAACCTGTCGCAATGCACGAGCCAGCTTAAGATCCGTCACTTTTGAAAT
CTGGAGATCTCTGATGGATCTACAATGGAATGAACGTCAACAGGCTCTGCGAGCAAAGTATGCTCAAATTGGCAGCCAGATTGACCACTCAGAGCTTCATGCTGGCAAG
CAGGCATTTGATCAGGACACATGGGACCGGCTTTTAGAAGCCGGTCTCTGGAAGATCATAGTCCCGGCAGAATATGGCGGCGCAGGCACAGACTGGTGGGACTTTACCG
CTGCTCTGGAAGGTCTTGCATGTTCCGTACGTACGCCGGGTATTCTGCTTTCTGTCATAGCACAGGCAGGCATGGTTCATGCACTCAACCTGTTCGGCACGGAAGCACAG
AAGAAGGATTATCTTGGCCGTATCCTCAAAGGAAAGTTGAGCGCGACTGCAATTGCGGACCCTGACACGGGCACCGATGTACGCTCCACCTCTTCTGTACTGACACCGGG
ACCAAATGAGACTTTCATGCTCAATGGTGCCAAATACAACATCGCCCATGCGCCCGTGGCGGATTTCACACTTGTCGTGTGTAAGCTGGATGGCCACGCTCGTGATGGCA
TCTCGCTCATCCTGCTGGATGCAGGCACTAAGGGCCTGGAAGTTGGCGACTACGACCAGAAGCTCGGCAATCTGGATTTGCCAACAGGGCCGATGAAGTTCACCAATAT

CCCGCTGCATTACGGTGACTTGCTTGGCATTCCCGGCAAGGGCCTGCGCAACCTGATCACCTTCGTTTCCCTCGGCAGGCTGTATTACGGTCTGGTTTCTGCATGGATGAT
TGAGCCGATGCTGAAGGAAGCGATGCGCTACGCTCAGGAGCGTCAGACCTTCGAAGTGCCGGTGCTGGAGTTCCAGTACGTGCAGAAAAAGCTCACCGACATGCGTATC
TCCGCAGAATCCGCCAAGTGGACTGCTTATGGGGCGCTCCAACAGCTTCTGTCCGGTGCTCCGGAAGCAGGTATGACCTGCTCTATTGCCAAGCTGGTTGGGGCTGAAGC
AATCACCAGCGGCGCGATCGACCTGCTCAAGCTTTACGGCACCAAAGGCTATCATCAGGGTGAGGTTTCGACCTTCCTGCGTGATGCACTGGCGTTTTGCGCCGTGGGTG
GCACTGAGGAAATGCACCGCCGCAATATCATCAGTCAGATGAAACGGCTGGAAGCAAAGGCTGCAAAAGAAGCAGCACCCGTGCGCGAAGAACGCTTGCTTGAAACCG
CATGATTTGATGCAGGCCCCGATATTGCCGGGGCTTTGCACTCTTTTGCATTTACAGGAGACTTTTATGGAATTCGCCAAAGTCGCATTAGCGGCCACTCTTCTGGCAGCA
ACAACCGCCGCTCCCACTCTGGCAAAGAATGGTGTTGAGCGTGTGCCAACCAAGATCGATTGGCAAGCAGCCAAAACCGCTGCAAAGCGGTCACTCGCGCGGGATCCGG
CCGCCGTGGAGAAGTTCGTCGCAGCCACCAATGGCAAGCTCAGCTACATTGAGCTGCCTGTCTTGATTTTCGGCAGTGATGCGAAGATGGAGACACCACAGTTCGGAAG
CCAGAACTCGTCTTATGCCGCGTTTTACACTCTGGAAGGTGCGCAGATCAGCATTCTCGGTTCTCACAGTGTGCTGACCGGAACGGAAGATCTCACCGAGCATCATAAGG
CAACAGCGCATGAATCCATCGGCAACGGGGCAGACTACAATCTGAACCGCTACGGGGCGCATTACACAATCCGCCTGACTTGCGATAACCCGATCAAAGACACGCGCTG
CCGTAAGCCTGATTTTCTTCAAAGCGTCGTTCAGACCCTCCTTGTTGTGAATGGAGCAAACCAATGATGCGATTTATCTATCTCTGTCTGGCCCTGACCCTTGGCTTCTCCG
CTCACGCAGATGACAGCTCCCGCTTTATGCAGCCGGGTGACCTGATTGCAGACACCGGCAAGGGTGTTGCAAGCCATGCCATCTTTTTCCCAAACATCCGGTTTCCGGTG
GAATCCGGCCCGGCCTATGCCAACTCTCACGCTTATGCGCCCGGCGGGTTCAAAGCTGATGGCGGATCCGAGTGTGATCCTGTGAACTACTCCTACCCATGGCGCGATAA
CTTCTGCGAGTTCCGTAGAGCCGATCTGCCAGTTTGCGCCTCAGGCGGCCATCAGGGGCAGGATATTCGTCCGGCATCATGCGAAGCCAGCAAGCACTGGGCTGTTGCAG
CCGAAGACGGCATTGTTGCTCATGTAGGTCGTTTTGGTGTGACCATCCAGAGCGCCACCGGCACGCTGTATCGCTATGTGCATCTGGACATGGGCACCCTTCAGGTGAAG
CCGATGGACAAGATCGTCAGAGGCAACCGCATTGGCAAGATCTCCAACAACTGGCTTGGACCGCTACCAATCCATCTGCATTTTGATGCAAAAGACACGGTGATGCTGA
ACGGCGAACCAAAGTCGGTTTTCCTGCCGCCTTACACCAGTCTCGTGGCGTCCTATCAACTGATTATGGATTGATCAGCTGAGATTTGATCGCAAAGGGCGCCTTTCACG
GTGCCCTTTTTGCTTTTCAATATTCTCCTGAAAAGGGAAAAGAAATGAAAGATATTCTTTGGGTCTTTGGCTACGGCTCACTCATCTGGCACCCCGGCTTTGAACCCGTTG
AGCAAGTACCCGCTCGTCTTGATGGCTGGAAACGCCGGTTCTGCCTGTGCTCCCTGCGCTACCGTGGCACTTTCGACAAACACGGTCTGGTTTTGGCACTGGATGCAGCA
CCTGATGCGTATTGCAGCGGTGTTGCTTTGGGCATTCCCCGTGATGGGCAAGAGAAGGTGCTGGAAGACTTACGCAAACGAGAGCTGTTCAATCCGGCGTATAGAGAAA
TCATCACCACGCTTAAACTGCAGGATGGCCGCGAGGTTGAGGCGCTCACCTATGTGGTCAACCGCGAGCACGAGCAATACGGGCAGTTTAACGCCGAAGAACAAGCTCA
GATCATCGCGACAGCTGCTGGCGAGCGCGGCCCGAACTGTGACTACCTCTGGAACACAGCAGATCACCTGACAGAGCTCGGCATCGCAGACCCGGACCTGGACTGGCTC
TCAGACCGCGTGCGCTCGATACGGTCTGCTGCAGCAGCCCCAGCCTAAGCAAACCAGCACGTTACCAAAACGTGCTGGTCTCCTTTTCAGCTTGCAGCGTTCGTTTCCGG
CGCTTGGGTGCAGATGAGTTCGCCCTGCTCCAAATGACAGACTTTGTCTGCCAGCGAGCGGGCGGCGGCATGATCATGAGTGAAGAGCACCAGAGCCATCTTCTTGTTCC
GCATTACACTGCGCAATATTTCCAACACCTGACTGCGGGCTTGCATGTCCAGCGCAGAGGTTGGCTCATCAGCAACCAGACACTGAGGCCGCGTTGCCAGTGCGCGTGCA
ATCACCAGCCGCTGCGCTTCACCGCCAGACAGCTTTCTGCTGTTACGGTGCAGGAAATCTGAATCCGTCGGCAAGCCAACCCGCTCCAGCAGCCGCGTGAGTTCAGACTG
ATCTCCATCAGACAGCATGAGTGCTTCTTTGAGGACCTCATAGACCGTAAAATGACAGCCCATAGCTCGGTGCGGGTGTTGTGGTATTAGCGCCACCACGGATTTAGCCC
GTCCACTCTCGTTTTGCCAATAAACGAAACCGTCCTGCACGGTTTCATAGCCGCTCAACAACCTCGCAAGCGTGGTTTTCCCGCTCCCACTTCCCCCCAATATCGCAAGAC
ACCGGCCAGCTGGCACGAATGCAGAAATGTTGGAGAGCAAGGCTCTTCCATCAAAGCTGTGATGCAGTTTATCAAGGGTCAATCCATGACCACTACGTGCCTCCAACTCC
GGATACCCCCCTTCTCCACTCGCCAGGATCGGGGCTGAGTTATGCGCGTTTCGGCTCAGTTTCAGCAGTTGGTGGGTGTAGATGCTGCGCGGTCTGTCGAGCACCTCTTCC
GCTGTCCCGATCTCAGTTAAATTGCCTTTGCGCAGTACGGCTACATGACTTGCCAGACGACGGTGCCAGTTCGATGTCGTGCGTGATCAACATCTGGCACCGGTCCTCACTC
CCAGTCTGAAATGAACGAATGGCGATTTCCCGATTGAAGGTATCCAGTGCTGCAGTCGGTTCGTCCATAATGAGAAACTCAGGCGAGTTTGCGAGTGCAAGCACCGTCA
ACACGCGTTGAATTTCCCCGCCGCTCAGAGCCGAGGGATACCGATGCAAAAGACGTTCCGGGATATTCCCTCTGGCGCATTCCTCATGGGCATCACGCTCATAAAGTCCA
TGCGCTCTCAGCATCTCCTGCACATGCTGCAGGACCGTCAGGTGCGGGTTCAATGCATCCGACAGCGCTTGGACGATCAACCCTGCAGAACATCCTCGCATTCCCTGCAA
TTGTGCTCGGGACGAACGCAGAACATTCATATCTCCCAGCCAAGCTTCTCCATTCCAACTGAAATGGTTGGAGGCCGATGGCTGTTCATTTTCTGCCAGCCTGTGTCCTTG
AATAAGACCCATCAGAGAACGGGCGAGGCTCGTTTTACCACCCCCAGGCTCTCCAATAACCGCCAGAGTTTGGCCTTTGTGGATCTCAAAAGAGATATCCTCCAGCAAA
GACTGGTGACCGTAGTAAAGATTAAGCTGCTCAACGCGGATCATGATGTTCGTTTCCGTTGAGATTGCCGTTTGACATGTAGCCTTCACCAATAATCAAAATACCAACGA
GAAACAGAGAAAGCGCTCCCGCAGGCGGAAGCAGGAGCCAGAGCCATGCCTCCGTATAAAGATAGTCCATTGAATCCTGTATCATACTGCCCCAGGTCAGCAGCCTTGG
ATCAGTCAGTCCCAAAAACCCGAAACCAGCTGTTTTCAGGGCCGCCTGCCGTACATTTTGCAAAAACAAACCAACGATCGCAGGCCAGATCGCGGGAAGAATATGATGG
CGCAGACAATACCTGCGACTGGCCCCCATGCGCCGTGCAAAATGCACGTTCTCACGGGTCAACTCACGCAGCAAAATCGGCAGAGCAACACGCACATCCTCGTGCCAGC
CAGTCAGCGACAGAAGCACCACGAGCCCAAGGAAACCAGGCTGTATCCAGGCCGCCAACAACAGCAGCAAAAGGATAGACGGGATCACCTGCAGCACATCCACCAGTC
GCAAGACGGCACTACGAATGCGCGGCCCGCCAAACGCGGCTAGACTAGACGCAGCCAATGCTAGCACCAGAGCTGCGGTGGCCGTGATGAGTGCGAGAAACACCGTGA
TGGGAGCAGCTCTTACAAGGCCGACAAAGACATCCTGCCCAATCCCGTTTGTACCAAGCCAGTGCTCCATGCTCGGAGCCTGCAAGACATCGGGGCTGGTTTTGTAAGAA
GGCAAACTTTCTGACAGCGCAAAGCCTGCCATAACCAGCACAGCTGAGGCAGCAATCACTCCTCGTGAAATGAAACCAGCACTTTGTGTGATTGGTGACCAATAACCAA
GCATGAGATCACCCTCTACGCGCCAGACTTGTTGCGACAGCATCAATGACCCAGTTCAGCAGTAAAACCAGACCAGCCATAACGATGACGGCACCCTGCAACAAAACGT
AATCACGATCCAAAATCGCATTGTAAATCAAATGCCCTGTTCCAGGATAAGAGAAGACAATCTCGACGAACAGGACGGCACCAAACAAGCCCGTCAGCGCATCACTCAT
CCGCGCCAGAAAAACAGGAAACAAGCTTCTTCCGTAATAAGCCGTGCGCAAGCGCCAGCCTTTGATGCCACGCCCACAGGCGTTGAGGATGAAAGGCCGCTGAGACAG
CGTAATCGCCTCTCCCCGTGCCAGAAAATAGAACCGGGAGATTTCGTGCAAGGACAGCGCCAGCACGGGCAGCACAGCGTGGCGGGCGATATCAAAATCCCGCTGTATC
GGGTTACTGCTGGGAAACAAAGGCTCTGATCCGCCAGATGGCAACACGGGCCACAAAATGCCAAAAGCAAGCAGCAACAGAACAGCAGAAGTAAATGGCGGGATGCTG
GCAATAACCGTCACAACACCCGTCGCGATCCTGTCAAAAAAAGCCCCAGGCGCCTTGCCTGCTTCAATACCGCCAGCAGCCCCCGCAAGCAGAAAGATAGGCGCAGCGC
AGGCGATAAGTAGCAGGGTCCATGGGATGGACGCTTTTAACAAATCCGAGACTGGTGCTGCATGCGGAACCGAATAGCCCAGATCGCCTCTGACAAGGTCCCAAAGATA
GACCGCCAGCTGCTGCGGCCAGTTCCCCGACAAGCCGAACTGAGCTCGAAGCTCAGCACTCTCAGCGCCGGTGAGGTTGCGCGAAAGGTCAGATGAAAGAAACACCTCA
ACAGGATCCCCCGGAAGCATTCGCGGTAAAAAGAACACCAGCAAAACCAGCAGAGCAAACACCAGGAGGCGAATGCCCAGACCAGACAGCATGTTTAAAATAATCGGA
GGTACCACAAGAACGGCCTGCCTGTTGTTGGTTGTCTAGTGCCAGCCACCCGGAGGCGGCTGGCAATGGGTGTCCAGCTGATCAATCAACCAGATAGGATTTATGCAGTG
CTGAAGGCACGCCTATCGCCACTCCCCCTGGCATAAACTGGAGGGAGGAAAGCTTGTCATTGTAGGCAGTCACCATATGAGGGTTTGCGATGGCGTAGGCAGGAAGCTC
CTGAGAATAAACCTCTTGAAATTTATCCAGAATTTCCTGCCGTTTCTTCAAGGTGGTCGCAACAGACTGCGCTATCACAAGCGCCTTCAGCTCTTGGCCGTCATAACCATC
ACTCGCCCACTTGCTTCCAAAAACACGGCGCATAATTCCATTCGGATCACCAAGAGTACCGGAGGAGTAAAGGGCCAGATCATAATTGCCCTCAGACAGACTGTCCTGA
AGCGCGGCTCGCTCTAAAATGCGAACCTCCAGCTCTATGCCAAACGTCTCAAGCTGCTCTGCGAGAACTTTCGCTGTTTTGCTCTCACGTCTGCTGGCAAGAAGGCGCAA
TGCAACCTGCTCTCCGTCACGGTACCACAGCCCATCTTCTCCTCGATCCCAGCCTTTGGAGATCAGAAGATCGGAAGATCGTTGAGGATTAAAGGAATACTCTGAATGCT
TCTCCTGAGAATGCCAGGGCGACTCCTCCTGAATGAAGCCAGTATCTGCCACCTCTGCCGTTTTTTGGAAAACTGTTTGGATGATGGCAGTCCGGTCAATTGCATAAGCA
ATCGCATGGCGAAGTGACTTGTCGTTAAAGCGTCCGCTATGGTTGAAGCCAAGCCGCAGGGAATGGTGGCTATACGCCTTGGCAACATTCAGACCAATCGCCTCGGCCTG
TTCCAGACGATCAGGAGGCAAGTAGGAAATCACGTCGATGTCCCCAGCATGAGCCGCCCTCAACGCCGCATCTGTCGCCATCTTCACAATAGCAATGCGCTTAAACTTTG
GTTTGCCCAAATAATAATCCGTATTCCCTTCCAGCACATAGCGCCCTTGCGCCTTACTGTACTCAACAAGGCGGTAAGGCCCACTGCCAATAACGGCATCCGGCTCCGCA

AAACGCTGTGGTTCCACCACGCCCTGATAGACGTGTTTTGGCAGTACAGGCATGGAAAGCAGAAGCCCGTTCAGCAACCCGGCATCAGGCTTTTTGAGAGAGAGCTTAA
GGTGACCAGGGGAGACAACCTCGGCTTCTCTCACATTATCGACGGAGACAAAGACGTAAGGATGCTCCTTCATATAACTGAAGGTGAAAGCTGCATCTGCCGCTGTAAC
CTGCTGCCCATCATGCCAGGATGCCCGCAGGTCCAGATACACATCACAGGTCAAACCATCCTTAGAAACAGTCCAGTGCCGTGCCAGAGCCGGGACCAGCTCACCGTCTT
TCGATTGGCTTAACAGACTATCAAAAACATATGTGGTGTAGAGATAAGCAATACCGTTTCTGCCATGAACATAAGGCGCCAGCATCCCGCGGTCTCCGCGGGCATCGGTT
ATAACAAGCTGGTTTATGGGACCAGCAATCCCTGCAGCAATTGTCAGCAGCAGAGCCGCGACGGCTGCGATTGCAACGCGCACAATGGGTTGATGAAAGATGTGCATGT
CTCAGACTTTCTGAGCGGGACGCACGGCAACATGCGCATCCTCAGGTTCCGGATCATGGCTGCCACGCAATTTGCGCAGTTGGCTGGTGGTCTTCTATAGGGACATAGGG
TTTGAGATCCAGAATTGGCGTGCCATCGATGCAATCCAATCCACTCACTGTCAAAACGTTGCCGTCCCGTCTGATTAATTCAACCGAGCTCAAACCAATCGGGTTCGGGC
GGTTGGGTGAACGGCTGGCAAAAACACCACGCGCCGCTTCCTCTCCACGGGCGCGCCGAACAAGCGCGGTTCCGTCAGAGTGGTGCAGCCAGTACATCACCCAGAGATG
AGAGACCATCGTGATGTTCAAAAGCGCATCCGCATAAGGAGAGGCAATTTCCAGTTGGCTCTGGGCAGGGTTGCGACGACTGGACTTCGGGCATTCGCGCGTTTCTGAG
AAGCCTGAGCGCACAAAGCCAATCGGGTTCATGATGATGTCTGAAGACACCGTGGTTTCCTGCTGATTATTCATTTGTGAAGCCGTAATAGGTTGCAAAATTAAACACTT
AACGTCTCCAAAGCCAAAATCTACCTCAACTACAGATCGATACTTGGAGACGATATCCGAGTAAATTCGATTTATCGAATGGAGCACCTCAGTTTTTCTGAGGTCTCCCT
GATCAGCATATTCGAAAATGCTGAGGCTATGCCTCTGATAAATCGATTGGATTATTGGCAAGCGCAAAAATCATAGTTCTGGCACCTCATTTTTCAAAATCAGAATCACA
GGTCGCTATGACAGATCACACCCAGCACGATACAGCCCCAGCCTCCGGGCTCAGTCCCACGCCCGAAAAAATAAAATCCTCCTTTAACACCAAACAGACCTGGCAGTTT
CTTGGAGTGGCGATTGCTGCAACAATCGCCGGCCTTCTGGCCTCCGCTTATCTGCCTGCTGGCATGGCCGTTGCAGCGGTCCTATCCGTTGGTTGTCTCGGCCTATGGGCA
ACGGGTGTGGTGCCGGAATACTGGACCGCGCTCGCCTTCTTCTTCGTTGCCATCGTAATCGGAATTGCCCCGCCGGAAGTGGTATTCTCCGGCTTCCAGACCTCAACCTTC
TGGCTGCTGTTCGCAGGCCTCGTGCTTGGCGCTTCCATCAAGCACACCAAACTGGACAAACGAGCTGCCGTGCTGCTGATGCGTTTGCCGCTCAACAGCTATGCGGGGCT
GGTCTCCAGCATTGTCGCCTTCAGTGTGGCGCTGGCGTTTGTGGTGCCCTCTTCCATTGGCCGTATCATGATTTTGCTGCCGGTGATCTCCGCCCTTGCTGCGCAGATAGGC
TTCAAACCCGGCTCCAACGGGCGTACGGGTATGCTTTGTGCGGCGGCCTTTGCGACCTTCGCTCCAGCGTTCACCATTTTGCCTGCCAACGCCCCCAATATGATCCTCGCG
GGGATGAGCGAGAGTCTGTATGACCTCAGCATTTCCTATTGGGATTACCTGCGTCTGAACTTCCCGGTTCTGGGCATCCTGAAATCACTCATCCTTATCGGTCTGACACTA
GTCCTGTTTCCGGACAAGCTCCCTCAGGACGCGCGTGAGACCACTGAGGAGGTACCACCAGTTAGTGCTCCTGAACGGCGCCTGATGATCATTCTGGCCCTGTGCCTCAT
CTTCTGGCTGACGGACAGCCTGCATCCGATTGGTCCGGCTTGGGTTGGTCTTGCCGCAGCTCTCGCCTGCTTCTGGCCCAACGCCAAGCTGACCTCCAAAACCTGCATCAA
TCAGGACCTGAGCTATGGCTCCCTGATGTTCGTCGCGGGCATCATGGGGCTGGGGGCCGTCATCTCAGTTGCCGGACTGGGTGAGGCTTTGGTTCAGGCTGTTGGCGGCA
AAGCCGGTTTTGCCGAGGACACACCAATCTGGAACTTATCCGTCCTCACTGGCCTGTCGACACTGGTTGCCATGGCGACGAACCTGCCGGGCGTGCCTGCGGTGATGACT
CCAGTCGCCGGAGATCTGGCAGATGCCACTGGCCTGCCCTTGCAGACTGTTTTGATGACGCAGGTTATGGCGTTTTCCAATGTGTTCCTGCCGTTTCAGGCGCCGCCACTG
GTCGCCGCCATTCAGGTTGGAAACCTGCCACCAAAAGCCGTGATCAAACTGTGCCTGACGCTGTTTGCGATCAGCCTGTTCCTGCTTGTCCCAATCGACATGCTGTGGTG
GTACCTGCTCGGATTTCTTTGAGCACCCTCAGGCGGGCCCACACAAATTCTGAGGCCCGCCTTCCATTCTTTCGATTTGTGCAAAGCGAATTTCCATTCCTAGCTATAACC
GTTTTCAAATTGCGAGGTTTTAGTATGACGACCCGATTTACTTGGTTGAAGACAATTGCTTACAGCGGCCGTCTGGCGCTTGTGATTACTGGTGGAGCCGCCAAAGCAGA
ATTGGCTGCTGACTGGCAAGACACCGCTATGGGCCTTTTCAACGCTGAAAAGTGGCCAGGTGCCAACCGCCAGCGCGCCCTGACAATTTTGCATACTGCAATGTTCGATG
CCGCAAACGCAGTCGAGGGCACCTACGCTCCTTACGCTTACACTGGCTCAACCAATCCGGATGCATCTGCCCAAGCAGCCGTCACTCAGGCCGCCCTGCATGTTCTGTCG
ACCCTTATGCCAGAGCGTAAAGCGGAGTTGGAAATTGCAGCAAAAGAGCGCCTTGCAAGCCTGAAGAGCCCTCAGTCCGGCATTGCTCTGGGCCAAGCTGTTGCGGAGC
AAATTCTGGCAGCGCGCGCATCAGACAACGCAAACTTTTCGACAGATTACACGCCGGGTCCAGCAGAAGCGGGTGTGTATCAGCCAACGTCTAAACGCCAGATGGTTGC
ACCGAAAATTCGTGAGATGAAGCCATTTGTTCTGTCCTCTGCAACGCATTTCCGTGTGCGCCCAGCTCCGGCTCTAGAGAGCATGCAGTATCAGCGTGGTCTGAGTGAAG
TCGTAAAGTTGGGCGGCAAGGAAACACAGACTAATGAGGCCAACGTCTTCATCGCCAAGCTTCATGCTGGCTCCGGCTCCGGTGCCTGGAACCAGATCGCCCGCAAAAG
CAGCGCGTCCTGTGACCTGTCCATGCTGGAAGAAGCCCGCACACTGGCGCTTCTCAACCTTGCACTGACGGACTCACTGGTTGCTGGTTTCAACGCAAAGTATGAGTACG
CGTTGTGGCGTCCGCAAACCGCAATGGAAGCTCTGGGCAAGACCTATAAGCACCCAACCCTTAAGGCCGATTTGGAGTGGAGCTCTCTGGTGCCTGCGCCAATGCATCCT
GAATATCCATGTCAGCATTGCACCAGCGGTAGCGCGGCTCAGGAAGTCATGGAAGCCGTCTTCGGCGACGGCAAAATGACCTTCACCTTCAACGGCAAGGGCGGTCTTT
CCAAAGACTATGTCAGCTTCCAGCAATTTGCTGAGGAAGAGTCAGAATCCCGCGTTCTCGGCGGCGTTCACTACCGCCGTTCCAACGCTGTTGGCGACATGTTGGGCTAC
CAGATTGGCTCTTACGTTGCAGAGACCTCTTTGCAGCCTCTGAAGCAAATCAGCGACATCAAGAAATGTGCAGGCAGCCGCTAGGTCAGCTCCTCCTGACTTAGCTTGCA
CTTGCAGGCCGCACTCCAATCCCCTCGCGGCCTGCTTTTTTTTATCACCCCAGAAAAGGGAGCACTTGCAAGCGCCACTTCCGGCAAAGCACCCGCAGGTCTAAAGGAAC
AGGTACTATGGAAAATCGAACTGGTTTAGCACTCCGCATTCTCTTTGTTGGCATTGGCGCAACGCTCTTCATGGATATCTGGGCCATCGCAGCAGATAGCCTCTTTGGCAT
TCCGCAGAGCAATTTCTCAGTGGTGGGTCGCTGGCTGGGGCACATGCTGGACGGACAATTCACCCACGAAAGCATCCGCCGCGCTGCACCTATCGCAGGAGAAGCAGCC
CTTGGCTGGACGGTGCATTATCTGGTCGGGATCATCTTCGTTGCAGCCTTCGTCATGATTTTTGGAAGTCGCCAGTTGCAAAAACCATCTCTGATCCTTGCGATCTTGTTTG
GCGTTGTGATGATCGTGTTCCCATTCTTCGTTATGCAGCCATCTCTTGGCGCAGGCGTTATGGCTGCCAACCGGCCAAACCCAGACTTCGCACGCCTGAAAAGCCTCATGA
CGCACTTCTCATTCGGCCTGGGCATGTATGTAGCGCTTTGGGTTCAGACCAAGATCTTCAAACCAATTGAAGCACCGGTTTCCTCCGCATCCTGATTGTTTCAGGATCACG
AGAAAGCCCGCCGTTCATCTGAACGGCGGGCTTTTTGTCGGTCAGGAAAGCCGTTGAAGTTAGCTTAGCTCAAGCTCTCTGCGAACCTGATACTGAGGCAGATGCTTCTT
GATCTGCTCCACCACCACCTCAGGATCGGGCAGTGACTGGTGACCATGGATCTTGCCCATGTGCGGGTCAAAATTCTCACGCCATACCGGATCACACACCTCGTAACCAT
CTTCGCGGATCTGCGCGACATTGCGCTTTACCGAGGGCTTCTTCCACATGGGTCCGCCCATCACCGGGAAGAAGAGCACCGGAAACTCAGCAGCAAGAATGACCGTGGT
CAGGCGATTGAGTGAGGAACCATTGGCAACACCAGACAGCGTGTTGGCCGTTGCTGGCATCACGGCGAGAATATCGTGCTCGGCTGCGATGCGTCCCGGCTTGTCTGTA
GACCAGTCTTTCGGATCATCGCCGCAGATCACTCGATCCACAAACAACGCGATGCTGTCCATGTTCACGAAGTTGATAGCATTAGGCGTAAACATGGCGGTGAGTGTACA
ATCAATCGAATTCTTGATCTCTCGCAGATAGTTGGGCAGCAGCGTTGCATCCAGAGATGCGCTCGCCCCAATCAGAATGCGGGCCTTGTCAGGAGTGTCATTGGTGACCA
TGGAAAAATCCTCTCTGTAAAATCCATTTCAAGATGATGGAGGTGCCGCGCGATTGGCGCCTTTGGGTGAGCTACGGAACCCTAGGGGAGGTTGGTGAGTGCACCTTGGC
CAGCTCATCGCGCGGTGTCTTTTTCTGGTCGCCAGATGACCTCAAATCATGCCTCATATTCCAATTGTTTCTTTGAAATGCTCTCCTCAGGTTTGCTGAGGGCGAGGGCTC
AGATTTGCTGATAGGCCGCCTCAGAACTTCCGATGTGCAATGAGAATGTACTGCTGGCATGGTTCCATTCAGATAAGCCATCAGGCTGCTGGAACTGGAGGGATTGATGA
GGGGTGCTGAAGTCAAAACCATTTGCGTTGCAGGAGCCAGCGGGCTGGCCGGGTCTTATATGGTCAAAGAAGCCCTCGCACGCGGCTATCGCGTCAACGGCACCATGCG
TGATATTTCCCAGCTCGATAAAGTCAGTGGCTTGATGCGGTTGCCCGGCGCGGCAGAAAGGCTGCATCTGTTTTCAGCAGATACAGCCAACTCTGAAAGGTTCGATGCGG
CAATGACACAAGCGAAGGCTGTCTTCATCGCCTGCATTCCAACAACCAAACACGGCGTAGACGGGACGCCTGCTGCGGAGCTGGATCGCGAGCGCGGTTTGAAAGAGAT
CATCCAACCTGCAGAACAAGCCTGTCTGAACATCCTGAAAGCAGCACATCGGCAAGGCGTTAAAACAGCAGTGCTCTGTTCCAGCACAGCCTCCGCAGAACCTGCCACG
CCCTCCTTTGAGAAGCAGGAGGACAAGGACATTTCAGATGCCGAGATGCAGATCCAGCAAAAGAAATACGCACAGGCACAAAAGACTGTAATGGAATCTGCAGCAGAA
GAGTTTACCACGCAACATGGCATGAAGCTTTGCACCTTCTTGCCAAGCATGATGGTGGGTGCCCCTCTCCTGCAGTCGCATTTGAAAGGTCATGTCCTCAGCTTTCTTGCA
AAACTGCGGCACGGCAAGACGGGTTGGCACAGCACTGTTCCAGCCGGATCCATGTCCCTCACCTCCCCGCAGGATCTTGCAAAGATGTCATTGGCTGCATTGGAACAGG
CAGAGGCCTCCGGACGATATTTTGCTGTGAGCGGCAGTTGGAGTTGGCGTCGGATTTACCAGGAGATTGCCGAGTATGTTCCGCTGGAGGCACTGCCTGCCCCACTTGGC
AGAGATCCTGAACCGCCAACGCGCTTTGACTTCAGCAAACGCGACAGTCTCGGCGTTCCGATACAAAGCATTCAAGAGATGTTTGCCGATTTTTATGAGGCCTATGAGCA
GGTCAGGCTCGTTCCTCAAGCCTGATCCTCAGTTATCTCATCGGTACGCGCGCTTTTTGCACGCTGTTACGCCACACAAACTTGGCAGGCAGCTTCACCGTTTCACTGGAC

GGCTGCTCAGTGGTGCACCATTCTAGGCACTTGTCGATGATCTGATCTGTTGGCTGTGCAAAGGTGGTCAGGTCATATCCTTCCCACTCGGCCTGCTGAATATTATCAAAG
CCAATGACTGAAAGATCATCCGGTGTGCGCAGACCAAACCTCTGGCGCGCCGCATCAATGACGCCGCACGCAATCAGGTCTGTGGTGCAGAAAATGCCATCCGGGCGGT
CTTCACGATCAAGCAGTGCCGTCCCCAGTTCCAGACCTGCCTGATAGGATGTGAAACCATCCACCTCACACAAAACATCAACTCCGGCTTCTTCAGCAGCGGCTATGAAA
CCCTTCTGTCGCTCCACAATGCTCCAGGTGGCATTTGCGGAGTTTGCGAGGGCGATTCGCTTGCATCCCGCACGCAAAAGAGCCTTCAGTGCGGTCTTGCCTGCTGCCCTG
TTATCCGGGCCAATCTGTAGAGAACCCGGGAAATCTTCCTCACGATTGATCAAAACCAGCTTCATGCCATTGCTAAAACATGTGTCGGCTAAGGAAGGGTCGGGCGCACC
TGACAAAATAATCGCCGCTTCTGTGCGAAAACTGATGGCCTGAAGCAACGCATCCTTCACACGCTCCTCATATCCACTGGTGTTGATAAGCATTGTTACCTTACCAGCGG
CCTGTAACCTGTCAGAGAGTTCAGCGATCATCGCAGAGCGGTGCGGCGTATCAATCTCAGCAGCAACCAGAGCCACGATACCACTTTCAGAGCGGTTGAGACTGCTCGC
AAGGCGGTTGACGCGGTACCCTAGTATTTCTGCGGATTTCAATACTTTAGCACGCGTCGCGTCGGAAACACTAGCCCCCGGAGTAAAGGTTCTGGAGACAGAAACACGG
GACACGCCAGCGTGGCGCGCGACGTCCTCAGCCCTGACCTGCTGCTTTCTTGAAGACACTTTTTCCGACATCAAATTTCGCTGATTATTGTACACGTTCGCATCTGCATAT
CAGGAAATATCCAGAAAATAAATTAGTCTTTCTTGGAATAAAGTTGACATTGCAAAAATTTAGCCGACTATATTATGTACACGTGTACACAACGTGTTTGGGAGGAGTAA
ACATGAAATCCATTAAGACAGCTTTGTTGGGTGCAACCCTCATTGTGGGTGCTGGTGCAAGCGCGGCGTCTGCAGATGAACTCAATGTCATCTGCTCCGCTGACAACGCT
CTGTGCGAAAAGCTCGCAGCAAAGTTTGAAGAAAAGTCGGGTGCTGACGTCAGCATGATCCGCCTTTCTTCAGGCGAAGCTTATGCAAAGATCCGTGCAGAAGCTCGCA
ACCCGCAGACGGACGTATGGTGGGGTGGAACCGGAGATCCGCATCTGCAGGCAGCAAACGATCAGCTCTCTGTCGCTTATAAGTCGCCAATGCTGGAAGAACTTCTGCC
ATGGGCTGTTGCTCAGGCTGAAACTTCCAACTACCGCACTGTAGGCGTGTACTCTGGCGCACTAGGCTGGGGCTACAACAAAGAGATCCTCGCGGAAAAGGGCATCGAA
GGCCCTAAATGCTGGTCTGACCTGCTAAATCCTGCGTTGAAGGGCGAAATCTCTATTGCAGATCCAAACTCTTCAGGCACCGCATACACCGCACTTGCTACTCTCGTTCAG
TTGATGGGTGAAGAAGAAGCCTTCGAATACCTGTTGGCTCTGGGCAAGAACGTATCGCAGTACACCAAGTCCGGTTCCGCTCCTGTTAAGGCAGCGGCCCGCGGTGAAG
CAGGTCTCGGTATCGTGTTCATGCACGATGCAGTGAAACAGGCTGAAGCAGGCTTCCCAATCGAGGTTGTCGCTCCATGTGAAGGCACCGGTTACGAAATCGGTTCAATG
TCTCTCATTGACGGTGCGCCAAACCCAGACACTGCAAAAGTGTTCTACGATTGGGTGCTAACCCCAGAGGCACAGAACCTGATGGCAGAAACTGGTTCCTACCAGATTCC
ATCCAACAAGGGTTCTACCCAGCCAGAAGCTGCTCCTAAGCTCGAGAACGTACGTCTCATCGACTACGATTTCGCAGAATATGGAAGCAGTGACCGCCGCAGAAACCTG
CTGGCACGCTGGGACATGGAAGTTGGCGCTGCTGCAAACTGATTGAGATCGGGGCGAGTTAACGCTCGCCCCGTTTCCCTAAAAAACCTGGTTTTTCCAAATGAAAAATG
AAAATCGTCGGCTGGACCTCGCACTCATATGCGGCGTCCTTGGGTTGTTTATATTGCCTTGGTTCCACATTCGCAGCGGCTTTTTTGGCGGCGAATGGGTCTTGGATCTGC
TCAGCAAAGAGCGTTATTGGCCGGGCTTTTTCCAGCTCTTCCTTGGAAAGCTGCAACTGCTCCCCGTGTTTGCTCTTGTGGGTGCAGCCGGATATCTGCGCCTGAAGTCTG
AGCCGGGTCATCGTGGCAAACTGATTGCTATTATCGGTGCTCTTGGCCTGTTCTGGATGGTGCTGGAAGGCCTTTCCATTGGCCTGAGAGGCTGGAACTGGAGCCTGCTG
GAAGCAACATTCGGCCGCGTAGATGGTCAGCCCGCGTTTGGCGCCGGTGCCGTCTTGCTCGCACTCACGTTCACCACATTCATTTCATTTGGCCTGGCTGAAAAGGGCGC
TCTCAAAGGCGACGCGATTGTACTAACCGCTGTCTCGCTTTTGATCCTGCTTGTAGCGACCTTTGTATTCTACCCAATCCTATCCATGTTCGCTGGCGCGTTTCAGGATTTC
CGCGGCAACTTCACCACTGAAGGCGTCATGCGGAACATCACCGATCCGCAGATCTGGAAACTGAACTGTCTGGTTGGCGGTAGTTGCGGCATTGCATGGCGCACCATGTT
CCTTGCGATTGCAACAGCAGCTGGTTCCACCCTCCTGGGCCTTGCGTTGCCCTGATCGTAACCCGCACCCAGATGCCATTCCGCAAAAGCCTGCGGACCCTAAGCATCC
TGCCAATCATTACGCCTCCTTTCGTGATTGGTCTTTCTCTGACCCTGATGCTGGGCCGTACAGGCTTTGTTACCGAGGCGATTGAAGCTGCGACAGGCCTTGAGCTGGGCC
GCTGGATCTACGGCCTGCCGGGGATCTGGATTGCACAGATGCTCAGCTTCACACCAATCTCCTTCCTCGTACTGGTTGGCGTTGTTGAGGGCATCAGCCCATCGCTGGAA
GAAGCAAGCCAGACCATGCGTTCTGGCCGCTGGAGAACCTTCCGCAAAGTGTCGTTGCCTTTGATGGCACCGGGTCTGGCAAACGCCTTCCTGATCGGCTTCATCGAGTC
CATGGCTGACTTCGGCAACCCAATGGTTCTGGGTGGTAATGGGCGGCGTGCTGTCGACTGAAATCTTCTTCGCGGTCGTTGGTGCACAAAGTGATCCATCCCGCGCAGCCG
TGCTCGCAATCATCCTCCTCGGCTTCACACTCTCTGCCTTCTTTGCGCAGCGGATGTGGCTTGGTAAGCGAAACTTCGCAACCGTTACCGGTAAGGGTGATGGTGGCCGCT
ACGCAAAACTGCCACGGCGTATCGCAATCCCGGTCGGTACCATTGTTGCCCTCTGGAGCTTGTTCACCATTGCTGTTTACGGCTTGATCCTGTTTGGCGGCTTCGTGAAGG
TCTGGGGTCTGGACCACAGCGTGACGCTGGCGCACTACATCACCGCGTTCTCCATCTCATTCGAGAATGGCTTTGCGTTTACCGGTGTTGCGTGGAACAGCTTCTGGACCA
CCATGTGGATTGCCCTTGCCGCTGCTCCGTTGACTGCTGCAGTTGGCATCCTGTCTGCATGGCTCATCGTTCGTCAGGACTTCCCGGGCAAATCTGCCTTTGAATTCGGCCT
GATGATGAGCTTCGCAATTCCGGGCACCGTGATCGGCATCAGCTACATTCTGGCGTTCAACCTGCCACCATTGCAGATGACGGGTACCGCTCTGATCCTGATCTTCTGCTT
CGTGTTCCGCAACATGCCAGTGGGTGTGCGTGGCGGTGTGGCGGCCATGTCTCAGCTGGATAAGAACCTGGATGAAGCTTCCCTGACACTGGGTGCAAACAGCGCGCGG
ACACTGAAAAAGGTGATCCTGCCGCTGCTGCGTCCAGCCATTCAGGCGGCGCTGGTCTACAGCTTCGTGCGTGCCATCACATCCATCTCTGCGGTCATCTTCCTTGTGAGC
GCGAAATACAATATGGCGACGGCCTATGTGGTCGGTCTGGTTGAAAACGGCGAATACGGCGTGGCGATTGCCTACAGCTCCGTCCTCATCTTCGTGATGATCACCGTCAT
TCTTGGCTTCCAGCTACTGGTTGGTGAGCGAAAACTCCGTCGTCCAGAAAACACTCAAAAATCTCCTGCTCAGGGTAAACTGAAGGAAGCCGCAGCATGACACATTCTAA
ATTTGGTCAGGTCCGCTTCGAAAATGTCGGCAAAAACTACGGCGACTTTACTGCAATCCCGGATCTTAACCTCACGATCGAACCGGGCCAGCTGGTTACCCTTCTGGGGC
CATCCGGATGTGGCAAGACAACAACCCTGCGTATGATTGCAGGGTTGGAGACCCCAACATCTGGCCGCATCATGATTGGTGGCGATGACGTAACCCGTCTGCCACCAAA
CGAACGTGACGTGTCCATGGTGTTCCAGTCTTACGCCCTTTTCCCGCACATGAGCGTGCTTGAGAACGTGAAATACGGCTTGGAATCCAGCGGCTTCAAAGACGCAGAAG
ATCTGGCGCTAAAAGGTCTCGACCTTGTGGGCCTGCGCGAAAAAGCACCGCGTCTTCCTTCCGAGCTGTCCGGTGGTCAGCAGCAGCGTGTGGCTGTTGCTCGTGCGTTG
GTTCTGGAACCGCAGGTTCTGCTGCTTGACGAGCCGCTCTCCAACCTCGACGCCCGCCTGCGCCGTTTGGTGCGTACTGAAATCCGCGAAATTCAGCAGCGCCTTGATTTC
ACAGCGGTTTACGTCACCCACGATCAGGAAGAAGCGCTTGCTGTTTCCGACCGCATCGTTGTTATGCGTGATGGTCTTGTCGAGCAGGATGGCACGCCTTACGAGCTTTA
CGAGAGCCCAGCTTCCGAGTTCATCGCGGACTTTATCGGCGAAGCGAACATCCTCGACTGTGACATTAGCAGCGTGAATACGGACACCGCAAATGTGCTGGTGGGCGGC
AAAGGCTTTGACCTGCGCAGCCGTAACCACAGCATTGGTAAGGGTAAATTGGCCGCGCGTGCCAACTCCATCAAGCTCAGCAGAGATGTGAATGGTGGACCGGCCGGCA
AGATCAAAAGCGCAGCTTATCTTGGAGATCACTTCCAATACGAAGTGATGAGCGACCTTGGCTTGCTGTTCGTCATAGACAACCACACTGATGAACTGCTGGAAAACGG
AACCTCTGTTGGCATCGACTTTGATGCTCGCGGTCTCGCACTGATTGGTGGTAAAAAATGACCACCGAAACAGTTCTGGACCAAGCTGCCGAGTTAGAGCAACGCTTCAA
CGCGGCCCAAAAGATGGCCGATGAAGCGGCAGCGCTCGCTCTTGAGCACTTCGCCAATCGTGATGATTTGAACATCGACACAAAACGGCATCCGCAGGATGTCGTTTCT
GAAGCCGACAAGGCTGTAGAGCAAAAAATCCGCAGCCTGATCGCCGAGCACTTCCCTGAGGATGCGTTGCTTGGTGAGGAGTTTGGCTTTGTGGAAGGGACAACCGGTT
ACACTTGGGTCATTGACCCAATTGACGGAACTGCCCCGTTTCTCAATGGCTTACCAGGGTGGTGCGTGTGCATTGCCGTAACCGACAAAGACGGTGCTGCCATCGGTATC
ATTGATGTTCCGGTCCTCAAAGAACGCTTTGCTGCCCTGCGCGGTAAAGGCACCACACTCAACGGTTCCCCTGTTCGGGTGACTGATCGATTTAACCTGCAGACCGGCAT
GCTGGGCGTTGGCGCGAATGATCGCGTTGCTCCTGCGCAGGTCGGAAAAATGTTCGAAGACCTGATGTCCGCAGGCGTGAGCTACACTCGTTATGGCTCCGGTGCATTGA
TGTTGGCATGGGTCGCTGCGGGGCGCTTGGTTGGCTACGTTGAGCCGCGCATGTCTGCCTGGGACTGCATGGCTGCATACTGCCTTATCACTGAAGCTGGCGGGAAAATT
CTCCCATTTCCGATGGCTGAGAACCCAAGTGCCCCGTATCCTGTTTTGGGCGCTGCTCCGCAGGTGTTCGATGAGGTTTTGGACCTCACAAATCTAACCCAGTTTGCCTAG
TCTCCCGAGTAGACAACGGGTTAGATCATACTTGCAGGCGCTGTTCACCCAGAGCAGCGCCTGTTTTTTTATGCACTCAGATGAAGAGCAGAACACTGCCGCCCGCGATC
AATCCACCACGGATAAGTGGTGAGAACGGTGCGATGACTCCGTAGGAAATAGCGGAGAGCGCGAGCCCCACCTTCACTGTTGCCAACAGCGACCAAACAGGCGAACTC
ACAAGCGTGATCAGCGTCGGATTTGCTATCATTCCCAATGGGATAAGGTAGAGGCCAATCCCCAGCGCCATAGCGAACAAAGCCACCTTCAACCAGTTCTCATTGACCAT
TCCCGCAGCGATAAACACCGCTCCGCAAACAGGGGGAGTAATGGTAGAGAGCAGCGCGAACCAGAACACGAACAGGTGTGCCTGAAGCGGTTCTAATCCGATATCAAT
CAGCGCTGGCCCTGCAACAGAAACGCAGATGACATAGGCTGCAGTTGTTGGCACTTCCATGCCCAGAACCATGCAGGCGAGTGCCGTCAGCAGCAAAGCTGGCCACAAC
ATGCCGCCGGAGCCGGAGAGGATAAGTGAGGTGATCTTCACCCCAAGGCCTGTCACCCCAAGTGTGCCGATGATGATGGAGGCACAGAGGATAATCGACGCGATCATTG

CCACCTGTTTGCCGGAGTTCATCAGCGCCTGCTCATAGCGCCCGCCGATTTTCCGCAGGTCTACCTTGAAGTTCTTGTCGATCAGCAACAGCACAACCGCTGCCATAATCG
CCAGACATGCAGCATACTGAGGGGTATAACCCAGCACAAACATCGCCCAGAGCAATACCGTGAACGGCACGAGGAAGAACAGGGATGTGATGATAACTGAACGTGTGT
CCGGTTGATCCTCCGTTGCCAACCCCTTCAACTCATAGCGGCGCGCGAACGCGTTGATGCCCATCCAGACCGCGAAGAAGTACAGAAATGCAGGCAAAATCGCAGCGGC
CATGATGCCGGTGTAGGGAACGCCGGTCAGTTCCACCATGACAAAAGCGCCCGCTCCCATCAGCGGCGGCATGATTTGTCCCCCTGAGGAGGCAACGGCTTCCACGGCG
GCGGCAAGTGCACGCGGGTAGCCTAGTTTGGTCATGGCGGGCAGCGTAATCGCCCCGGTAGAGGCTACGTTGGCGGAAGCAGAACCAGAGATGGAACCGAACAAAGCC
GATGAAATCACTGAGACTTTCGCCGTTCCGCCTTTCAGCTTACCTGCGGCGGCGGCAGCCACATTCATGAAACCAGCGCCTGCTTCACCGGCATTCAAAAATGCACCGAA
AATGACGAAGATCGCGACAACACCAACAGATACTCCCGTCAGCTTGCCCCATAATCCCCCTTCAGCAATGGTCATGGTGCCGAGGAAACTTTCCAGCGGAACGCCAGCA
TAGCCGAACTCACCGGGGATATATTGCCCAAACATGCCATAAGCGAGCGCGATCAGTGCCACCAGCGGAAGCGGCCAGCCGATGGCCCGGCGCGCCATTTCCAGCACCG
TGATCAGCAGGCCAAATGCCAGATAAAGCTGAAAATCATCAATAAGATAGCCGTATTGATCCGACAGATCCGAGTGGTTCCATGCAATCCAGCCAGCCGCGAAAACACC
AAAGGCTGTGAGCACATAACCCGTCACCTTTTCAAGGGGATTACGTGCCATCAGGATGAGCGCCCAAGGCAATGCCAGTGCCATGTGTAACGGACGACTGACCAGATTG
GGAACCAGACCAGAAAAGATCAGACCCAGATGAAACAGCACGGTGGCCAGACCGAGCACCAGCCAGACGGAACGGTGAATTCTATTCAACGAAATAACCCAGAATGA
GTAAGGCCCGAAGAATATACTCCGGGCCTGAATTTTTGATAAGATGAATGCGCTTAGCGCTGCGCGTCGGTGATTGTGTACCCAACTTCATCATAAAACTTCAGCGCACC
CGGATGGATTTTGCCGCTGATGTTGCCGAGCAGATCCTTGTCTACGCCGTTCCACCATGCTGCGCCCTTGCCCATTTTTTCCTTCTGGTCCCAGAAGCTTTTGGTCAGCGCA
TACGCGGTTGCATCATCCATCTGTGTTGTTGTGTAAGCCACAACCGGCAGAGAAGTTGTTGTGATGTCAGCGTCTTGACCAGCATAGGTTCCAGCCGGGATCACCAGTTT
GGTGCGCTTGGTCTGGGCGATCTGATCATCAGAAAGAGACAGAACATCTACACCGGTACCAGCAGCAGCCTCAATCACGTTTGGCGCCGGGTAAGATCCGGCGGTCACG
AAACCATCAATCTGACCATTCTTCAGCGCAGGCACAGCGTTGGAAAGTTCCACATCCGCCAGCTCTACTTCACCTTCAAGACCAAACAGCTTCAGATACTTAGCGCCTTC
GCGCGCACCAAAGGAGCCTTTGCCGATGAGGATTTTCTTACCCTTCATGTCCGCAAAGCTCTCAATGCCGCTGTCTTCACCAACAACAAAGTGCATGGTCAGGGACGGGA
TTGGGAACAGTGCACGGATTTCACCGAACTTGGGATTTGCCCGTTTTTCGAACATCGCCTTGCCGCCTTTGGCAAGCTTCACCAGTACCGGCGGCGTGGTGAAAACGTAG
TTGCCCGGACGCACCGCAGCTTCCATTACGTTTTGAACAGAACCCTGACTTTCTTCAACGGTAACAATGATATCCCCGTTTGAGGCTTCCTTCATTGCTTCCGCAATCTGT
ACTGCCATCTGGTAATAGGAGGATGAGGTTTTCGCTGATTTGTAAGTGACGCGGGTTTCCGCCTGAGCGCTAAACGAAAACGTGACGAGTGCTGCAGCAGCCACGCATG
AGTGTAATAGTTTCATGGATCATCCTCTGGAACGACAAATACTTAAAATTGCTGAGAGATTTTATCGAGTATGTGGGGTGTTTCAAGTAGGGGGATCAGCTTACCGCTGA
TCTGAGGCGGAAATTTTCTCATGAATGCATGACAGAACGAACAGCTTGCGCGATTTTGCGTTTTTTACGTGAAATCTGCGTTACTTGCGTTTTGGGTGCGTTCATATGAGT
GTTTGTCGTGCAATCATATCAGAAGGCCGTTTTTGTCTAAAGTCTCCTGAAACTCTTCCTTGGGCAACAGCCGGAAATGGTCCCCCGTCATCTCCGCCGTCAGAATTCGAT
CGCACCGAAATGTGCGCGGCGCATTTCGCAGATGGTCAAATGCGATCACGTACCAAACCGGATAATTTAGCATGAGGTAGTGCGGTTCGATCCGCCGTTCAGAAGATGC
ACCATCCTCTCGCTTATATCCGACCGTGAGCGTCACCTGATCGATGAAGGCCTGATGCAGTGCCTGCACAACATCTTTGGGCGGAGCATTGATCGTCCGCTGCACATAAG
TGGAGGCTGTCGACCCGATCAGGATGCGGGATTTCAGACGGTTTACCTTATGCCGCTTTTCCGGCGAGAACGACGCTACCAGTTGCCGTCTGACCGAACCGAGGCTCGCC
AGAAACATGGGCGAGTCCATTTGCTCTGAGACCGCGATGCTGATCAGCAGATCGACCGCTTCCGGGTACGCCAGATTCATGCGCCCAACACCCCAGTTGCGATCGAGGC
GGACACCACCACCGCGACCACGATCAGCATCAATAGGCAGGGCCCTGCTCGCGCATCAGCTGAAGGTCCCGAGTGATCGTGCGTTCGCTGACATCATGCTGAAGCGCCAG
ATCTTTGACTGTGCAAAGCTCTTCCTGCTTCAGCTGAACAGCCAGCAATTCCATCCGTCTTAATCTGCCATTTGAATTTGTGCGCGCCATCGAAAAAATAGGACATTAAAT
GTCCTATTTGTCAATACAACCGAGGTGACCTGAGTTCAACACATGGAGATCACACGATGAAAATTGGCTATTTCACTGTCGGAACAAACAACAGAGAAGCCTCCGCCAA
ATTCTATGACTCACTCTTCGAGCAAGGAGGACCAAAGCAAGTTGCCGGCCAATGAAAGGCTGACCTATTGGCAAAATGAGAATTTTGCATTTGCGACTGCCTTGCCCATTG
ATGGGAAGCCTGCGACGCATGGAAATGGCGCAATGGTTGGGTTGGCGTCGACACTGTCGAAGAGGTCAAAAGGCTCCACAAACTTGCCATCGACCTGGGCGGCACCAA
TGAGGGCGACCCAAACCAGCGAGGTCCTTACATTTCAGCTTATGTGAGAGATCTCGACCACAACAAACTCATCTTCTTCACAGAGGCAATGAGCTGAGCACCCTGCCCAC
ACCTCAGACTGTCTTCATCCCCCAGAAGCGGTCATGAGCTGTGAGAATGCAGGTATTGCAACCTTTATACATAGGTCGCCTAATGCCTGCGGGAATGATGCTCATGTGCA
GCGGTTTCGACGCGAAATCGCTGCACGTTTCAGCTTCAGTTGGGGGACAATATGAACGCGCCAAACGCCGATCTGCAAAACAAGATCAACCAAATCAGAGAGCAACTGC
AACATCTGAAAACGACTGATACAAAGCCAAAAGCCGTCGCAGGGCCTAGCCTGAACCAAGCCTTGCAAGCCAGCGAGCGAATGTTGTTCGGCGAGCTCAACAAGCAGG
GTGTTTTGAAGTCTTCGCCGCCACAAATCGATCTGGAAGGCGGAACCATTGAGCTCGGCCTGTCTCTCTTCTGGACCTTCAATCCACCCGAAGATTACGAGGATGTGTTTG
CAGTCGTAAAAGAGTTTCTGGAGTGGGAAACAGGCGGACAGCTAAAGGCACCTCTGCTGGATGATTACAGCACCATCAAACAGCAAATCGACCTTGGCGGTATTGTCAT
GTCAGACGGCTCAATTTTGGGCTTCGGCAAATATGAGATGGCCGATGTTGGGTGGGCCGTCTGCCTCGTCTATGTGATTTACTACAGCCTCTTCCCAAATAAAAAGGCGC
CCTTCAACACCGCGCCAGCAGGCGTCATCGATCTGGCGGCGGAGACCTCTCAAAACGAGGTGAGCATCGCAATCCTTGGAGATTGGGGGACTGGTGATTACACCAGCGA
CGGCGGCCCCGCCAAGGCCGTGATGGACGCGATCAGCAAGAGTGCCACGATTGATAATCCTGACTACATTATCCATCTCGGGGATGTTTATTACGCTGGCACGGAATCCA
GTGAAGAGAACGATAACCTCCTCCAGATGTGGCCCGCAGACCGCTTTGCTGGTGGCAAAATGGCCGGAACGTCCTTCACACTCAATTCCAACCACGAAATGTATGATGG
AGCGAACGGTTACTTTAAGATCGCTTTGACACCTGCAAACACGCCTTTCAACGAGCAACAGAAAGCCAGTTTCTTTGCGCTAAAGTTTGGCGACTGGATGATGCTGGGCC
TCGACAGCGCCTACTACTCGTCCGCCTCAACACTCTTCATGGACGGCACCTTGGGCGGCAGTAAATCGAGCGTCCAAAAAGACTGGTTACAGCAAGTCGTTGGAGGTAG
GAAAGGCAAGGATGTCATTCTCCTGACCCATCATAATCCGATCAGCTTTGATGGCCTTTCAAGAGAACCACTGTGGGATGAAACTCTTGCAGATCTGGGCGGCACGCCAG
ACTATTGGTACTGGGGTCATGTTCATCTGGGTGTCGTCTACGGCGACCAACTGCCCGCTGCGAAAGGTCCTGATGGACGGTTTGCGAAAGTAAGATGCTCTGGAAATGCA
GCGATCCCGAATGGAGCTCCCTCCGGTCTGGATGGCAAACCAAACATAGATTACTACTCGCATACGCCATTGAAAAATAAGGAACCTCGGGTTCGCAACGGGTTCGCGG
TCATTACACTGTCTCTTGATGGGCAGATTAAAGAGCGCTTCTACGAAGTCTCTGATGGTTCAGAAGAGCCCGTCGAGGTCTGGAAGTCATCACATCAGAAAACAACGGTA
TAAATGTAGTATCAGCCTCCATGTTCAGGTGGGAAAGTAGCCTGCGTGAGCATGGAGGCTCATTGATTGAATGCTATCTCAGCACGTGCTCCACTGCGCGCCCTGCCGCC
AGCACCTGCTCATCATGGACCGCAGCACCACAGATGGAAATGCTGTGAGATGGCTTGCTCGGATCAACCGGCAGAGTGAGCGATGGCATGTTCAGGAAGGCGGTGATCA
TGGTCAGTCGCAAAGCCTGAGCGTTCGCCTGCTGGAACGCCTCAATGCTTTCCTCAACCTCAGAAATGAAAGGTGGCAGGTCAGGAACAGTCGGCATCAGCAGCAAATC
ATCTTTCAGCTGTGCTTCCACCTCTGACTGAAGCACAGCTCGTTGCCGACGCAGGTAAGCCAGCTTCTTTGGCGTCATGGGCATGGCATGTTCCAGCCGGGCTTTGATCTG
GGGATCAATCTTTGTATCACCTAAATCGACGAACTGAGAGATCTGGTCCCGCGCTTCAACGCTGACCAATGTGCCGTGCTCCGCAAACAGCAGGTGCGATTTCAAAATTG
GTCGCAAATCACGTCGGACGATATGGATACCAGCTGCAACAAGCGTATCAATGCAATGCTCAAACATTTCCATAATCTCAGCGGAAATGCCAAGGCCGTCCAGATCAGG
AACAACAAGACGGAATGCTCGTTCCTCTACATCCTTGTTTTCAGGCAAGGTTTTGCCGGCCATCACCGCATCAAGCGCGCAGATGCCATTAACCGTTCGCGCCATAGGTC
CCAACACATCAAGGCTCGCCCCAAGCGGCACAACACCTTCCATAGGATACCGCCCCATGGTCGCCCGATACCCGCAAATGCCCTGAAACGCAGCGGGCACACGTGTTGA
GCCAGAAGTGTCACTGCCAATCGCCAGATCACAAATGCCGTTGGAAATCGCCGCCGCGCTGCCAGAGGACGAGCCACCCGGAATTAACGCCGCATCCAACCCGTTGGGC
GCATTGCCATAATGGGGGTTGAGCCCGATGCCAGAAAATGCAAACTCGGTCATGTTTGCGGGTCCGATGGGAATGCCCCCTGCTGCAATCACCCGCTCCACAACCGCTGC
ATTCTCAGCCTCAACCTGCGTGGAATAAACAGCAGAGCCAGCAGTCGTTGCCCACCCTTGCAGAGCAATATTCCCTTTAATCGCAACAAGCTGTCCATCTATGGAACTCA
GCGGCTGACCTGCACCACGGCGCGTATCCGCCGCCGCCACTAACTCCTCTACCTGTTCCCGATTGAGTCCAATAAAGATACCCCGGAGTTCCTCCCCCCCAAGCCTCTCCC
AAAGCCCCGCAAGATCACCGATATCCACCATCCCAGAGCTCCATCCAACCAATCACCCCACCCATACTAAAACCAGCAATTCCAAGGTGTCAAAGCTGGAATGAGGGAG
AGCCGATCACATATTACAAACGCTGGAGCCGGTTCCCCAACCACAAAAAAGGCCTCTCAGATGAGAGGCCTTTTGGTTTTCTTGTCAGGTAACCCAGAGATGCTTAGTCG
TAGAGGACGGCTGCGATTTCTGGGGAGTCCATGTTGGAGGCGTCGTAGTAGTAGAAGCCGGTGTCGATGTTGGATGGGACGTCTTTGCCCTGTGAGGCCATGACGGCTGC

CTTGACGGTTTCATAACCGATGCCGACCGGGTTCTGGGTGATCGCACCTGCCATCAGGCCGCTGCGGACCGCATCTTTCTGGGCGCGGCCAGAATCGTATCCGATGATCA
CAACATCTTTGCTGCCCATTTCACGCACGCCGTTGGCAACACCGATGGCAGACCCTTCGTTGGTGCCGAAGATGCCTTTGATGTCCGGGTTAGCTGTCAGGATCGCTTTGG
TCAGCTCGGTAGATTTGAGCTGGTCGCCGCCGCCGTATTGAACGGTGACAACTTCGATATCAGGATAGTTCTTCATCTGGTTGAGGAAGCCATCGCGGCGGTCGATGCCG
GTGCGGCTGGTCTGGTCGTGCGCAACCACGGCCACCTTACCGCTTTTGCCGATCAGCTCTGCCATTTTGTCCGCCGCAAGGCTTGCTGCTGCAACGTTGTCAGTTGCGGCC
AGCGCCACCGGAATTTCGCTTTCCACGCCGCTGTCGAATGCGATGATAGGAATGCCAGCAGACTGTGCCTGACGCAGCAGTGGAATAGCAGCCTGACTGTCCAGCGCGG
CAAAGCCGATTGCTGCCGGGTTTTTCGCCAGCGCTGCGGCCAGCATGTCGATCTGGCGATCCACCATGGCTTCGCTGTCTGGGCCTTCGAAGGTGATTTCCACATTCAGCT
CAGCGGCTGCATTGTCGGCACCTGCCTTCACAGCTTGCCAGAACTGGTGCTGGAAGCCCTTGGAGATCAGCGGGATGTAAACCTTATCCGCTGCGAATGAAAGCGAGCT
GGTTGCAAGCAATGTTGCAGAGGCCAATGCGCCAATCAGAGTACGTCTTGTAAACATAACTATTTCCTCCCAATGAAATAATTTAGAGCGCGTTCCGTTGGATTGGATTC
AATCAAACGATAAGAATTCGCTCCACTAAAGATGCTGGGGCAAAAGCGCCCCAACCGAACTTATGCATGAGCTTTTTTCTTACGGGCCATGTCCGCATAGACGGCCAGA
ATAATGATCGCGCCAGTCACCACGGTCTGCCATTCCTGAGCAACGGAGAGCACACGCAAACCGTTGGTCAGAACGGCCATGATCAGTGCCCCAATCAGCGTTCCCATGA
TGGTGCCGCGACCACCGGCCAGAGACGTTCCGCCAATCACCACAGCAGCGATTGCTTCCAGCTCATACCCAAGACCAAGCGCTGGCTGAGCGGAATTGAGGCGGGACGC
GATGAGAATTCCCGCGATGCCGCAGATCCCACCTGCAAGCGCGTAAATTCCCATTTTCCAGGCATCCGTGTTGACGCCGGACAGCCGCACGGATTCCTCGTTGGAGCCAA
GTGCGAAGGTGTAGCGGCCCAAAACAGTTCTTCCCAGAACGTAGGAAGCTGCAATCGCAACAAGGAACAGGATCAACACACCATTGGGGATCTGAAAGCCCGGAAAGA
CAGCACCGATCACGGAGCCGCGGGAAATGAGATCAAACCCCGGCGTGTCGTTGAAGTAGATTGGACGTGTGCCGGAGATGACGAGGGACAGGCCCTTCAAAATCAGCA
TCATGCCAAGCGTTGCGATGAAGGGTGGCACCTTCATTTTGGCAACGAATGTGCCGGAACACAGACCACAGAGTGCACCAGCACCAATGGCCGCGACAACACCAAGCGG
CAGGGGCAGGCCCCAATATGTGAGCACCACACCGGCAATCACGGCGCAGAAGGTCATCATGGTGCCGATGGACAGATCAATACCGCCGGTGATGATGATGAGCGTTGC
AGCAATGGCCAGAACGCCGTTGACCGATGTTGCCTGCAAAATGGCGATCATGTTGGACATTTGCATGAAGTTTGGCGAGGCCAGACCAAAGCCAATGAGCAACACAATC
AAGCTGGCAAAGGCGAGCAGCTTCTGGTGCGCCCCGGAACTCACGAAACGACCGATAAGGTTCTTTTGCGGGGTTAAGGTGGTGTCTGTCATTGTACTTCACTCACAATT
TGCGCTGCTGGGTCGCGTTGTGTTGCGAGGCGCATAATGTCTTCTTGCGATGTTGAACTGCCACCCGGCAAAATGCCTGTCAGGCGGCCTTCACACATCACCGCGATGCG
ATGGCTGAGGCGCATAACCTCTGGCAGCTCTGAGGAGATGACGATGATCGCGCAACCCTGCTCTGCCAGCGAGTTCAACAGCTTATAGATTTCAGACTTCGCGCCGACGT
CGATGCCGCGTGTGGGCTCATCAAAAATGAGGATGTCGCAGTCGCGCAACAGCCACTTTGCGATGACCACTTTTTGCTGGTTGCCGCCAGACAGGAACTGAAGCTCCTGC
TTGTCGGAGGGCGTTCTGATTTCCAGTTGCTTGATGTACTTGGCGGCAGCCTCTGCCATGCCAGCTTCATTCAGCACGCCAAGGCGATTGCAGAACTGATCCATACTGGCC
ATAGCCACATTGGCACGCACATCCATACCCGTTGCGAGGCCGAAGTGCTTGCGGTCCTCTGACAGATAGCCAATGCCAACCTGCACGGCATCCGACGGGCTGTTGATTTT
GCACTTACTGCCATCGACCCAGACTTCGCCTGAGTCCAACGGATCAGCGCCAAAGATCGCGCGGGCTACTTCTGTGCGCCCTGCCCCCATGAGCCCTGCAAAGCCGAGGA
TTTCGCCTTTCTTGAGCGCAAAGCTGACATCACGGATCTCAGTGCCCCGGTTCAGGCCTTTCACTTCCAGCGCAACCTCTGCCTCTGAAAGGTCAGGCACGTCGACGGCTT
CTTCTTTCACCTCGCGGCCAACCATCATGGAGATGATCTCATTGACCGGAGTGTCAGCGGCAGGAACGGTGCCCACATACTCGCCATCGCGCATGACGGTGACGCGATCG
GAGATCTGTTTCAGTTCGTCCATCTTGTGGCTAATGTAGATCATGCCCACGCCTTCCCGCTTCAGACGGCGGATGATGGTGAACAGCTCCGCGATTTCCTTGTCGTTCAGC
GCGGCTGTCGGCTCATCCATGATCAGCACCTGCGAGCGATAGGAGAGCGCCTTGGCAATCTCCACCATCTGCTGTTTGGCGATGGTCAGGTTGCCGACAAGCACATCCGG
GCGCAGGTCCAGCGCCATGCTCTTGAAGATCTCAGCTGTCTTTTGGTTGAGCGCAGCCTCATCGAGACGTCCAAATCTCAGGCGCGGTTCGCGGCCAATAAAGATGTTCT
GCGCTGCTGTCAGATCATTCATCAGGCTGAGCTCCTGATGAATAATGCCTATTCCGAGGTCCTGCGCTTCTCGCGGCCCTTTTGGGTCAGCCGCTGTACCGTTGATTTCAA
TCTTTCCGCCGTCACGGGTGTAAACACCGGAAAGCACTTTCATCAAGGTGGATTTGCCTGCTCCGTTCTCTCCCATCAGGGCGTGAACTTCGCCGGCACGCAAGTCAAAC
TGCACATTCTTCAGGGCGTGCACACCCGGAAACGTCTTTTCAATACCCGTCATCCTGACGATGGTATTCATTTATTTTCTCCCCCCAGTGCTGGCCTGCAAACCTCAGTGA
GAACGCACAGATCCCGACTGGTGCTGCGCTTGCAGATCAGCAAAATGAAATGGTTGAGACACCCGTTTGATACCTGACTGCTGGCGCTGCTCCTCCCAGAGTTTTGCCGC
AGTCACCGTTGTCTCATTTTTGCGCGCCTCCCCATCCGAAAATCGGATCCACTTGGCGGCGAGGCGCTTTAGTTCTGAATGACACCCTTCTTGATGATGACGTTGCCGTAT
TTGCGTGTCGTGCCTGTCACCACAACGCAGAAAGCTTTCTTGGCTTTTCATAGAAGGCAAACCGCTCGACAAAGGTGATCTCAGCACCTGCTGGCAAAGCACTCTGGTA
GTCCTCCACCAAACCATCCGGCAGTTGGTCACCTTCCACCGGGGCCATCATCATCACATTTTCGGGGTCATACTGATCCAGCTCCCAAAGCGGCATGACGGCTTTGAGCA
GATCCGTCGCCTCGCACCCATCGGCGCGCAGCACCTTGTCACAAACAGAGTAGGCCGGAAAATGCGCATCGCTGAAGATGATCTCATCTCCATGTCCCATCTCATGCAGG
ATCTTCAACAGGTCCGGGAAAATGATCGGGTTGATCCCTTTAAGCACTTAAGCCCCCACTCACACCATGTGCCCGGCACCAACCGGGCAGGTAAAACAAACACCAAGAC
TTCCGGCAGGGTTGAAGAAAGCTCAGAAGTCCGCCGCAAATAAATGTACTATGAAGAACCAGAAAGGCTGTTCGCCTCATGATTCCGTCTTCAGGATGACGCTGTCGTGT
TCCGATAAACTCTCTCAGGACCAGCCCTTCAATTTCCTCGGCAAGAACGACGTCAAACCATTGCTAATTCGTGTTAGCAAGTTAGCAAAGCTTTCCGTTTAATCAAGAAC
AATTTTTCAATTTGCCGATTTTGAGCGCATTTTAAGAAACGGCCCTAACGTCAGATTTACCTGAATTTTTAGTTGGTCGAAAACCAGTAATAATAGCTTCAGGCCAGCAAC
TTATACTGCGCAAGTAACCATAAACCGGCGCAGATCACGGCCAGTACCAGAGCCTGATACCGGTAATAAAACGTGATGCTTCAGGAACAAAAGCAAATGCCAAAAGCG
TACTGGGTTGCCCGCGTCAACATCAAAGATCCTGATATCTATCGGTTCTACATGGAGGCTGCGCCCAGTGTCCTGCGAAAACACACTGGCCGAATCCTTGCGCGAGGCGA
TAAGGCTGAGCAACTGGAGGGCCAAGAATACCATCGGCATGTGATCGTAGAGTTCGACAGCATGGAAGACGCGCTAAACTGCTACCACTCCCCAGAATACCAAACTGCC
CGAGAATTCCGAAACGGGATCGCCGATGTGGAGACCGTGATCGTCGAAGGGGTGGACTAATCCGCAACCGACATCAGACACAAAAAAGCGGGTTGGTCTTGCGACTTAC
CCGCTCTCGTTTTGCTACGTGCTGACTAAAGCAGTCCGCATTTAAGTTGAAGCATCTGAACTGCTTTAGCCTTTGTTTTTGAGTATCTTTGTCCGAAAACTGGTTCCCACTT
TTCGGCGATACACTGTAAGTTTACTCAGGCAGGTTTTGGTGCCTTGAAGGCGGATTCGGTGAGGTGCTCCAGCAGCAAACCGGATTGCTTGACGGCTTCGTCTTCGTTGCC
CATGGCGACCGCACGCATGATGGCCTCGTGGTGGTAAGAGCACAGCAGATTGTCTGATTTGTCTTTGTTCGCAAACCAGAACCGGCGTGACTGGCTCTGCACACGGGCGA
GGGAGTGATGGAAATATGGGTTGTTGGTCGCTTCCGTCACGATACGGTGAATGTCCTGCAGATTGGACAGCAGCGATTTCTCATCCTCATCCTTCGCGCTTTGCACGATG
AGCGATCCAAGCGCCAGCATCTCGCGGCGCTGATCAGTTGAGGCGCGCATGGCCGCAAATCTCAGGCAGATCGGCTCAATCTGTTTGCGCACTTCCAACAGCTGCAATTG
CTGCTCTGCGGTGATGGAGAGAAACTGCACGCCTTTGCGCGGATGGATGGCCAGCATCTCGTCCTGTTCAAGAGCCTGAAGGGCTTCACGAACGGGCGTACGGCCCAGA
TCCAACAGAGAGGCGAGTTTGTTCTCAGAGTACAGAGCTGAAACATCAAGCTCGTTGAAAAACAACTTTCTCTCAATTTCCGTGCGCGCAAGCTCAGCCTGAGTTTTCTG
GTGAGCCACGGTTACCCCCTGTTTTAAAAGCGTTATTTCTGCGTCAGTTGTATCACAGTGTAGCACAGTGCATTCCCGATTTGCCAGCGGTCTCTTCTTTTCTTTTGCTGCT
ATGTCAGCCAAATGGATCTTGTTTAGCGGTTTTTCAATCGGCTGCCATGTCAGTTCTGGACCCCTTCTAGATAATGTAATCCATCAGACTCCGGCATTGTGTCATTCTGTCA
CGCAAAAACTACATAGTTTTCATGGGTTCCTCGCGCCTTGCAGATTTTAGATAAACTAAAAAAGCTAAAGTAATTAGTACTTTATCGTGAACTGATACTCTGAAAAATGTT
GAACTTCTCTGGCATGTCAGTAGATCAATATTGGTACGTCAGTGCTCATTATTTCGCAGTTTTATCGATTTCTAAGGGTTATTACCATTGCTACAGCCGCCTGTTTTTGCT
TTTAAGGTTGCTAATATGTCAGATTTGAGGATATGTTTCTGAAAACAGTCCTCATGGACTCAGAGATAAAGCAAACAGCCATAAAGTGTGGCATACAAGGTAAGCACCC
ATGAAAAACTGGAAAGAGACTGCCCTAGCAGAAACCGCATCTCCTGTTCTTGCTGAAGCAGATGTTGTTGTTTGCGGTGGTGGCCCATCTGGTGTGGCGGCGGCTGTTAT
GAGTGCCCGCAACGGCTTGAAGACAATCCTCATCGAGAAAAACGGGTTTTGTGGCGGCGGTGCGGTTGCCGGCCTGTCAGGTACTATCTGCGGTCTGTTCCTCACCCAGG
ACGACATCGAAAACAAAAACGCCAAGCAAATTGTTTTCGGTTTCGCAGAAGAGTTTCGTGCTCGCCTGCTGGCCAAAGGTGGCGCCACCGCCCCACAAATCTACGGCAA
CACCCACGTGGTGACTTTTGACCCGCTCATCTGGCGCGAAGCAGCGGACGATCTGCTTGAGGAGTCCGGCGTTCAGTGCCTCTACCACACCTACATGACCGAAGTGGTGC
GCGATGGCGACGACATCACCGCGATCCGTGTGGAATCCAAAGCAGGCAGCACCTTCCTCAAGGCAAAAGCATTCATTGATACGTCAGGAGATGCGGCTCTCGTGACGAA
AGCGGGGTGTGAGTACAGCTTCGGCGATAATGGTGCCATCCAGAACCCAACCATGATCTTCCGTCTGTCAGACGTGGTGGAAAGCGATTTCTACGATTACTTCGGCAAGA

ACACCATCTGCCCGCCTGACATGATCGAACAGCTGAAAACAGCGTTTGAGAGCAGAGAATACGACACACCCCGCAACAAGGTCTGGGTGTTCCCGACCCCGCAAAAGGG
CGTGTTTTTGATGAACTGCACCCGTCTAGCAGGACAGGACAACCGACTGCTGAACGTGGTGGACCCTGCTGATTTCACATACGCGGAAGTCTATGGCCGTCGTCAGGCCC
GCGAGTATCACCGCTTCTTCAAGGACAAGATCAAGGGTTTCGAGGAATCCCAGCTGATTGACATGCCACCGGAAGTGGGCATTCGCCAGACACGGACAATCAGCGGTGA
AAGCACCCTGACAAACGAGCAGGTTGCAAACTGTGACAAGCCTGAGGATGGCGTGGTGCGCTCCTCCTGGCCAATCGAAATGCACGCCGGTGAGAAGAGCAAACTGCA
CTGGCTGATGGAAGATTACTATGAAGTTCCATACGGCACGCTGGTGCCGAAGGAGCTGGACAACGTGATTGTCGCGGGACGGTGCCTGAGCGCTGAACATGAAGCCCTC
GCTTCTGCCCGCGTTACTGCCCAGTGCTTTGAGCTGGGCCATGCGGCGGCTGTGGCGACAAAACAGAAAATCGAAAAAGCGATCAATTACAGAGACGTAGACACAAAAG
AGCTGCGCGAGGCGATGCGCGCGGCTGGCAGCGCTCTGTAAAAAGTCGGGCAATAAGGAAGGTGCCTCAAGGCACACTGGGAGTTAAAAATGGAAGACTGCAGAATTA
GCCGGGTCGAGCTTTATGCTTTGGCTGATGAAGATGCGCCCCCTATTCCCTGGGCGGATAATCAGGAACCTCTGCTGTACACCAACAACATTGTACGCCTGTTCACCGAG
GATGGAACCGAGGGTCTTGGCGCAACTATGAGCTACACAGAGAACTTCTTTGATCGGTGCATCATTGAATCTCTGCGGACTATCGTTCCGGGCCTCATCGGCAAAAATCC
GCTGATGACCCAGGAGCTCAACAACTGGCTTGGAGCACGCTGCACCTGGGGTGGACTGCCAGCGAAGTCTCCAATCGACATTGCAGCATGGGACATTAAGGCCAAGAAA
GCCGGAATGCCGCTTTACATGCTGCTTGGCGGCGCACGGACCAAAATCAAGTCCTACGCCTCCACACCAATGTTCGACACTGTAGAGGAATACTTCCCTTACATTGACGA
TTGCATTGCCCACGGCTTCACAGCAATCAAGCTGCACTGCTACTGCATCTATGAAAAAGATGTGGCACTGGTGGAAGCTGTGCAGGCCAAATACGGCGCTACCGGCATCC
GTTTCATGCTGGACACCGCTACGTTCTACACACCTGCACAAGCCATGAAGATGGCCAAGTGGATGGAACGTCATGATTGGGAATGGCTTGAAGCGCCGGTTTCTGACTAC
GACTACAAGACCTATCAGCGCCTTGTCGCCAATACGGATCTTGAAATTTCCAGCCATGGCAACTGCCTGCTCACACTGCAGGAAGTTACGCACGCACTGAGCACTGGGAT
GTGGTCAGACGTGCGTCAGGACGCGACTGTTTGCGGCGGCATTACTCAGCTGAACAAATGCTTCGCAATCGCGGCCGGCCACAGCAAAAACCTGGAAATCCAGAGCATG
GGATACACCCTGACACAGGCTGCCAACCTGCACGTCGCTCTGGCTCACGACAACTGCAATTACTTCGAGCAGTTCTACCCTTACGAGGCGTTTGAGCTGGCAAGTAAAAC
CCAAATCCGCACAGACAAAGAGGGATATGTGCACGCGCCGGCGGGCAACGGGCTCGGCGTGGAAATGGATTGGGACGCAGTGAAGGAAGCATCTTTCGCTTCTTTCGTC
TTCGAATAATAAATGGGGGCATCCATGGCTAAACTACAACTGGATAAGACGAACTTCTTCGTCGGGGGCATTGCGGTACTCCTACCGACAATGGGCATTCTGGCAGCGG
ATATTGCTGCACCGGGCTCCGGTAAAGCTGTCATCTCTTCTTTGAAGAGCTATCTGACAAACGAATTTGGCGCGTTCTACATGTGGGCAGCGATCTTGTCGCTGATCTTCT
GCCTGAGCGTTTCTTTCAGCAAGTACGGCAACATCGTGCTTGGTGACCCTGATGAAAAGCCTGAATACAACACGATCTCCTGGGTCGCGATGATCTTCAGTGCGGGCATC
GGCGCAGGCATCCTCTGGGGTGGCGCCAGTGAGTGGATTTACTACATGTCCACACCGCCACTGGGTCTTGAGCCGGGCACCGAACTGGCCAAAGAATGGGCGATCACCA
CGGGCATGTTCCACTGGGGTCCAAGTGCATGGGCACTCTACTCCGTGTGCTCCATCACCATCGGTTATTACTATTATGTGCGTAAAGACCCGGTTCTGAAGGTTTCAGAAT
CCTGTAAGGGTGTTCTGGGCACACGCGCAAAAGGTCCTCTGGGCAACGTCGTTGATATCTTCATCATGATCGGCTTGGTGACTGCAAGTGCGACCAGCCTTGGTCTGGGC
ACTCCGGTTGTTGCTGCTGGCTTCTCCAAAGTGCTTGGCATTGAACACGGCACTTCCATGAACCTTGCAGCTCTTGCAATCGTGGTTGCGCTGTTTGCAGCAAGCTCCGCG
CTTGGCCTGAAAAAAGGCATGGCGAAGCTGTCCAACTTCAACATCTTCCTGGCTGTTGGCCTGCTGATCTTCGTTCTGTTCTTCGGCGGCTACTTCCTGACCACCATCAAG
TCCTTCACCACATCCGTTGGTGCGCTCGTGACAAACTTTGTGCGCCTGAGCACCTACATGGAGCCATACTCAGAAAACTCCACCTTTGTGTCTGACTGGACGATCTTCTAC
TGGGCATGGTGGGTTTCTTACGGTCCATTCCAGGGTCTCTTCTTCCGCAAGATCTCCCGTGGCCGTACAATCCGCGAAATGGTTCTGGGCACTCTGCTGTTCGGTTCCCTC
GGCACTGCAGTGTTCTTCGCAATCTTCGGCAACTACGGTATTGGTCTTCTGGATGCGGGCATCTTTGATGCACCTGCAGCACTGGCAGCTGGTCAGAGCTCCTTCACCGTG
ATCATCGACATGTTTGCAAATCTGCCATTGGGTACTCTGTTCATCATCGGCCTGTGCGTTTCCGCGATCGTCTTCATGGCAACCACCTTTGATGCGGCATCCTTCTCTCTTG
CAGCGATGACCACCAAAGAACTGCCAGATGGTGCAGATCCAGCACTTTGGAACCGTCTTGTTTGGGCACTATTCCTGGGCATTATCCCAGTGGTGATCAACCTGATTGGC
GGCGACTTGGGTACCATCCAGATCTCAACGGTTGTTGGCGGATTGCTGATGGTCATCATCGCTCGGCATTGTTGCAATCTCGTTCCTCAGAGAACTGCAACGCGACATTGCC
AACGGTGTCGAACTTCCAATGCGCAAAAAATAATAAAGCACATGAGGGCGCCAGCAAAAACTGGCGCCTTTTCTCTTAGGAGCACATCATGTCACTGGCACTTAACCCAG
CAGTCGCACCCATCAAGTCCGTCGAGTTTATTCCGGTCAATTACAAAGCCAGCAACTGGAGCCAGAACACGGTTGTCGTCAAGATTACAGACGAAAACGGTGTCTACGG
CCTCGGCGAAGCTGACGGCTCTCCGGACGCAATTCTCGCTTATTCCAATATTGAAACAGAACATAAGTGGCTGACGAACCTGACCGAAAGAGTCATTGGTCGTTTGCCGA
TTGAAATCAACGCCATCTGGAATGCCATGTATGACGCCACGCAGTGGCAGGGCATGCGCGGTCTGGGCATGTTCGCCCTCTCCGGCATCGACATGGCACTTTATGACCTT
GCGGGCAAACAGCTCGGCGTTCCGGCTTACCAACTGCTCGGCGGCACCAACAAAGACAAAGTCCATCCGTATCTGACGCTTTATCCGGCCATTCCGGTAGATGCGTCTTT
GGACGTCGCCATCAAGGGCTACGCCCCACTGCTTGAAAAAGCCAAGGCAGACAACATCCGTGCGGTGAAAGTGTGCGTGCCGATCAAAGCAGACTGGAGCATCAAAGA
AGTCGGCTATTATCTGCGCGAGCTGCGCGGCATCTTGGGCTATGACACCGATATGATGGTCGACTACCTCTACCGCTTTACAGATTGGTATGACGTTGCCCGCCTACTCAA
TTCCATTGAGGATCTGGAGCTGTACTTCGCAGAAGCAACTTTGCAGCACGATGATCTGAGCGGCCACGCCAAACTGGTAGAAAACACCCGCTCCCGCATTTGCGGTGCGG
AGATGAGCACCACCCGCTTTGAAGCTGAGGAATGGATCACCAAAGGCAAGGTGCATCTGCTGCAGTCTGACTATAACCGTTGCGGCGGCCTGACAGAGCTGCGCCGGAT
CACCGAGATGGCAACTGCTCACAACGTGCAGGTAATGCCGCACAACTGGAAGACCGGCATCACCTCAGCAGCGGCGATCCATTATCAGTTCGCCGTGGGCAACGCACCA
TACTTTGAGTATGTGCATCCGGACTTCTGTGACGGGGAGCTGCGCAAGTATCTGGTGACGCCGGAAGCCGAGTTGGTGGATGGGGGCTTTGAAAAGCCAACCGCACCGG
GCTACGGAATCGACCTGAACCAGCAATTTATGGCGAACCTGTGATGACACTCTTCACTCAGCACCCCTCTGCGGGCTGAAGATACACACAGGAACGTTCTCCGCCCCGCC
TCCATCTCGCTTTACAGCCTCGCGTATCTGCAAAGTGTGCCTAGGGGCGGAGGATCTTCTCCTGAGCCACTCAGGAATATATTTCCCCTGGGTAGAGACGTACCCTGCCCA
CTACTCACCCTGGTCAAGCCCCATTGACCGGGGTGTTTTATGAGCTCACTTCAGCATACACACTTCCCTATACGTCAAGTTGCCGGATGAGCTCCGTTAATTTGATAAAAC
GTTCCATCAAAATCACGCATTTTCACGTATATACACGCACCTGACCAGCAGAATCACAGTTTCGTGAATTGATGGAGCTCTCGAATCAACAAATATAATTAAGTATTAAA
CATCTCTTTATTAGAATAGATCAGAAGAAAAACATACTACAATTATTTGATCAAATGGTCAGTTTTTAGAATTATTTATCAGGCGGCTAACGTATTTTATCCAAATATATT
AACATTTTGAATTGATATCATTAAAAGATTCAGATGTTTTATACCTAAAACAATATATTCTGCATTCTATTCACACATCGAGAGTTCGTCAATGATTCAACCAGTAGATGT
GTTGATACTACTTTCTTCTGTGAATCCAAAGATCATTTATATTGCCTTTCAGTGCATCCCTCATACGTTTAGGTGGTTTACTACTACTATTACTCTAGGGGAGATAATATGT
TTAGAGATAATCACGTTGAAGAGGCAAGAGTTATTGAAGGAGGCGATGCGCCATATAGCACAACGACCCCATACTCTATTTCATCAGGGGACACTTTTGAGGGAACGCT
GAGTTTTTCTGGCGATAAAGACGCAATTAAACTCGAGGTACAGGCTGGAGTAACTTACACAATTTCAATGGAAGGCATCGGAAGTGGGGGCGTTTCCGATACATACCTG
TACCTTTACGATTCCAACACAAACGAAATCGCAACAGACGATGATGGCGGCAGCGGACGTAACTCCGAAATCACCTTCACCGCAACATCCACCGGCACCTATTACATTG
GCGCTGCGTCCTACAATGATCGCTACACTGGCGATTACAGAGTCTCCGTTGATGCAGACGGCGGCCCAACTGACCCGCCAACCGATCCAAACGCGTTCACCAACGACCA
GATTGCAAGCCAGCTGACCGATGGGTACTGGAACAACTCAGGCCGTGGTCGTCGTGCGTTTGACGTGAGCCCGGGTGAGTCTCTCACCGTCAACCTCAATGGTCTGAACG
CTTCTGGCCGTGAGCTTGCAACCAAAGCTCTGGAAGCCTGGACAAAAGTCACCGGTATCGACTTCCAGACTGTCAACTCCGGCGCTCAGATCACATTTGACGACAACGAC
AGCGGCGCTTACGCAAACTCGACAACACGCGGCAGCACCATCGTTTCTTCAACCGTGAATGTGAGCACTGACTGGGTTGCGCGTTATGGCACTGGCCTGGACGGCTATGC
TTTCCAGACTTACGTACACGAAATCGGTCACGCGCTCGGTCTGGGTCACGCTGGCAACTACAACGGCAGCGCGACTTACGGCGTGGACAACCACTACTCCAACGATTCCT
GGCAGGGCACAATCATGTCCTACTTCGATCAGGTGGAGAACACCAGCGTCGACGCTAGCCGTGCCTACGTCGTCACGCCAATGATTGCAGACATCCTCGCTATGCAGGA
CCTGTACGGCACAGCAACCGATCTGCGTACCGGGTCCACCGTTTACGGTGAAAACTCCACAGCAGGCGGTTACTACGACGACATCGCAGACCTGAACCCAGCCACGTTC
ACCATCATGGACAATGGCGGCACCGACAGCATCAACTTCAGCTCTGTCCGCGCTGACCAGAAAATCTCTTTGGTGGCAGAGACTTACTCTGACGTGAACGGCCTGACCGG
CAACATGACCATCATGCGTGATACCGTGATCGAGAATGCTTATTCTGGCTCAGGTAATGATGAGCTGATTGGTAACGATGCCGACAACGGTCTGTATGCCAATGCTGGTA
ATGATACCCTCGAGGGTGGTCTCGGTGATGACACCCTTTCCGGTGGTTCTGGTCGTGATACCTTCGTGTTCAACATCAACTGCGGTGCGGACACGATCGTTGACTTCTCCA
ATGGCCTTGACTTTATGAGATTTGAAAGCGGTGCAGACAGCTTTGCAGACCTTACAATTCAGGACACTGGTGCAGACACAGTCGTCAGCTATGACGGTGGCTCCATCACC

CTCTCAGGTGTTGATTCCAGCCTCATCAATGCAGACGATTTCACGTTTGTATAAGGGGTAATAAAATGTGCGACCTGCCGCAATGGCAGGTTGCATAATTATTTAATCAA
ATCATAAGTTTGCGGTATTTCCCTCAAATTCAATAGATGAGCAAAATTAAGTGTCTTACCCCTTTCAAATAGTGTGCGATTTACATACTATGGGGACCACAATTTTTGTTG
GGGCATAACCACGAGTTTTGTAGCGCAAGAGGCGTCCAATCGTTCTTTGGGAGGACGTGCCGCGCACTGAAGTTTGGGGCCAGGTGGTTTCGGGAGTATCTATGAGTAC
ATCGAAGAGTGGGCAGCCAAAGCGCCCTATGGTGGCTGAGGCGTTTGCGGCGTCTCGGTCCGGGTATGCGGGTGTTGGGGTCTTGAGTCTGGTCATCAACCTGTTGATGC
TGACGGGTCCGATGTTCATGCTGCAGATCTATGATCGCGTGCTCGCCAGCCACAGCGTGCCAACGCTGGTGATCATCGCCAGCTTCGCCGGTATCCTCTACCTGTTCTTCG
CCCTGCTGGAAGGCATCCGTGCCCGTGGCCTGTCGCGTATTTCTCAGGAAATTGATGCCCGTCTGTCGCGCACCGCGTTTCTTGCCAACCTGCGCGTGCCACTGCGCATGG
GCAACAGCGGCAAAAACGTAGACCCTATCCGCGATCTGGAGACCATGCGCCAGTTCTGGGGTGGTCAGGGGCCAGCCGCGCTCTTTGATATTCCATGGATGCCGATCTAT
CTGGGTCTGGTGTTCGTGCTGCACTTCTGGCTTGGCATGCTGGCACTGGCCGGTGCGGTCATCATCCTGTCTCTTGTGCTTGCCAACGAGTTCATGAGCCGCGATCCCTCC
AAAGAGCTGGCCTCGCAAAACGCACAGCGCAGCAACACCCGCTCCGGCGCCCGCCGCAATGCGGAAGTGGTGCGCGCCCTTGGCATGAGCTCCACGCTCACCAACAAGT
GGGATGAAACCAACGCCCGCTTTTATAATGCCCAGACCAAGGCAGGCGACCGTGCCTCCCTGTTCGCCACCGCCATCAAAGCCTTCCGCTTCCTGCTCGGTTCTGCCGTTT
TGGCTGTTGGTGCATGGCTTGCGATCCTTCAGGAAGTCTCTCCTGGTGTTATGATCGCAGCGTCCATCATCACCTCCCGTGCTCTGTCTCCTGTGGAGCAGGCGGTTGGCC
AGTGGCGCGTGTTCATCGCAGCCCGTCAGGCCCGCACTCGTTTGGAAAAATCACTGGAAGGTCTGGATGAGACCCCGCCAGAAACCGAACTGCCGCTGCCAAGCAAGAC
CCTCAATGTACAGCAGATCGCAACAGCGCCTCCGGGCGCTCGTGCAGTCACACTGGCTGGCGTCAACTTTGAGCTTAAAGCAGGCGACGGCATGGGCATTATCGGCCCG
TCTGGATCAGGCAAAACCAGTTTGGTCCGTGCGCTTGTGGGCGTCTGGCCTATCCTGCGCGGTGCCGTTCGTCTCGATGGCGCGGAGCTGAGCCAGTGGTCGGAAGATCG
TCTGGGCACAGCCGTTGGCTACCTGCCGCAGGATGTGGAACTGTTTGACGGCACCATTGCCGAAAACATCTCCCGTTTTGCAGCAGAACGTGATGACGAGAAGATTTTGG
AAGCCGCCCAGCTTGCCAATGTGCATGAGCTGGTCACCAGTCTGGCTGACGGTTATGACACGCAAGTTGGTGAAGGCGGGGCTCTGCTTTCTGCCGGACAGCGCCAGCGT
GTGGGTCTGGCGCGTGCGCTTTATGACAGCCCGTTCCTGATTGTGCTTGATGAACCGAACTCTAATCTGGATTCAGAAGGCGAAGCAGCCCTGACCCGTGCCATCAAACT
GATGCGTGAGACCGGCAGCATCGTCATTGTCGTTGCCCACCGTCCAAGCGCGATCTCAGCCGTTGATCTGCTGCTTGTTGTGAAGGATGGCCAACAGGCCGCCTTTGGCC
CGAAAGACGAAGTCCTGCCAAAAATCAGCGCCCCTGCTGGCGTGCCAGCCCCCGGTGCAGTTCCAGCTCCCGGCGCAAAACCAGCTCCAGCTGCTGCAAAAACAGGACC
ATTGAGCGTGGTGCCACATGCCTGAGCAAACAACACTTACCACTATTCAGACCTCTCGTAAGAAACACATGCGTTGGGGCCTCGTCCTTGTCGGCGTTCTGGTTCTTGGCC
TTGGCGGCTGGTCCTCCGTTGCCAAGATCAACAGCGCGGTGATCACCTCTGGCGCCATCGCAGTCGATGGACAGGCCAAGAAGGTGCAGCATCAGGAAGGCGGCATTGT
CGGCGAAATCCTCGTCAAAGACGGCGATGAAGTCAAAGCCGGTGATGTGTTGTTCCGTCTGGATGAGACCGTGGTGAAAGCAAACCTCGCCATCACCCGCAAGCGCCTT
CACCAGATGCAGGCTCAGGAAGCCCGCCTTTCTGCCGAGTGGCGCAACGACGATCAGGTCGCCTTCCCGGACTATCTGGTCGAGCTGGCAAAAACAGACAAGATTGCCG
CCACCGCGCTGGATGGTGAGCAGGCTCTGTTCGCTGCCCGTCAGCAGGGCATCAAGGGCCGCAAGAAGCAGCTTGGCGAGCAGATTGCTCAGCTGGAGCAGCAGATTGA
AGGCCTGCAAGTGCAGCGCGATGCCAAGGCCGAAAGCATTGATCTGGTCACCCAGCAGCTCGCCGACTTCGCCTCTCTATTGGAAAAACGCCTGATCAACGCCTCTCAG
GTCACCGCCATCAAGCGGGAGCGGGCCGAGCTGGTTGGTGATCGCGGCGGGTTGATCTCGCAGGTCGCCCAGACCAAGGAAGCCATCAGCGAAAAGCGTGTGCAGATC
CTGCAGCTGGACGAAGAGTTTTTGGAGAAGGTCCTATCCGACCTGCAAAACATCCGCTCGCAGATTGCCGAGCTGGAAGAGCAGGAAGTGGCTGCACAGGACCGCCTGA
AGCGCATCGACATTCGCGCCCCGCAATCCGGTTATGTGCACCAGCTCAACGTTCATACCGTTGGGCGGCGTGATCGCACCGGGTGATACCCTCATGCTGGTGGTGCCGGAA
GACAGTGATCTGGTGATTGAAAGCCGCGTTCAGCCCGTTGATGTGGATCAGCTGCAAGTGGGACAAGACGCGTATGTGCGCCTTGCCGCCTTCGACCAGCGCACCACACC
AGAGCTGGGTGCAGCTGTGCTCAACGTTTCCCCGGACCTGACCCGTGATGAGGTCACCCGGCGAGCCATTCTATCTGGCACGCCTGAAGATCCTGGAGAGCGAGATGCCG
AAGCTGAACGGCCAGACCCTCGTCCCGGGCATGCCCGTCGAAGGCTTCATCCAAACTGGCGAACGCACCGTCCTCTCCTACTTCATAAAACCAATGCAAGACCAAATCG
CTCATGCATTGAAAGAGAAGTAAGGAGCCCCTTACTTCCTCGGAACGAAAGCAAAAACTGTGACGGAAAGGGATCTAAAGCCCCTTTCCGTACACTGCTATTCAACAGA
GCTCTGTTAGCCAGAACAAATCACGGACATCTCATGATGGACCATGCAGAAGTCTATGCACGGAAATGGCGGCTTAAAGAATTTTCCCATGAGGGTGAAACCACACGTA
GCAATTTTTACTCGTGTCGTTCCGAAGAGTTTGGCCCAGCTCTACTTAAGATTCTTAATTCAAATGGTTTAATCACCGAGAGAGGTGGCTTCGCACTATTAGCGAGATATA
ATTCCAATTCTGCCGTAAGAATATATGAATACGATGAAACCGCAGCATTGATGCAATGGCTCGATGGCCCACGCTTGCTGGATTTTATAGATCAGGGGCATGAAGAAGAT
GCACTTAATATCCAGTTAGAAATCACCCGTCAGCTTTTAAAATCGCAAGTCAGTTTGACAGGTCTTAAGACACTCGAACACCTCATGGCAGACAGCCTTAATATGAAACC
AAAGGCCGTGCCAGTTTGGGCCCAGGACGTCGTTCCAATCGCCGAAGCGATAGCACACGACGTCATGGCAGATCGTTCTGACTGGCAGGCTCTTCATGGTGATCTTCACC
CCCGAAACATCATCAAACACGGAACAACGTGGCGCGCTATCGATGCACGGGGCCTATTAGGGCCTCCAGCACTTGAGTATGCCAATCTTTTTATTAATCCATGGGATAGG
AAAGAGCTGATCTTTAGACATGGACGAATGGAAGATCTAGCTAAGCGGGTCAGTATGCAGTTGGGTTGCGACATTAAGGAAGTTGTCTCATGTGCGATCACCAATGCGC
TTTACTATGCACAATCAAGCTTCCTCATTGGACAGGGGCGTCATCCAATCAAATGCATTCGCCAACTGCTCACGCTTCTTTAAGCAACGTGCTCATGATACCATAAGCCCT
TACACCTAGCAGAACAGATTACAAGGTAAACGAGTTGCAGGGCATAACAGCTTTCATCCTCAAAACCCAACCACCAAACATTACAACCTTCACCTCCCTAATTATAAAA
ACGATTTTCATTTAACCAGTAAAAAAGCGGCCTTGAGGGCCGCTTTTTCTGTTCTTTAAAAGGCCCTCGCATATTCGCGAGGGCCTTCTTTTTTGCTTAAACGAGCATTTC
GCTTGGGCTGCTAAGGTCCAGGCCGATGACAAGTGGGTCATGGTCGGATGCGCCGTACTGGTCGTCAGTGTAGAAACCAGCATCAGTGAAGCGGCTGTTGTAGCTGAGC
AGGTCTGGCTCATCAGCATTGATGTGCCACTCAGCAAGGCCGGTTACCTGATCAGCAAGGCTATCGGATGCCAGAGCATGGTCCAGAGAACCGCGCTGACCATCAAACA
CGAAGCTGTAAGCTTCATCGCCCAGGTACTGAGACAGAAGATCGGTCATATCGCCGCCCTCTACCAGTTCCTGTACGGAGTCGCCTTTACCGTAGGTGTTCAAATCACCG
ATCACGAGGATATCCGGATCAAGACCCGCACCCGCATATTCGGTTCTCAGCCACTCATGCAGCTCTTTTGCTGCGTCTTCACGTGCCTGAGCCCAGAAAGCCTGCCCGTC
GCCCTGATCGTAGTTTGGATCAGCGATCAAGGCATCCAGATCTGCCTGTACCTGGTCGACATCAGCTGCCGGAATGGTACCTGCATCCAGTTTGCCCTGAATGTCTTCCAC
CAGATCGCTCAGGCCACTTGGACCCTTGGACTTGTAGTGGTTGACTGCGAGGGTAAAGCGCTCGCCGTTTGCATCCTCAAAAGTGGTTGCAACGGTTGGGCGGTTACGCT
GGAAATCACCTACATATGAACGATCAGCAAAGTCCTGCAGACGATCCGCAATCGCGTAAGTCGCATCCGCACTTGTTTCCTGATAGGTGTAGGTTTCTGTCTCAATGACA
GTCAGCTTGCCAGGACGGTAGATCATACCGGTGGTGATAGCATCAGACCCAATCGGAGCCCCGTCAGTCGGGTTCACGAAGGTGTAAAGCTCATCAGTCCGGCCCTGAG
CGGTCAGCTCCGCATTCAGTGCATCAACAAGAGCGCTGATCGCAGATGCATCGTCGAAGCCGTTGTTCTCCAGTTCCTGAAGACCAAAGACGTCCGCATCGATATCAAGC
ATTGCCTTGACGATCTTGTCGGTCTGACGTGCAAGGTCTGCCTCGGTGCTCGCACCGCGTTGACCAAGGGTGGTGAAGTAGTTCAGCACGTTGAAGCTGGACACTTTCAG
ATCACCGCCAACATCAGCAGGCTGGTCAGGACGTGCACCGGAGTTGGTACTCTCGTCGATGGAGATCTGACCGTCAACGATCAGGGTATGACCATCAAAGCCTTCAGTG
AGAATGCCTTCGATCGGAGCATCGATCTCTGTGCCAAGACGCAGAGTAGCCCCTTCAGAGAAGTCATCACCACTGTCCAGATACCCATTGCCGTTGTCGCCGTTTTCTTCG
GTAACTGGCACATAGGTGAACTCGTCAGGGTTCTGTTCAGAAGAACCATCATCGATGAGGATGCGGTTGTTCTGGTTCATTTCCTGATGTTCGATGATCTGGTCAAGCTCA
TCCTGCGGATCGTACAGATGGGTTGCCTGATACTGGGCGCCAGCGGAAACTGTGATCTGACCGTAGCGGTCAAAATTGAAGTTTTCGATGACTGTCAGAGCATCGTCTGT
ACCGGAGCCTACGCTGACGCGCATGCCTTCCAAAGCTTCCAGGTTATGCTGATTTGTTGGAGACAGAAGCACCTGCATGTATTCTGGAAGTTCGTTTCCAGAGGACAGAA
CGCTCAGATCACCAATGTTTGTCAGCTGTGTCTGACCACCAAATTCAGCAACCGTTCCACCGAAGGACACTTCATCACCAACAGCAACGGTTGGAGCCGCGCCGGTGTAA
ACGAAGATACCTTCAGAAGTTGCGGAGTTGCCGTCACTGTCAGTTGTTTCTTCCTGCAGATAGAAACCACTGCTGGTCACGCTTGTTACAAGCGCGCTGACGTAAACAAA
CTCACCAACCATGCCGGATTCTGCACCGTCACCCTGAATGGTGGAGATGAGCGTAGGCGCGGTATCGTCGTTCACAACAGTGCTGGTTGCAGTCTTGGTTGTGATCTTGC
CACCCGTTGGGTTGCTCAAGGTCACAGACAGGGCTTCGCTCAGCTCGGTTGTTTCATCGCCAGTGATGTTGACTGTGATGGTCTTGGTGGTTTCGCCAGCTGCAAACTCGA
TGCTGCCAGATGGCAGAATGCCACCGAAGTCAGCTGCATCAACGTCGCCGCCAACAGCAAAGTCAACAGTGCCAGCCTCAGACAGATCACCGTCACGGGTCACAGTGAA
CTCGAAGGAGGTTACGCCGCTGTCGCCTTCAGCTTTGGAGTCAGCTTCTGCTGCGATAGCGTAGGTTGTGGACGTACCGCCACCGCCTGTGGTGTCGCCTTCCAGCTGCA

GGTTCTGCAGAACGAATGCTTCATTGCTGCCGTTTGCAGAGCCTTCAACGATGATGGTCAGCTCGGAGCCAGAACCTTCCAGATCAGCAACAAGTGTTTGCAGCTGGTTG
GTGAGAACAAGGCCGGAACTGGTTTCAACCATTGGATCGTTCAGATCAACAACCTTACCGCTGGCAAGGGTGTAAGACTGAGCGCCGTCATTGTCAGAGGTGAAGGCGA
ACAAAGTTTGCTCTTCGCCACCATCAATGGAATAGCGGATGGAAATAGTGTCATCAGACTCGAAGTCACCCATTGCACCAGCATCCAGAGACAGGCTGATATTGCTGTAA
TCTGAGATGTCGAAGCGCCAGGTCGCGACCAGCGGGTCACCATTAGGATTATCGTCGTTGACGGTATCGGTCAGACCGAAGAACTCACCAGTGTTGGAGACTTCATCAAT
AACGCCGACTGCATCATTCTCGAAGCTGCCATTGGTGGAATCGACCAGCGAGAATGGAGCATCTGCGGCAGTGTACTTGCCAAAGCCGTCTGCAGGACTGGTGAAGTCT
GGTGCAGTGCTGGTGTACTCGATCAGACGATCATTCTGTGAGTCCACCATGTCAAACGCGAAGGTTTCGCGCACAACAGCCGGGTTGCCGAGGATCTTCAGGTTCTGCAG
CGCAAACGCTTCAGAACTTCCGTCTGACTTGCCTTCAACGCTGATGGTCAGTTCAGAGCCAGTGCCTTCCAGATCAGACACGAGGGTCTGCAGCTGGTTGGTCAGCACAA
TGCCTGCGTCTGCATCAGTGAGCGGATCATCCAGATCAACAGTCTTGCCGCTGGCAAGGGTGTAGGTCTGTGCACCATCATTGTCAGAGGTGAAGGAGTACAGAACCTGC
TCTTCGCCACCATCAATGGAATAGCGGATCACAAAGGTGTCTGTGCCTTCAAAGTCACCCATTGCACCAAGATCAAGGGAAAGACCCAGATCTTTGTAGCCTGCAATATC
GAAGGTCCAGGACGCTGTCAGTTCGCCATCATTGTCATTGTTGATGCTGTCGGTTACGCCGAAGAATTCACCGGAGTTGGTGGCTTCATCGATGATACCCAGAACGTCAT
TCTCAAAGCCGCCATTTGTGGAATCGACCAGTGAGAACGGCGCGCTGTCTGCAGTGAACTTGCCGAAGCCATCAGCAACGCTGGAGAAATCAGGCGCAGTGTTGGTGAA
GCTCACCAGGCCCTGATTTTGAGAGTCGACCATGTCGAACGCGATTACGTCACGCACAACTTCGTTTTCGTCTTCAACAGTGACGGAAACACTTTGAGTGTCAGAACCGC
CGTCAGCGTCGGTGACAGTGACTTCAACCTGATAAACGTTGTCGCCATCAGCATCTGCTGGTGCTTCAAAGTCAGGAGCTGTCTTAAAGCTCAGCTGACCGGTTTCAGCG
TCGATTTCAAACAGCTCTGCATCAGTACCAGAAACAGCATAGGACAGGTCTGTGTTGTCTACGTCGGTTGCAGATGCAGTTGCCACCTCAGTGGTGTTTTCTGCAACGGA
AACGTCAGAAGCGACAGACAGAACCGGAGCCTGATTGGAACCAGTCACGGTCACAGTAACGGTGACGTCTTTGGTGCTGCCATTGCCGTCGGTGACTGTGTAAGTAAAG
GTATCTTCTGCAGTCTCGCCTTCACGTAGGCTATCAAACACACTGCCTGGGTTGTAGGAGACCTGACCGTTCTCAAGAGAGACGGTTGCGCCAGAAGCGCTGCTTGCGGA
AACTTCAGTGATTGTCAGCGTGTCGCCGTCATAATCCTTGTCGTTTTCCAGCAGAGTGGTAGCTTCAATCACAAGACCTTCATCGCCGGATGCAACAGCATCTTCACCAGC
GCGGATCTCGTTGTTGCGACCATCCAGATAGTAAGTGTAGTCATCAGCTGCGAAGTGAACTGCCTCGACATTCATCAAGGTGTCTTTTTCAGAGCCGTCATTTGCAATGAC
TTTGGTCTTTGGATTCCACCAGCCACCACGCTGCTGGAATGAATAGTCATCGATTGAGCCTTCATAAACGGCTGTATCGAAACCACGACCACCATTGATGCTATCATTGCC
TGCACCACCATGGATGACATCGTCGCCACCAGCTGCGCTGATGCGATCGTTGCCTTCACCAGCAAAGATAAGGTCATTGCCCTTACCTGCCCGGATCCAGTCATTTCCAG
CATATGTGGAGATGGTATCGTTACCGTTGAAACCCCAGACCCAATCAATACCGTCGGTACCCGTGATGACGTCATCCCACCACGTGCCAAAGTTTACGCTGATCGAGTAG
AAGTTAAATCGAGGAAAGAAATTGAACATGTATCAACCCCGTTACAGTTCATCTTATCCGAGTCCGTCACTCCCCTCCCAGTTTTGAGCCGGGAGCCAAGACGAACGATT
ATTTTTTGCTAGATGCCCGAAAGATTTTCTGAAGCATGCATCGACAGTTTTCTCGATGTAGAGGGTGACGTAAAGGACGTATTGATCAATTTTTCACTTTTCGCATATTATT
TCGAATAATATTTTCTTTTGTGAAATTTTTTTCAACTTCGCGTTGCATATCAATGTCAAGGCGACAGGAAGTTACCCAGTGCTGTGCATAACTCGGGATGCTTTTAAACAC
AGTTTGGCTCCTGCGCGACTCATATTTCACAAAGACGCACTGCGAATCCATAGATACAATTAAAATTTAATTATCAAAATCACGATTTTTGCGGGTTCATGGCGAGGGGC
AGGCGCCTGCTTGAGCAAAATCATACCAATTCTATATTTTATCTCATTTAGAGTATCATTTTCTAACTTGACTTATCTTTGTAGTATGTAGTTTCTTATATTCAATCTTTTCA
GCAAACTAAGCGAACTTAGAGTGTGTTCGCCATAACAAGAATCTGATGACACGGAATAAACGGACGGCGGCAAGCATGTGTAATTTGATCGATAAAATCAGAAAATCAG
TGATCGGTGAGGACGCAGCCATCGCCACGGCATTTGGTCCTAAGCCATTGGTTTACGCGGATTACACAGCCTCAGGCCGTTCGCTGAGCATGATTGAGGATTACATCCGC
GATCACGTTCTGCCTTACTATGCCAACACCCATACAGAGACGTCCTACACTGGCGCTCAAACCACCGCACTGCGCGAGCAGGCCCGTCAGCAGATCCGTTCGTCAGTTAA
CGCCACCGACGATCATTCCGTGATCTTTTGCGGTTCTGGTGCAACCTCCGCGATCCATAAACTGATCGACATTCTCAACCTGAAACTGCCAGCAGACCTGAGCGCACGCT
ACAAGTTTGAACAGCAGATCCCTGAGGAAGAGCGCCCCGTCGTTTTCATCGGCCCATATGAGCATCATTCCAACGAGCTGCCATGGCGCGAGAGCATTGCAGACGTGGT
TTCCATTCCCCTTTGCGAAGGCGGTCAGGTCTGTCTAGACGATCTTGAGAAACAACTGATCAAATATGCAGACCGTCCCCTCAAGATTGGCAGCTTCTCGGCGGCTTCAA
ATGTCACCGGCGTCAAAACCGACGTCGACACCGTCGCCCGCCTGCTTCACAAACACGGCGCACTCTCCTTTTGGGACTACGCCGCGGCGGCTCCATATGTGGGCATTGAT
ATGGCTGGTGCTCAAAATGAGCAAGGTGACAGCACCAAAGACGCTATCTTCCTGTCACCACACAAATTTGTCGGCGGACCGGGCACACCGGGCGTTCTGATCGTCAACA
ATGCCATTCTTCAAAACCGAGTGCCAGCGATGCCAGGCGGCGGCACCGTCATGTACGTGACGCCAGAGGATCACCGTTACGTGGCAAACGCAGAACGGCGCGAAGAAG
GCGGCACACCAGCGATCATTGAATCCGTGCGAGCAGGGCTGGTGTTCAAGCTTCAGCAAGCCGTCGGTACTGATGAGATTGAACGCCGGGAAGAGAGCTTTGTCTCCAG
AGCTGTAGCGCGTTGGTCAAACTGCGAGAACATCGAGATAGTCGGCAAGACTGAGGCACCACGCCTGTCCATCACCTCGCTGAAGATCACCCATAAAGGCAAAGACCTG
CACTACAGCTTCATTGTGGCGCTGCTCAATGACCTGTTCGGCATTCAGGCTCGTGGTGGTTGCTCCTGCGCAGGTCCATACGGTCACACCTTGCTTGGCATGGACATGCAC
TACAGTAAGGCACTGGAGGAGCAACTGGTGGAAGGGCACATGATCCTGCGTCCGGGTTGGATCCGCCTCAACTTCAACTACTTTATTGATGAAGACACCTTCGAATATCT
GGTGCAAGCTATTGAGTTGGTTGCGGAGCATGGCTGGAAGTTGCTGCCATTTTACCGCTTCGACCAAAACGGCGCTGTCTGGCGCTACCAAAACCACAAAACAGAACTC
GCAGCATCTCTCGAAACACTCGATTTCAACGCACTGCCAGAAGCCGCAGAAAAATCAGAGCGCCCTTCTCTGGCTGAGACAATGAAAGCAGCTGAGGCCGAACTCCTGC
GTAACCACAGCACTGAGGAACGCTACAACGTCATTCAGCCGCAAAAAGTAGAAGAACTCCGCTGGTTCGTCCGCCCACAAGACGTTGCTGTGGCTTAACAAGAAGCCAA
GGCGACAGAGAAAGACCGTCGCCTGTGATGCCTCACAAACACTTCACCCAATCCTCATAGCGTCTATCGACATCTGGAGAAGATATGGCTCGCGACCATGCTATACCCGC
AGAGTTTCGGGCTTCAGTCGGGGTCATAAAGAAGCGTTTTTTGAAAGCACGACTGAAGGTACTCTCATAGACAAAGCCGAAGGAAAATGCGATTTCTGAGTTTGAGCGG
TGGGTATTGCTGTTGCTGATGAGCGCCATCATTGCGCCCTTGAGACGTTGCTCCAGAATGAAGCTTTTCACACCGCCGTGATTGCGGAACATCTCATAAAGTTTTGACCGC
GAGATACCTGCTGTTATAGCCACAGTTTCTGGCGACAGATCCGGGTTCGACAGTTGCTCATTGATTGTGCGCTTAGCCAGCATAATCTGTGCGGCATGGCTGTTTGCATCC
ACCAGAACGACGCCTTCAGACGTGCTGTTGAGACAAGCAACGATGAGTGAAATCGTTGCCGGGTTAATATCTGCAGCTTGTGACAAGCTCAGTTGACCCGACAACTCTTT
CAACGATTTAATATGACTGTAAAGAAGCTGAACCATGGGCTGCGCGGCAGGTAATCTCCGCATATGCTGCCCATCTGGATAGACCAGTTTAGGTTCGAGCAACTCCCTGG
GAATAATCAGGGATAGGTTTGAAAATGTATTGGTCCAGGATTTTGATGTCCTGGCTAAATCAAAAATGATGAGGCCATCCTGTGCAACAGAATGTTGTGCACCTTCATGT
TCGTAAAACATGTCTCCTGTTTCGAACAGTTGGATCATGAAGTGATCCATTCCATCTTTTGCAATGGTCAGTTCACTCCTCTCCCAATACTGCTTTCTACTTTTTGTCCGGG
CCAGAATAATTTGCCCAAACATATGTGCATCTACTTCGGCGGTGAACGCATCCGACCTAACGGCATGGTCTGCTTCCACCTGAAACATGAAGGAGATACTTTCCTTCCAG
AGAGAGAAGCGCTCATCTGGCGCAACGCCGTCAACGGAAAACACAGCCTGAGGAATAGGCTCCGAAATTTCCACCATCGGCGCTCCATCAATGAGATCCACCTCACTCA
AAGAAACTGTCCTTGTAAAACCAGCACAAGTAATACGTGCTTTTACTAAAACCTTATAGCAATTGTATATATCTCCAAGCCCCACCCAAACAATGACGTGGCGTCCAGAT
CATTGTGCATTTTGTTTGGAACCTTCTGAAAACTGACTTTATTCCTGAGAAACTTGCATTGATGCATGAAGAGAGCACAACGACGAAAAGACCCCAACCACCGGTTTCTG
GCACTTCGCCTCTCATTCAGGTGAACGTAATGTGATTAAACTGGGAACGAGGCCGTCTTGCATTCTTCTTTCCAAAATCACACCTAAAGAAGTGCTCACTGCACATATCTT
AGTTTGTGCGCTTCCCCATCGAGCATCTTAAGGATCTCCCAATGGCTTTGGATAGTTCTGTTTCCTCAAACTCTTCATTGAGAAACGTATTGCCTGCTTTGCTGCTGGCTGG
TGTCATCGGCGAAGTCGCCTTTGAAGCCTATGCTTGGTTTGTCTCACCCATTCTGTTTGGCCCTGTGCTGGAACCCGCCTATCTCATCAGGGCCTTGGCAACCAAACTGTT
CGGCCTCACCATTCCCTATAGTGTTGCCTTCATCGTCCACTTTTCCGTTGGTGCAGTCGTCTTCCCACTTGCGGTGTTTGCCGTGCATGAGGTCACCAGGCTTCGATTTTGG
CTCTCAGGCGTCGTTGCTGGTGTGCTGCTCTGGTACATCGCGCAGGGCATGCTCGCGCCCTTCATCGGCAGAAGCTTCATGATGGATTTTGGCGCTTATACGCAGTCCTCA
TTCATCGGACATGTCGGTATGGCCATGGTGATGGGTGTCATCTTCCAAAAATTCTCTGAGAACCGGAGAGAGCGCTAAGCGCCCTCTCCTTCATCTTTGTGTCAGCTCACG
CAGAATACTTAAGCCGGATGACCAACTGCTGCGTATCCTCGTTGAAGTCCTCATGCATCTCCTCAAATCCGTAGCGCGCATAAAAGCCACGCCCGATTGGATTGTGCTTA
AAAACTTCCAGTTCAACGCTGTCATGAAGCGAGGTGGCCTTGTCCATCAAAGCCTTCCCAAGCTTCTGCCCGTGAAAATCAGGGTGCACAAACAGCCCGCCCACTTCATT
GCCAAGCAATGCAATGAACCCGACAACGGCCTCGTTTATCTCAACAACCCAGGTTTCAGCGTTGGGCAGATACAGGTTTCGAATATTGTCGGTTTCGTGGGCTAAAAAGC

TGTCGCTCAGAAAAGGATGAGCAAGTTTGCTGGCCTTAAGCCAGACGTCAACAATGGACTCGGTGTCACGAGTTTCAAATTTACGGATCATTTCTATGTCCTGAAGAGTG
CCGTTAATCGGCGCTTGATGTTCTGAAAATGAAGGCATGCAAAACCGGCCGTGCAAGCTGCGCGGTCTTAAAGTTGAAAGCATGCGAAGGGGTTCTTCTTTGAACCCTCA
GATCCGGGACATCAAATCTCCTGTGAAACATATAAGAGCTGGCAATCTATACCTCAGTTGCGCAATGTCAAGGCACCTCAAACAACTGCTTCAGGTGCACATATTTAGGA
GCAGTTTCGCAACTCGTGATTTCCGGAAAATATTAGCTGGCTTGGGAACTTGATGGTGAAGACGCAACCGTTGCCCGTTTCCCATTCCGGACGCGTCTTCACCATCAAGC
AGGCTCGATTCATCGTCAATCTCGAGCGAAAAGCACGCTCGGAGACAGTTTTAAACATATTATTGAGCTAAAAAGGCAAGTTTCTTCACTGGGTGAAAACCTTACCTTTA
ATGGACCCTAATCCTCTGAGTTCTCTAGTTCCTTGATTTTCGTTAGGTCCATTTGCTTTCAAGAGAACAACCCCAAGATGCCAATTGCGCAAAAACACTTGCGTTAGATCA
ACGTCCTTCATCTACGCAACAGTACGATTGGGTTGATCGTGGGAGAGGTCTGTAAAAATGTCTACTACTATAGTTAGCAGTGAAGTGGAGAAGCTCGCTGAAATTCCCTT
GCTTAAACGGTTGAGTACCGAGGCAAGAGAACAGCTGCTTCGGGGTGCCACTAAAAAATCTGTCTCGCGCGGTGAGATGATATTTTTTCAGGATGATAAGGCAGACGCG
TTCTTTATCGTTCTGGAAGGGTGGGTCAAAGTTTTCCGCGCAACAAGCAGCGGCGATGAAGCCGTTGTAGGGGTCTTCACCCGTGGGCAAACCTTTGCTGAAGGTGCTGC
CTTTACAAACTCAAACTATCCAGCTTCTGGTGAAGCGGTCACAGATAGTGTGGTGCTTTGTATTCCTGTGGCCCATCTGGCCAAATCAATTGCAGCAGATCCGAACATCG
CGCTGGCGATGATGTCATCTTTGTCCATGCACCTGCATCAACTCGTGTTGCAGATTGAAGGACTGAAGACGCATACGGGCGCGCAGCGTGTTGCTGAATTTCTATTGAAC
CTCAGCACAGAACCAGCTGGTTCCTGCACATTGGTGCTGCCTTACGATAAAGCGCTCATTGCCGGTCGCTTGGGCATGAAACCGGAAAGTCTGTCTCGTGCATTCCAGAA
GCTGAAAAGCTATGGCGTCGAGGTCAAGCAAAACCGGGCGATCATTTCCAACGTTCAGGTTCTGCGCACATTGGTAGAGCAGGAACGCACAGTGGTTATGCATTCAAGC
CACGGCTAAGACATTGCGATAATTGAATTTACTCGTTCGCCTGTTCCAAGGGAGTTATCCCTCATAGGCCTGTTTGATGAACGTGCCAAGATGTTGTAGCTCCTCTCCTCG
CGGCGAGCTTTTTCGCCACACAAGCGCTATCTCTCTTGGCAATGCCTTATCTACATCAATGAGGCATACATCATAATTATCTGTCAGTCCGGCAGTGATTGCGACGTCAGG
CAGAAGCGTTGTTCCCAGCCCTTCTGCCACCATACCCACTAGAGTGGAGAGGCTCGTCGCTTCAAATGAGCTATCTTGAACCGGGCCATCAGGTAAAAAAGCAGACAGG
GCATGTTGTTGTAAACAATGCCCTTTTTCCAGCAACATCATTTTGGTATCGGTCAGGTCTTGTCCATCTACATGTTGCCGGTTCGCAAGCGGGTGCTGAGGTGTGACGGCC
AGTTTGTATCCATCTTCAAACAATGGAAAATGAACAAGGTCGCCTATCTCAAATGGAAGCGCTAACAAAACCGCATCCAGATATCCGCTCTGCAATGCCGCCAACAGGT
CATGGGTAATGCCTTCTTTGATAAAAAGGCGAAGGTCCGGAAATTTCTTCCGAAGCTGCGGCATAATCCCGGGAAGCAGATAAGGACCAACTGTAGAGATCACACCCAG
TTTCATGTCTCCAGACAGGAACTTTGAGTTCGCCGCAGCGAAATCTACAAAGGCTTTTGTATCAAGCAGAATGTGACGCGCACGATCAGCCAGAGAGTTGCCCAGTGGTG
TAATAGTGACAGAACGCTTGGTTCTTTCTGCCAGATGAACACCAAGATTCTGTTCCAGCTCTTTAAGCGCTGTGCTTAGAGCGGACTGGCTCACATTACATTTTTCTGCTG
CCCTGCCAAAATGGTGTTCCTCAACAAGGGCGACCAGAAATTCCAACTGGCGAAGGCTTGGTAGATGACTCATATGATTTTTAGATCACTGTATTATTATTAATCAATTTC
ACTAATTAACATAGATCGATATTCTGCATTTATCAAATCAATTGAGCCCGAGGAAAATAAATGTCAGAAGAAACTATCATCACATTGCCACAGCTGAACACACCAGCAC
CAGATTTCACGGCCCCAACAACAGATGGCACGAAATCACTGAGTGATTACAAAGGCCAATGGCTGGTTCTCTTTTCGCATCCAGCGGATTTTACGCCGGTCTGTACAACT
GAATTCATTGCTTTTGCGAATGCAGCAGATGATTTTGGCAAGCTCAATACAGCATTGCTTGGCCTGTCTATCGACAGTCATTACTCACACATCGCGTGGATGCGAAACATC
GAGGAAAAGTTCGGCGTTCACATCCCCTTCCCGATCATTGCAGACCTTTCAATGGAAGTCGCCAAAGCCTACGGTATGATCCAGCCTGGAGCCTCAGACACCTCTGCTGT
TCGCGCAACTTTCATCATTGACCCTGAGGGAATTCTGCGGGCGATGGTTTACTACCCAATGACCAACGGGCGTTCAGTGGATGAGTTCTTACGCTTGATCAAAGCCTTGC
AAACATCTACCGAACACGGTGTCGCAACTCCAGAAAACTGGCAGCCTGGTGACAAAGTAATCGTTCCTCCACCAGCAACAGCGGATCTGGCTGAAACACGTAAATCAGC
AGGTTATGATTATGTAGACTGGTACTTCTCAACGAAAGAAGTGCCTGCTTAATACAAAACCAGAGACGGAGTTCAGGCGATCAATTGCCTGAACTTCTCACACAAAAATC
AGATCTACCTCAGAGCAGCCCTTGTCGGCTGTGCGGGTAAATCCCGTAGTATTGCAACTCTTTGGTTGGCAGAACCACGAGTTAAGCCATTTTCTTGCAAGTAGCTCTCCA
GATGCTGCGCCAATGCGTGCTCACAGCCAATCACCGGAGCAAGTTTGCATCCGGTTGGTGCGCTCTCCATAATTTTGGAGATATCTTCATGCCCGTGCAAGCCACCGCTG
GCGAACAGCGGAACAACAATTGCGTCATCGGAAAGTGTCAGGACCGCCTTTTCGGACGAAAGGCTCTTCCGAAAGGAATGCCACGCGTAGTTGCTTGTAGCCGTAGCGGT
TTTCCAGATGGTCTTTCAGTTGAAGGGCACAAAGCCGCGATTCCGGGCCAACTGTGGAGCCATGAGCAACGATCAGCACTTCCCGGTCGAGTTCCGGCTTTAACAACGTG
CTTTTGCTGTTTGGTGCCAGTGTACTTTGAATACTCTGAGACACCAGATCAGGTAGGAAGGGCAATACACCAAGCGGGGGTAACAGCTTGATATACTGCCCGTCTCGGTT
CAGATTTTGCCATGCACCTTCTTTCTGTTGTTGCAGGTTTGGGTCCAGCCCAGTCTGCGAGGCAGGACAGTGTTGCAGAAGTAGCCGTTACTTAGGAGGAGAGGAAGCA
CAAGCACATGCTCGTGCGGGAGGCTTTCGACTACTTCCTGTAGGGGCGGGGTCCCTTTCATCACCCCAGTCACCACATCCAGTTCAGGCAGTAAGGGCGCCAGAGCATTT
TTCAGGCGAACATGGCCTGGTGTGTTGTCGGCGCCCTTGCCACCCAGATCTCCATGAGCAGCAAGAACAAGTGCCAGATCGCTAGTGGCTCGTACCAGCGGAGAAGGTG
ATCTTGTTCCAGACATTGTACTTTCCAGATGGCTTACTCATTGGAATGCGCTTCACTTCTTCAAGCGTTTTCGCATCGTAAATGATGATCATGCCGTCATTTTCCCAGATAG
AGAGAATGGCGTGGCTACCATCGCGGTCAAACTCGATATGAGCAGCGGTTTTACCCGGCAGCGGTTTCAGCGTCTTTACGATCTCAAGCGTTTCTTTATCGATCACGTAC
ATCTGGTCCTTGTTTGGACCAAAGAACACGTCCGCCCAAGCGTACTTGGAGTTTTCATGAGATCGCATGAAGAAGCCCGGACCCGCCGTATCAAGTGTCTTCAAGACGGA
CCAGTTGGACATATCGATCACAGAAATCTTGCTCTCGCTCAGGTGCGGCGTCACCATCACTGGTTTACCCTGCCACTCCCAACTGATGCCGGAGCCCAGATGCGGGAAGC
CTGGCAGATCGATTTTGGCGATGGTCGCGCCTTCATCCATGTCCACCACAACGCCGCTCTTACCGTCACGAGAAGAGCCCATCAAATGCTTGTAGGACTGGTCGAAGAAG
AAATCGTCCATCGGCGCATCAACCTGAATACGGCGCAGGGCGAACAAACCTTCGGATGTTGGTTTCGTACCATCCTTGCCGTAGTTTTCAGCATCAACATAGGCATCAAA
CTCAGGCTCTGCCTTAGGATTGGTTGAGACCTCCCAAACCTCAGGTGTATCCTTCATGGCGACAACAAAGCTGTCACGTTCAGGAGCCTGATAAACTGCAGAAACACGTG
TGGTTTCGCCGCCCAAAGATTCCGCTTCAAAAATCTTTTCAACGGAAAGATCTTCAGCGGACAGCATCACCAGAGAGTGCGGCAGGTAGTTTGCCAATGCAATGTGCTTG
CCATCGCGTGAAAGCGCGATATTGCGGGAGTTGATGCCGCCGCGCACTTCACCAACCACTTTCAGCGCCCACATGTCGTATTTGGTGATCCAACCGTCACGGGACATGAA
GAACACGTAGTGTCCGTCCGGTGTGAACTTTGGACCGCCATGCAGCGCAAACTGGGTTTTGAAACGTGTGAGCGATTCAAACGTGTCACCGTCCATCACAGTGGCGTAGT
GATCACCGGTTTCCACGACCACAAACAGGTTCAGAGGATCGGATTCATAGACTGGCTTTTCAACGGGCTTGTATTCTTCGTTGAACACACGGGTCGCGTTGATTTCCTTTT
CGCCCCAAACCGGCATCGCACCCAATGGGGTCTTTACCAGGTTCACCAGATCAGCAACCTGCTGATCGTTGAGCTGATCCCCGAAGGCTGGCATCTGCGTGTGGACACGA
CCATCCTTGATGGTCGACGTCGCAGCACTGACCCGCAGACGCTTCAGGCTTTCCGGAATAAGCGCAGGGCCGATACCGCCTAAGCGATCACCGCCATGACACTCTGCACA
GAGCTCTTGATACGTGTCTTTTGTGGCGTCAGAAACGGGGTAGAGTTCCAGCTTTTTACCGCGAGCTTCCGCGGCCCCTGCCAACATGCAGGTTGCCAGCAAAACTCCGG
CGAGTTTAGGCATAAACGTGTTTGGCTTCATGACGAGCTCCCCGGAATGGTGTTGTTTCAACGCGCTGGTGGTTTGCATCCACGAGGCCGATCTCTTCATCCGTCAGGTAA
CAGGATGGATCTTCCGCCCAAGGGTCGCCGGTCAGCTGCAGCGCACGCACGCGCGTGTTGCCATTACAAATGTTGAAGTACTGGCAAGATCCGCAACGGCCTTTGATCTG
GCGCGGGGTCTGCTTCAAACCTGCCATGACCGGGTCGGATACATCCTGCCAGATCTCGGAGAACTTGCGCTCTTTGACGTTGCCGATGGTGTAGTGCCACCAGAAACTGT
CCGGATGCACCTTACCGGTGTTATCGATGTTGCCGATGTTCACACCAGAGGCATTGCCACCCCAGCGCTCCAGCATATCTTCCAGGTAACCAGCTTTTTCCGGCATGTTCT
CACGCACCCACTGCAACAGGTAGATGCCATCCGCATCGTTGTTGCCGGTCACGTATTCGTTGGACTTGCCAGCCAGCACATCTTCCCACGCCTTTTCAATCAGGTGGGTCA
TGGCTTTGCGGGTCACGTTCAGGTGCGCATCTTCACCACGGTTCTTGTTGCCGCGTCCTGCATACACAAGGTGAGACAGGTAGATCTTGTCCACGTTTTCACGGGCGCACA
GATCCAGATAGTCATCCAGCGTATCGGCATTGTCTTCTGTCAGCGTAAAGCGCAGGCCAACCTTCACGTTTTCCTCGCGCAGGTAACGCACAGCTTGCAAAGACAGGTCG
AAGGCACCATCCATGCCACGGAAGTCATCATGCACTTTGCCGATGCCGTCCAGCGAGATGCCGACATAGTCGTAACCAATCTCACCAATCTTACGCGCCATATCTTTGTC
GATAAGCGTGCCGTTGGTGGACAAGGCGGTATAGAAGCCAAGCTGCTTGGCGTGGGCCGACAGCTCCAGCATATCCGGGCGCATCAGCGGCTCACCACCAGAGAAGAT
GATGACCGGGACGCGGAAGCCTTTCAGGTCATCAATAACCTCATAGGCCTGCTTGGTTGAAAGCTCGCCGGGAAAATCTACATCGCCTGAGGTTGTATAACAGTGCTTGC
ACTTCAAATTACAACGGCGGATCAAGTTCCAGATAACAACCGGACCAGCAGGTTTACGCGCTGGTGGAAGTGGATACGGATCCGCTAGAGCTTTGAGATAATGGGTCAG
ACGAAACACGCGGGCTCTCCCCTCTGCGACGAAGTCTAAGTCCTGTTTTTTTCAGGGTTTTTGTCGAAAACAGGACATCATTTCCATTGCATGATGTCCCTATGAGCGCCT

CTATTTCCTTGACTAAAGTCAGACACTCTTCGCGAGTCTTTCCATGCACCATGGCAAAGACATTGTACGGCCATGTGGGTAGCCATCTTGGGCGCTGATAGCAGTGGGAC
ACAAAGTCCAGAGAACCCATCATTTGTCCAACGCGCTCAACGTCCTCATCAACGATATTCCAGACGGACATGCCGTTGGCACGCATGCCAAGGGCGCGATGATTTGGCA
CAGCCGCAATACGACGAATTGCACCGTTTTCTTTCATGGAAGCGACTCGCTCCAGCACCAGTTCTTCGCTCACGCCAAGCTGATGGGCAATCTGCGCGTAGGGCTCGGAG
ACAAGCGGCAAGCCTTCTTGCAGCGCAGCCACCAAAGCACGGTCAAATTCTGAGAATTCGTCGATATTATGCGTGCTCATACTGGAACCTTAAACTCAATGAAAAACTCG
CGTTCTTTCGGGAACAGCCGGACTGGAATCCCAGTCAGCTCTTCAATGCGTTTCCCGCAACTCTCAATTTCTTCAGGGGTTTCCGTGGCAAGCACGAACCACATGTTCAAC
TCGTGATCACGCGCATAGTTGTGTGCCACCTCTTCCAACGAGTTCACCTGTTCGGTCACCTCGTCATAGCGCTCTTCCGGCACGGCCATGGCACACAGGCAGAACGCCCC
GCCCATTTGAGCAATGTCGAACATAGGCCCGAAACGAGTGAGGATACGCGTTTTAAGCAGCCATTGCAGGCGTTCAAGCAGCAAGTCTTCCGCAATACCCAGTTCCTCTG
CAACCGCTGCATACGGCTTGGAGACAACCGGTAGACCATCTTGCAAGCGGTTGATCAGCTTGCGGTCCAGATCATCCAGATGCACCTTGGCAGGTTCATTGAGGTGGGCA
ATATCTGTGCTCATGCACTTACTCCAACACTGCGAGACAGACGGGCACCACGTTGCTTGAAGCACTTGGTGGAGAACAGCACATCGCTGCTCATCACATGCTGCGTGGCC
ATGGTAGAGAGCCGGTTGATGACCTCAAGCGCCTCTTCACGACTGTTCGCATGGGTCATGGTGTACAAGTTGTATGGCCAAAGCGGCATCTCTCTGTTGCGCTGGTAGCA
AAGGCTCACCATCGGGTCAGCAGAGAATGTGGCGCCGACGTCATCAATATGCTCATCACGAACATTCCAGACCGTCATAGCATTTGCCTTGTAACCGATGGACCGATGCC
GCACGACCACACCGAACCGGCGAATGATGCCAGCATCCTGCAGCACCTTCAGCCGTTCAAGCACGGAGGTTTCCGAGCGACCCAGCTTCAGCGCAATGCGCAGGAACGG
GCGGTCCGAGAGCGGAATGCCATCCGACAGCTTTTCCAGCAGGTCGATATCACCCGGTTCCAGCACTGACATGTCGACAGGAGGAAGCACGGGGGTCTTCTTTTCCAGCC
GCTGCTCTATCAAAGAAAACCCCAGATCTAAATGGAAAGCCTTTTCCATCCGAAAAGCATGAACCACCAGATTGGTGTTCTGCTCAATACGGGCGAGCGTGGCCTCTAAC
GTCGCCCGATCCCGTTCAACAACCACGAACCAGATATTGAGATCATTCTCACGCAGGTAGTTATGATTGACCCCCGGCTCATCGTTGATGAGCTGGGCTGCTGTTTCCAC
AAAGTGTGGAGGTGCAGACACGGCGGCCAGACAACTTGCGCTCACTGTATTTGGAGCCAGTACTGCGCCAATGCGCGAAATAACGCCATTTGCGACCAGCCGTTCGTAA
CTGCTGATCGCGTCTAAAGCGTCCACCCCGGTTTCCCGGGCTATGAGATTGAATGGCTTAGGAACAAGCGGAAAATTATGCTGCCACTTGTCCACCAATACGAGATCAAT
ATCACTCAGCTGTGCGCTCATATCGCAAACTCCCCAATCGCTCGGGCGGCCAGCGCCGCCCTGGGAAAGCCGGGGGTTAAAGCCCCAGCTGGTGTGCACGTGAAGTAAA
GAAGATGCCGCTTGGTTTTTCCGCTGGCAGAACTTCTTTCAGGGAGAAATCCGTTGTGTCGTAGACATGAATTTCATTCTTGTCACGAACGGACATCCAGACTTCGTTGCC
GCGTGGAGTGAACTCCATATGCAACACGCCCGGTCCCGGTTTCATCGTTTTGATGATCTTGCGTTCCAAAATGTCGATCACCTGAATGGTGTCGTTGTGTGGAGGCGCAA
AGTTCACCCAGATCTGACGACCATCAGGACGCGCCATGGCAAACACCGGCTGGCCATGCACTTTCACACGATCAACTTCAGTGAAATCACGCTCATCCATGAACAGAAC
TTCGTGGCGACCAATTGCTGGCAGTGCGTACATGCCATTGGCTTTCGCCCAACCTTCCAGATGCGGCATCTTGTAGACCGGAAGAGGCTCTTCCCCCTTGCCGTAACCATC
CATGATGCGCTCAACCGTCAGCGGCTCTTTCCAAAGGTCAACATGGGCAAGACCGTCTTCACCAAACAAACCAACGATGTAGTGGCGGCCATCACCCGTGACCAGCGCG
TCATACGGGAAAACACCAATGTCTTCGAGTTTGGTAATCTCAGGCTTCAGGCTCTTGGAATAGTCCAAAATCCATGTTTGCTCACCATCCCAAAGGGAAAAGGCAAAGCG
ACGACCGGGCAGATCTACGAGTCCAATGGTCTTCGACTGATTGCCCACTTCGTCATAGATCGCAGGCACTTCGGAGACCAGCTCCAGAGTTTCCGCATCAAATACATTCA
CGCCGCCCGGTTTGTAGTTGGCGACTGCGATCAGCTTGCCATCATCAGAGATAGCGCCGCCGATGGAGTTGCCAGCCTGCATGACGCGGCCGACCAGCTTGCCGGTGAGC
ATATCCACTTTGGACAGGCCACCATCGCGGCCAAACACGTAGGCATAACGCTCATCTCGCCGACTAAACGACGGAGGCGTGTGAGAGGTTGCCCAAACCCTCTACCTGAG
CCAACACAGTACGGTCTGTGCTTTCGACAATCAGGGCTGTTCCTGTGGCTCGTTCGATAATCAGACCAAGATCGCCCGTGGCGCGCAGGTCGTCAGAACTGGCTGCAGAA
GCAAATAAAAGACCCGCACCTAAGGCTGCGGCAGCAAAAAACGTCTTCACTTGAATGCTCCTGTTTTAAGGCCTTCCGCAATGAAGGCCGCTTCAGTTTCGGTAAGGAA
GCCGCGCCATGGAGGCATCGCTGTCCCTGGCAGTCCGCCAAGGATCACTTCCTGAATGACATCAGAGTCCATATGCTCCAGATTTTCTGGAAGAATTGGCCCACCAAGAC
CGCCCTCAAGCGTCAGACCGTGGCAGGACCCGCAGTCCTGCTGCACGAAATCTGTCAGCTTTTCGCCATCGATAATCTTGAGAGCCTCTTCGTCGGCGAGAGCTCCGATT
GGCGCTCCAGTCAGCAACGGAAGCGCCAGCAACCCAGCCTTCAGAAAAGAACGCAGCGGGTTTTGCGCACTCCGCTCTGTCTTAGCCATTGATAACCTCCAGATGAGGT
CTATCGCCCCATTTTGCCTCACTGCCGGAGAAGGATGGCAGCTCCAGAGAGTCTTTCAGACCCACAACATCACCAATCATAAACAGTGTCGGCGCTGGTAGTTTAGCGGC
CTTGGCCTGCTCACCGATTGTGGTGAGTGTGCCAAACACTTTCTGCTCTTCGGATGTGGTGCCCTTCGAGATTGCTACAGCAGGCATATCCGCAGGCATGCCGTGTTCCAC
CAGCTTGGTAGAAATCTCGGCGATGGCTGCAAGGCCCATGTAAATCACAAGGGTGGTGTTAGGGTCCGCAAGGCTGGCCCAATCCAGATCGAGGGGTACATCTTTCTGG
CAATGCCCGGTTACGTATCTAAGCCCCGTTGCAAGGCCGCGATGCGTGAGTGGCAGGTTTAGGCTTGAGGCACAGCCTTGCGCAGCTGTGATGCCTGGCACGACTTCATA
GTCGACGCCGGCAATGACCAATTCAGCAGCTTCTTCACTGCCGCGGCCAAAAATAAACGGATCCCCACCTTTAAGGCGCACAACATGCTTACCTTCTTTGGCGAGATCAA
CGAGTAACTGGTTGATTTCTTCCTGAGGGATCGGATGGCTCTTTGGCTTCTTTCCGACAGGATAAAGAGCTGTCTTAGGCGGAACCAGATCCATAATTTCCTGAGAGACC
AATCGATCGTAAACAACGACTTCAGCGTTCTGAATGCGTTTCAGAGCTTTCACTGTCAGGAGATCCGGGTCACCCGGTCCTGCGCCAATCAGACTGACCTTACCAATGCC
TCTCATAACTCATTCCTCAAATCGAAACTGGCGAAAATAGAGGGGCAATCCGAGCCCCTCTGGTGATCTGGAGTGTTGCCGAACCGATGTCTTTTAAGGTGTTGGAGACA
TAGACCATGCCCGACGGAGCGATCAGGCGTTCGTCTTTGATGACACTTTCAAACCTGAGAGTGTGGTCATCAAATGCTGCTGGCTTCTTGCGAGCGCTTACAGCTACGGC
GCTGATTGCGCATTGCTGCGCGATACTGGCCAAAAAACGGAGAATTGATTTAAACGACACCATAAACTCCAATCGAGACGAACCTTACCGGCATCGTTATGTAAGCTTGC
CCTCTCGCTTCTCACATTGACCATCGTCAAGTTATTGTGAGAGCCTCGTGAGTGGTCATAATTCCTATATATTTCAATAGGTTGATGTGTGTAGAAATTTCACAGAATTGC
AACCGGATGTGGTGATATCCCAGCTAAATTGACATAGGTCACGCACAGGCAAAAAATACGTAGACTTCCCACAAAAAATAATCGTTCTAAAAAGGAGAAAGAGATTCCG
TTTCAGGTATTTACTTCCTGACAACGGCAAGTTGGCACCCATCCAGAGCTGCACGGAGTGCGCGTTTAGGGAATAAACGTAAAGTCAGATCGGAACCGACCTTTTTACCT
AACGGTACCTTAACCCCATTGCCAAGCAAAGATTCCACCACTCTCCCATCCTCAGTAGGTTGAAACCAAAATCAACAAGCACCCGAGTGCTGTCTACGTGAGGAACCGCT
ACATGAGTGAGGCGAAGCCCTTTTCCTTTCTGAGCCCTATCAAAAACCCGCCAGCTTACGTAATGGATCCTGAAAAGAAGCATCTGGAAGTCACTCTGGATGTTGAATAC
GACACCAACAATGTTTTCAGCTTTGAGAAGCCTCACAACGTCCCAACCTACACCCGGGTCTTCACAAGCGACCTCACCAAAGAAGCAGGCAACACGCTGACCGGTCCGA
CCATTGAAATAAGACCGGGTTCAAAACTGCAGATCCACACGCACAACAAGTTGCCCGAACAAGACCAAGCCGAGCCGGATGACATCAACTTCCCCTACGGTTTCAATAA
CTACAACATCCATACCCACGGCTTACATGTGACGCCGGAATCGCCCGGCGATAACGTCTGCATGACCATTCGTCCCGGCGAAGCCTATGACTACGAGTACGACATTCCGG
CGGACCACCCACCGGGCGTCTATTGGTATCACCCACACAAACACGGGTCCGTGGCCCTGCAGGTTGGCAGCGGCATGGCGGGCATGATTATTGTGCGCGGCCCGTTTGAT
GATCAGCTCAAAGAGATGGGCGTTAAGGAATACTGCATGGTATTCCAGAACATAGCCGTGAATGCTGAGGGTAAGGTGGAAGACCAGAAGCAGGTCGCCAAGTTCCCT
AACAGCGAATACAACATCACGGTCAACAGTCAGTACTGCCCGATCATCAAGGTAAAACCGGGAGAGCTGATCAATCTGCGCTTGTTGAACGCAACCACACGTTCATCGA
TCAACTTCGAGTGCCCCTTCTTCGGGCAGATGTTCATCTACGGGTATGATGGCAACCCAACTGACAACTACCGCAGACAGGGAGGCATCAGCCTTGCCCCGGCAAACCGC
GCGTCTATCCTGCTGAAAATCAGCGAAGCGGCCACTCCGGGCGCGTTTTATTACGTGCAGAATGACTCCGTCGAGTATCAGTTCATTCCGCACTACAAGTACCACAACAA
CCCACCGCTCTTTGCGATTGAAGTGGAACGCGGCGAGGAGATCGGTGCCGCGATCTACATGCCGAGTGAACCTGCTGACCGCACCACGCCACCACCATTCCCGGAAAGC
CTGCGCGGTGGACTGCTGGAGCCAATTGGCCCCGAAGAAGTCACCAACAAGCGCACGCTGACCTTTCAGGCGAACACGCTGGACAATCCCGAGAAATACTATCGAGGGG
TGAACGTCACCATTGACCACGGCCAGTTTGGCGGTTGCGAGGTGGTTAATCACTGGGTCCAGCTGGACGCGGTGGAAGAGTGGACACTGGAGAACGAAAGCGAGTTTCC
GCATCCCATCCATATCCATGTGAACCCCATGTATGTCTTTGAGATTCAGGGCGAAGACCCTCTGGATTACACCCTGGAACTCCCGTTCTGGGCCGACACCATTCAGGTTCC
ACCCCATGGGCATGTGAAATTTCGCATGCGCTTCAAAGACTTCGTCGGCAAGTATCCATTACATTGCCACCTGCTTTATCATGAGGATGGCGGTATGATGCAGCACGTGA
AAGTCGTTAGAGACCTGAAAGAAAAAGACGACTGGAACAACAACCAATAACGGCCAAAGACAGCACCAGTTTTACATGACCGAAGCACACGGCATCAGCCCTAACAAC
GAGCGGAAGCCAAAGCGAATGCGCAGCTTTTCTCTAAGCCATTTTCTGCTTACCAATCCCATTGGTTATGGAGTGATGCTCGTGCTTCAGCTGGCTGTGGTCGCGATGATC
ACGCGGCCAGATACCCTTCATGTATTTGAAGTTGCTTTCGTTCTCGCCATATTCTTTGCAGGCATGGTCAGCCTGCAAAATGCAGGTATGCCCAGCTACCTCTTCACGTAT

TTCTGGAGCGGGTTTGTTTATGGCGGACTCGGTATGATTTTGGGCCACCAGATTGACCTGTCCATTGACCAGAACCTGGCGATGCAACAGATGAACAGCCTGACACTAGC
TGCTGTTTATCTGAACTGGATGACCCTCATCATGGTGGGCGCTTGTATTGTAACCTGCTTTTTCGCGGATCACTGCAAAGGGCATCCGGCTTCAAAAAGCGCAAAAACCTT
CATACACGTCTTCGCTATTTCCATGATGATGTTTGGCGTTTACGTAGGCCATGCACTCAACCTGATACTGCACCGGAACACGTTTGACCGGTTCTATCTACAAGTCCTTTTC
ATGAGCATCTTCTTTACCATCGGCTATTACTGCGCTCTCAAGTGTCACAACTTGCTTCGCGCCAGAAAAAAGGCCTGAGCGGCCTTCTCTCCTTTTGCTTATCCAACCTTCT
GCGGCGCAGAGTACCGCTTCTCCAGCCGTGCCTGTATCATCTCAAAGACAAAACTGACCGCCCAGTAGATGAGTGCGGCACTGATCAGCATTTCAAAGTGTTTGAACTCT
GAACGCCCGAGCGTCCGCGCCAGATACATAAGCTCCATGATGCCCATGGCAGACACCAGTGAGCTGTCCTTCAGCATGGAGATGAAGTTGTTGCCCGTTGGCGGGATGA
TAATGCGGATCGCCTGAGGCAAAACAACTTTCCACATGACCTGCCAGGGCTGTAGCCCCAGAGCAGCAGCAGCTTCCGACTGCTGCCGTGGAACCGCCTGTATGCCTGCC
CGGAAAATCTCCGTCATGTAGGCCCCGTAACTGAGCGACAGAGCCAGTATGCCGGAGGGGATTGCATCCAGCGTAATACCAAGCTGCGGCAATCCAAGATAAATCAGAT
AAACCTGCAACAACAACGGCGTACCACGAAAAAACGAGACATAGAACGTTGAAGCCGCATAGGCTGGACCTGAGTTGGACAGACGGGCAAGCGCGCCAACAAATGCCA
GCAAACTGGCGCAAATGATGGACGCAAACGAGACCAAGAGCGTGTAGAATAGTCCCTTATAAAGCAGGATCGGTAACCGCTCGATGATGAACGGAAACGACAGGTCAA
ACGAGTAGAAGAACCCCAGAAACAGCGCCAGAAGCTGCGCCCAGTAGATCGGCACTTGCACCTTCATCGGCAGCTTGGAGACGAACACCACATCTGCTGTGATCACAGC
AGA

Genes used forTDA analysis from *Pseudovibrio* sp. P12

>KU760700.1 Pseudovibrio sp. P12 LysR family protein transcriptional regulator (tdaA) gene, complete cds

ATGGACATTCATCAACTCAAAACCTTTGTGACGGTTGCTCGTGAGGGCAGCATCACCAAAGCTTCTGAACTACTTTTTCGAAGCCAACCGGCTGTTAGTGCGCACATAAA
AAACATGGAAGAAACGCTGGGACTCACCCTGTTCCAGCGCACCCCAAGAGGTATGGCGGTGACGCCAGACGGGCAGCGTTTGCTGAACGAAGCCCATCAGCTGCTCGAC
AGGCACCAGAAGTTTCTGGAGGAAGCCTCCCGGCTGCGTGGCAGTCTTGCAGGCCAACTCTCACTCGGCGCAGGCCGCAACTCCAGCACCGGCGTGGTCAGCAACCTTC
TGGCAGCCATGTCCCAGGAGTTTCCCGAGGTGGAGGTGGAGCTCCAACATCGGGCATCGGTCAAAGTGCTGGAAGGCATCCGCAATGGAAGTCTGGATGCAGGATTCTA
CACGGAGACTGGCGAGGATGATTCCGATCTCTCCACGCTGGAAATTTCCAACTTCTGCGTTTACCTCGCTGCACCTGCCAATCTTGTGCCGTCGGCTTCTGAGGTAAACTG
GAAGGCGTTGGAAGACGCGCCTTGGATCTACCCCAGTTGCAACTCCTGTTGCGGGCGGCTTGCGCCGCGATCTGCTGGAAACACACGGCATCCGGCCAAAACGCGTGATC
AACGTGGATGACGAGGGGGTCACCCGCGCATTGATCGCCAGCGGCGTTGGCATCGGCCTTGTGCATGCCAATTCAACCAACCCGCGGGAGCAGGAAGGCGATATCGTGC
TGCTGACACAGGTGCAAAAATCAGCCCGCGTGCTGTTTGCGCATTTGTCCTGCCGGATCGACGATCCCGTAATCGCTGCGGTTCGTTCCATCCTGGCCACGGACAACTGC
ACCAGCATTCCGTTTCAGGGTGACCAAAACATAGAGGCGGCGCTGATCGCCTCCTGA

>KU760701.1 Pseudovibrio sp. P12 glutathione S-transferase domain protein (tdaB) gene, partial cds

ATGCAAAGAATTCTATACAGCCTCAGCGGTGCCTCCGGGCAGCACTATTCCCCGCATGTCTGGAAGATCATCATGTCCCTGAAGCACAAGGGACTCGCGTTTGACCTGAA
GCCCGTTTCCTTTGGCGACATTGCCAGTATTGAAGACGGTTCTTATACCAGCGTGCCCGTGCTCAACGATCAGGGCACCTTGCTGGGCGACAGCTTCGAGATCGCGCAAT
ACTTGGAAACAACCTACAAAGACACGCCGACCCTGTTCGATGGCGACGGTGGTCTGCAGCTCTCGCGTTTCGTGGAGAGCTATTGCAAGACAGTTCTGCACCCAACACTA
GCCGTCATTGCCGTGATGGATATGTACAACATCATGAGTGATGCGGACCGCGCTTACTTCCGCACTGCACGTGAAAAGCGCTTCGGAATGCCGCTGGAAGACGTACACG
CCAACCGCGATGCGGAGCTGGCAGCCTTGCCGGAGAAGCTCAACCCGGTGCGCGACCTGCTTGCAACGCAAGACTGGCTTGGCGGCGAGAAAGCCATGCATGCTGACTA
TGTTTTGTTCGGTACGCTGCAGTGGTGCAACGTCTGTTGCGCAGATCCATTTTTTGCAGCAACGGACCCTGTCCAGATGTGGTTCGAACGCATGCTTGATCTTCATGGTGG
TGTGGGCCGCAAAGCAGTCAACGAAGCCGTCGCAGCCT

>KU760702.1 Pseudovibrio sp. P12 prephenate dehydratase (tdaC) gene, complete cds

ATGGATGCACATCTTGAAACACCGGAATATTTGGCAAAAGGAACAGCCCCACTGTCCTTTGATTGGGTCCGACAAGTACACACACTTGGACCCACCGGCACGAACTGCG
AGAGAGCTGCCCTGAGATGGGCTGCACGCCATTGCAAGAATGCTGAGGTCTTTCTTCATGCCACGATGGAACTGGCAGCGGAAGCCGTTGCCAAATGCAGAAAATCCGT
TCTGCTAAGCGTGGTGGCCTACCCGCATTTGCATTCCATCATTTACGAGCACATCGGTGAGCTGAAGCTGATGGACGTGTTCATCATGAACACCGATAACATGGTTCTGG
CGTCATCTACTGGCGAAACACCCAAGCTCTGCGCCACACACCCGGCACCTGAGAAGCTGATCCCGGCCTCTATTGAGCGCAAGTTCGTCACCAGCAATGCGCAAGCTGCT
GCTGTCTGCGCCAAAGGCGAAGCGGATGGCTGCATCACCACACTGTGCGCCGCAAACGAGTTTTCTCTGACAATCCTAAGAAAGCACGGTCCTGTGCCCATGGGCTTCAC
TATTCACGGCCCGTTACAAGCCACCCCGGCCTCCCACCCATCGTTTACGAGTGAGTTTCAAAACCCCAACAGGAGGGAATTTCATGACACAGAAAGCTTATGA

>KU760703.1 Pseudovibrio sp. P12 thioesterase (tdaD) gene, complete cds

ATGACACAGAAAGCTTATGAAACACCAATCGAAATCCGCTACCGCGATACTGACTCCATGGGCCATGTGAGCAGCCCAATCTACTACGATTACATGCAGTCCGCGTATCT
GGAATACATGCATGACCTGCTCAAACTGCCAAAGTCTGAAAAGCTGCCGCACATCATGGTGAAGACCTCCTGCGATTACGTTTCACAGGCGCTCTATGGCGATCGCTTGG
TGGTGCTGAGCCGCGTAACCAAGTTTGGTGGAAAGAGCTTTGAGATGCAGCACGAAATGCGTCTGGATGATGCAGATGGCCGTGTCGTGGCCAATGCGTCCTCTACGCA
CGTGATGTTCGACTACGAAAAGCAGTCAACCTATCCGGTGCCGGATGACTTCAAAGCTGCAGTGGCTCAACTTCAAGAACCTGTCGCCATGCACCAGCCAGCATAA

>KU760704.1 Pseudovibrio sp. P12 acyl-CoA dehydrogenase (tdaE) gene, complete cds

ATGGATCTACAATGGAATGAACGTCAACAGGCTCTGCGGGCAAAGTATGCCCAAATCGGTAGTCAGATTGACCACTCCGCACTGCATGCTGGCAGACAGGCTTTTGATC
AGGACACTTGGGACCGGCTTTTAGAAGCCGGTCTCTGGAAGATCATTGTACCTGCCGAGTATGGCGGTGAAGGCACCGATTGGTGGGACTTTACGGCAGCACTGGAAGG
GCTAGCCTCCTCCGTGCGTACGCCCGGCATTCTGCTTTCCGTCATCGCGCAGGCGGGTATGGTGCATGCACTCAATCTGTTTGGCACGGAAGCACAGAAGCGGGATTATC
TTGGCCGCATTCTGAAAGGCAAATTGAGCGCAACAGCGATTGCCGACCCGGACACCGGCACGGATGTGCGCTCCACATCCTCTGTGCTGACACCGGGACCAAACGAGAC
TTTCGTGCTCAACGGCGCCAAGTACAACATCGCCCATGCACCCGTGGCGGACTTTACACTCGTCGTGTGCAAGCTGGATGGGCACGCCCGCGATGGCATCTCACTCATCC
TGCTGGATGCTGGCACCAAAGGCCTTGAGGTTGGTGAATACGACCAGAAGCTCGGCAATCTGGATCTGCCAACTGGGCCGATGAAGTTCAACAACATCCCGCTTCACTAT
GGTGATCTGCTTGGCGTTCCGGGCAAGGGCCTGCGCAATCTCATCACCTTCGTGTCTCTGGGCAGACTGTATTACGGCCTCGTTTCTGCGTGGCTGGTGGAACCAATGCTG
AAAGCAGCCATGAGTTATGCGCAGCAGCGCCAGACCTTCGAGGTGCCGGTGCTTGAGTTCCAGTACGTGCAGAAAAAGCTGACTGACATGCGCATCTCTGCAGAAGCTG
CAAAGTGGACCGCCTATGGCGCGCTTCAACAGTTGCTGGCTGGCGCACCGGAAGCAGGCATGACCTGCTCTATCGCCAAACTGGTTGGCGCAGAGGCGATCACTAGCGG
TGCGATTGATCTTCTCAAGCTTTACGGCACCAAAGGCTACCATCAGGGCGAAGTTTCCACCTTCCTCCGTGATGCACTTGCCTTCTGCGCTGTGGGTGGCACCGAAGAGAT
GCACCGCCGCAACATCATCAGCCAGATGAAACGACTGGAAACCAAGGCTGCGAAGGAAGCAGCAGAGCCACCGCAGGAGCGCATTCTGGAAACAGCATGA

>KU760701.1 Pseudovibrio sp. P12 glutathione S-transferase domain protein (tdaB) gene, partial cds

ATGCAAAGAATTCTATACAGCCTCAGCGGTGCCTCCGGGCAGCACTATTCCCCGCATGTCTGGAAGATCATCATGTCCCTGAAGCACAAGGGACTCGCGTTTGACCTGAA
GCCCGTTTCCTTTGGCGACATTGCCAGTATTGAAGACGGTTCTTATACCAGCGTGCCCGTGCTCAACGATCAGGGCACCTTGCTGGGCGACAGCTTCGAGATCGCGCAAT
ACTTGGAAACAACCTACAAAGACACGCCGACCCTGTTCGATGGCGACGGTGGTCTGCAGCTCTCGCGTTTCGTGGAGAGCTATTGCAAGACAGTTCTGCACCCAACACTA
GCCGTCATTGCCGTGATGGATATGTACAACATCATGAGTGATGCGGACCGCGCTTACTTCCGCACTGCACGTGAAAAGCGCTTCGGAATGCCGCTGGAAGACGTACACG
CCAACCGCGATGCGGAGCTGGCAGCCTTGCCGGAGAAGCTCAACCCGGTGCGCGACCTGCTTGCAACGCAAGACTGGCTTGGCGGCGAGAAAGCCATGCATGCTGACTA
TGTTTTGTTCGGTACGCTGCAGTGGTGCAACGTCTGTTGCGCAGATCCATTTTTTGCAGCAACGGACCCTGTCCAGATGTGGTTCGAACGCATGCTTGATCTTCATGGTGG
TGTGGGCCGCAAAGCAGTCAACGAAGCCGTCGCAGCCT

7.8. Appendix VIII: Well-diffusion anti-microbial assay results