Author/s:
Byrne, Janet

Title:
Investigation of rare actinomycetes for novel antimicrobials

Date:
2020

Persistent Link:
https://hdl.handle.net/11343/243116

Investigation of Rare Actinomycetes for Novel Antimicrobials

MASTER OF PHILOSOPHY-BIOMEDICAL SCIENCE

Submitted with total fulfilment of the requirements for Master of Philosophy Degree

Janet Byrne

ORCID Identification Number: 0000-0002-9467-9338

Submitted: 26TH February 2020

MPhil Thesis – 2020

Candidate: Janet Byrne

Student Number: 962177

Title: Investigation of Rare Actinomycetes for Novel Antimicrobials

Status: Submitted with total fulfilment of the requirements for Master of Philosophy Degree

Date: 26th February 2020

Supervisors: Prof. Tim Stinear[1]

Dr. Sacha Pidot[2]

University: The University of Melbourne

Faculty: Medicine, Dentistry and Health Sciences

Department: Microbiology and Immunology

Abstract

Nocardia are a genus of ubiquitous environmental bacteria belonging to the phylum Actinobacteria. Genomics has revealed that *Nocardia* species are endowed with extensive and varied arrays of secondary metabolite biosynthetic gene clusters with the potential to produce natural products that have antibiotic properties. Furthermore, the abundance of such gene clusters within the *Nocardia* rivals that of *Streptomyces*, the signature genus among the Actinobacteria, owed to the fact that *Streptomyces* species have yielded many clinically used antibiotics. This project aimed to address the current antibiotic resistance crisis and the shortfall in new compounds within the drug discovery pipeline. A range of natural product discovery techniques were utilised amongst different Actinobacteria with a particular focus on a collection of species within the generally overlooked genus *Nocardia*.

This study had three primary objectives, the first was to use a traditional, high-throughput, empirical screen of 169 pathogenic actinomycetes predominantly from the genus *Nocardia*. These isolates were screened for antibiotic activity on 19 distinct growth media against a panel of five highly prevalent, multidrug resistant pathogens (*Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Enterococcus faecium* and *Acinetobacter baumannii*). Secondly, whole genome sequencing and bioinformatic interrogation of 100 *Nocardia* species was conducted to assess their genetic potential to biosynthesise natural products. This facilitated the selection of a single *Nocardia* isolate which possessed a non-ribosomal peptide synthetase locus that appeared to be unique amongst other *Nocardia* species. The locus was also transcriptionally silent. Bioengineering using promoter refactoring was employed to activate expression of this gene cluster, the product of which might have potential as a novel antimicrobial. Thirdly, by utilisation of liquid chromatography-mass spectrometry (LC-MS), bioinformatics and molecular networking, a metabolomic approach was employed to gain a

global secondary metabolic footprint of ten predicted “biosynthetically talented" *Nocardia* species grown on five distinct media types.

This project identified:

(i) A *Nocardia* sp. with activity against multidrug resistant *Acinetobacter baumannii*.

(ii) Two *Streptomyces* isolates (*Streptomyces cacaoi* and *Streptomyces* sp.) which exhibited antimicrobial activity against multidrug resistant *Escherichia coli* and *Acinetobacter baumannii* respectively. Secondary metabolite extracts from each of these producing isolates were investigated by LC-MS/MS and the resulting spectra was assessed for uniqueness through a dereplication data platform developed specifically for bacterial natural product identification. No hits for previously discovered metabolites were obtained suggesting that the antimicrobials discovered within this project appear to be unique and have potential as new drug leads for today’s ever-decreasing antibiotic discovery pipeline.

(iii) Four distinct families of bioactive secondary metabolites that were produced by multiple *Nocardia* species following LC-MS/MS and molecular network analysis. The identified secondary metabolites were correlated with genome sequence data to identify their probable biosynthetic origin in *Nocardia* species.

Personal Declaration

I. This is to certify that I am the soul author of this thesis in its entirety.

II. The intellectual content of this thesis is a product of my own work.

III. All assistance and sources have been duly acknowledged.

IV. This thesis has been submitted in less than the maximum word limit (50,000 words), inclusive of footnotes and exclusive of tables, maps, bibliography and appendices.

Acknowledgements

I would first like to thank my supervisors Prof. Tim Stinear and Dr. Sacha Pidot for backing me as a MPhil candidate and for all their support and advice along the way. Their faith in me has allowed me to live and study in Australia which has always been a huge dream of mine and for which I am very grateful.

To Marion Herisse, thank you for your advice and guidance in the laboratory, particularly regarding troubleshooting within the molecular cloning component of this project. Thank you to Dr. Liam Sharkey for taking such an interest in my project and for all your helpful advice along the way. To Dr. Koen Vandelannoote, thank you for all your hard work regarding to Nanopore sequencing which has allowed me to add additional depth to my findings.

A special thanks to Prof. Helge Bode and Dr. Nicolas Tobias in our collaborating laboratory at Goethe University, Frankfurt, Germany for performing LC-MS/MS analysis on all *Nocardia* isolates within the molecular networking component of this project.

Finally, to my amazing sister Ann Byrne (soon to be Dr. Byrne), thank you for all the laughs, your constant support, advice, and your amazing ability to lift my spirits.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

antiSMASH	Antibiotics and Secondary Metabolite Analysis Shell
ARC	Antibiotic remodelling compound
ArsR	Arsenic response repressor
ATCC	American Type Culture Collection
AURA	Antimicrobial Use and Resistance in Australia
BHI	Brain heart infusion
bp	Base pair
CDC	Centres for Disease Control
CRISPR	Clustered regularly interspaces short palindromic repeats
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FAS	Fatty acid synthase
gDNA	Genomic deoxyribonucleic acid
GNPS	Global Natural Product Social Molecular Networking
HA-MRSA	Hospital-acquired methicillin-resistant *Staphylococcus aureus*
HPLC	High-performance liquid chromatography
ISN	Isolate specific node
kb	Kilobase
KR	Ketoredutase
KS	Ketosynthase
LB	Luria broth
LC-MS	Liquid chromatography–mass spectrometry
LC-MS/MS	Liquid chromatography- tandem mass spectrometry
LHS	Left-hand side
m/z	Mass-to-charge ratio
MDR	Multidrug resistant
MDU	Microbiological Diagnostic Unit

MRSA	Methicillin resistant *Staphylococcus aureus*
MS	Mass spectrometry
MS/MS or MS^2	Tandem mass spectrometry
NCBI	National Centre for Biotechnology Information
NMR	Nuclear magnetic resonance
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OSMAC	One strain many compounds
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
PHID	Phenotypic identification
PKS	Polyketide synthase
QAC	Quaternary ammonium compounds
rDNA	Ribosomal deoxyribonucleic acid
RHS	Right-hand side
RNA	Ribonucleic acid
RNAP	Ribonucleic acid polymerase
RP-HPLC	Reversed-phase high-performance liquid chromatography
rRNA	Ribosomal ribonucleic acid
SBSPKS	Structure-based sequence analysis of polyketide synthases
SMBGC	Secondary metabolite biosynthetic gene cluster
SOA	Super optimal agar
SOB	Super optimal broth
SOC	Super optimal broth (catabolite repression)
sp.	Species
SQID	Sequencing identification
ST	Sequence type
T1-PKS	Type 1 polyketide synthase

TAE	Tris-acetate-EDTA
TAR	Transformation-associated recombination
TB	Tuberculosis
USD	United States dollar
UV	Ultraviolet
VISA	Vancomycin-intermediate *Staphylococcus aureus*
WHO	World Health Organisation
XDR-TB	Extensively drug-resistant tuberculosis

Chapter 1: Introduction

1.1.1 The antibiotic resistance crisis

The serendipitous discovery of penicillin by Alexander Fleming in 1928 sparked a revolution in the treatment of infectious diseases, saving countless lives and improving public health globally (Fleming, 1929). The discovery that microorganisms produce chemicals that kill other microorganisms gave rise to industrial scale bioprospecting to find bacteria (and other organisms) that produce valuable bioactive agents.

Existing and emerging multidrug resistant (MDR) bacteria are arguably one of the most troubling and prominent issues facing humanity. Rising antibiotic resistance is challenging the efficacy of our most robust antimicrobial agents in clinical use today. This is an unfortunate trend that has occurred with every new class of antibiotic discovered to date (Fig 1.1). Twenty new classes of antimicrobial drugs were brought to market between 1940 and 1962 (Coates et al., 2011). The past five decades has yielded few new drug classes; consequently, this period in history has been dubbed "the innovation gap" (Silver, 2011).

Each antibiotic drug class has its own distinct core structure (scaffold). The development of semi-synthetic derivatives of these scaffolds has been successful in prolonging efficacy of some antibiotic drug classes. This means that the vast majority of clinically utilised antimicrobials are based upon scaffolds of drug classes discovered between 1930-1960. Within the past two decades, scaffold repurposing of less efficacious drug classes has been too slow to keep pace with the onset of antibiotic resistance (Coates et al., 2011). Additionally, these repurposed antibiotics predominantly exhibit activity against Gram-positive bacteria, with many exhibiting poor efficacy against prominent Gram-negative pathogens such as: *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* sp. Sourcing antimicrobials capable of penetrating the outer membrane of Gram-negative bacteria

is particularly problematic due to constitutively overexpressed efflux pumps and highly selective porins (Domalaon et al., 2018). Moreover, these drugs are currently ineffective on multi-drug resistant Gram-positive nosocomial pathogens such as methicillin resistant *Staphylococcus aureus* and vancomycin resistant *Enterococci*. These persistent and worrisome human pathogenic bacteria are collectively termed ESKAPE pathogens (***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumoniae*, ***A****cinetobacter baumannii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* species) owing to their nosocomial predominance and their ability to "escape" the effects of the most efficacious antibiotic drugs in clinical use (Rice, 2008).

In addition to prominent nosocomial pathogens, extensively drug resistant *Mycobacterium tuberculosis* (XDR-TB) has been referred to by the World Health Association (WHO) as "the most profound challenge facing global health". It is estimated that 1.7 billion people are infected with latent tuberculosis (TB) worldwide (WHO, 2018). XDR-TB has been detected in 113 WHO membered states and exhibits resistance to two of the most potent anti-TB drugs in clinical use: isoniazid and rifampin, as well as all fluoroquinolones and at least one of the available second line injectables: amikacin, kanamycin or capreomycin (CDC, 2019).

The indiscriminate use of antimicrobials is a fundamental factor driving antibiotic resistance in bacteria (zur Wiesch et al., 2011). Bacterial resistance to antibiotics can occur de novo through spontaneous mutation in chromosomal loci or acquired through lateral gene transfer (Martinez and Baquero, 2000, Wang et al., 2001, Ochman et al., 2000). Pre-existing resistance determinants from a bacterial gene pool can be acquired through mobile genetic elements such as transposons, intergrons, insertion sequences and plasmids (Partridge et al., 2018, Lerminiaux and Cameron, 2018, Read and Woods, 2014). These factors are exacerbated under the selective pressure of antibiotics heavily used within a clinical setting.

Additional contributing factors to the current antibiotic resistance crisis include, drug misuse, incomplete dosage regimens, prophylactic consumption and over the counter availability in the developing world (WHO, 2017, AURA, 2017). Antibiotic overuse in agricultural and animal husbandry industries is also worrisome as many of these drugs are utilised clinically for the treatment of human disease (Smith et al., 2002, Peak et al., 2007).

These factors, coupled to low economic incentives, lack of novel drug development initiatives and constant re-discovery of previously known antibiotics within the research field has resulted in fifteen of the world's largest pharmaceutical companies abandoning this endeavour (Bartlett et al., 2013, Renwick and Mossialos, 2018, Simpkin et al., 2017, Projan, 2003). It is abundantly clear that "the golden years" of antibiotic discovery are truly over. The dire need for new drug leads to tackle this ever-progressing global issue has never been more apparent.

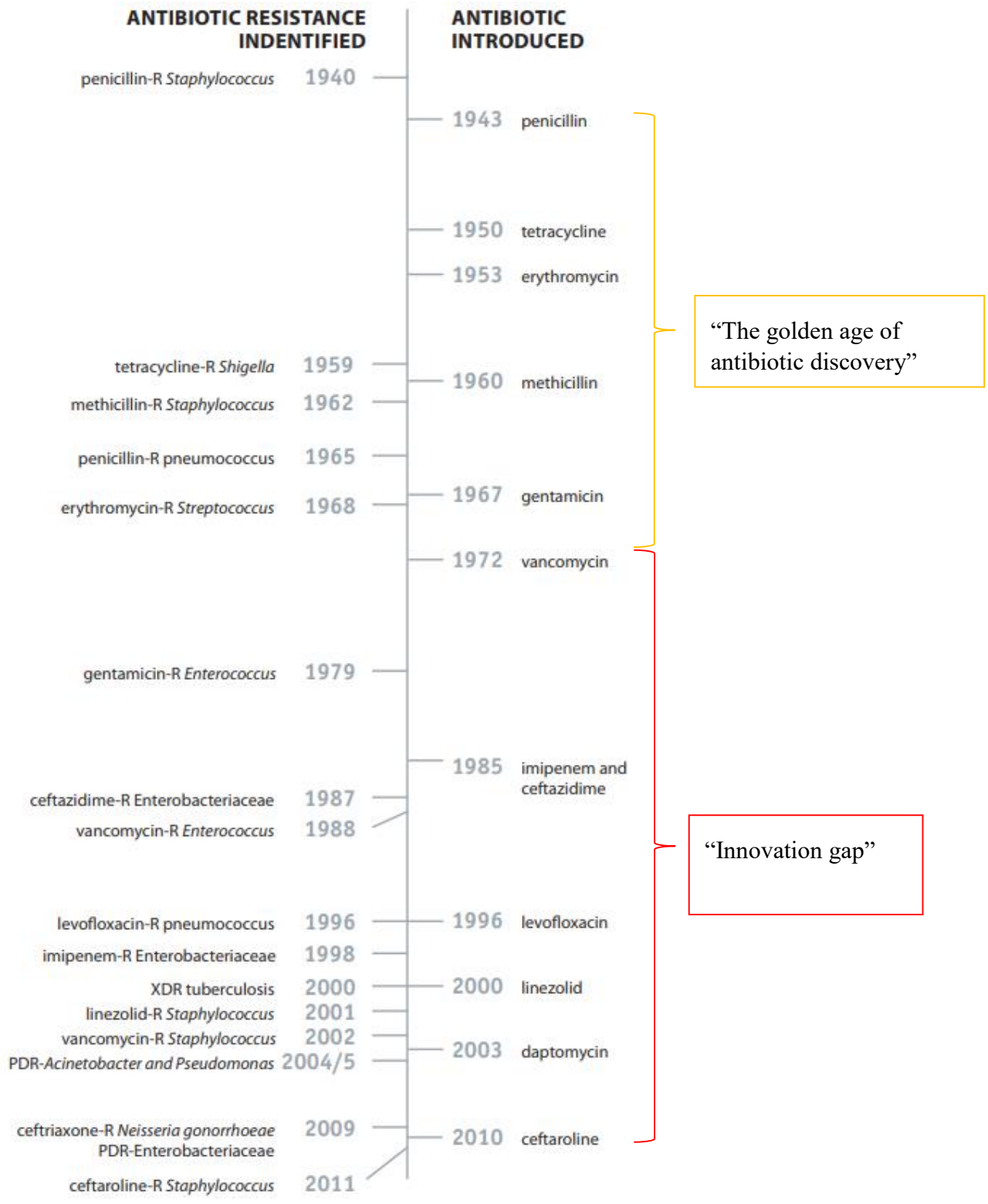


Fig 1.1 Antibiotic discovery pipeline from the 1940s to the present day. Years in which novel classes of antibiotics were discovered (right-hand side), with subsequent year of resistance development (left-hand side). Timeline adapted from (CDC, 2018).

1.1.2 The search for novel antibiotics

A single gram of soil can harbour up to 10^{10} bacterial cells and up to 5 x 10^4 different species (Torsvik et al., 1990, Roesch et al., 2007). Considering this, and the competitive nature of an environment with such a high bacterial density, it is not surprising that the soil contains an abundance of bacteria equipped with an array of genes responsible for antibiotic biosynthesis. These antibiotics are thought to have evolved as a means of defence against other microorganisms facilitating the maintenance of a favourable ecological niche. This has directed discovery efforts towards the exploitation of soil bacteria for new drug leads.

1.1.3 Empirical screening of soil bacteria for novel antibiotics

To date, empirical screening of bacteria present within soil samples has been the most fruitful source of novel antibiotics. This process involves screening whole bacterial cells, their fermentation products or chemical extracts for antimicrobial activity against a target organism. These screening methodologies only identify antimicrobial agents that are abundantly expressed and have easily visualised zones of inhibition. Such assays are not suited to identify potentially efficacious antimicrobials that are expressed at lower levels. Another important factor that hinders discovery efforts is the fact that empirical screening of bacteria is a culture-dependant process. This undoubtedly limits opportunities to find new antimicrobials given that less than 2% of environmental microorganisms are believed to be culturable under laboratory conditions (Wade, 2002). Despite this, empirical screening has yielded most antibiotic classes in clinical use today.

As the antibiotic discovery pipeline narrows, researchers are forced to think of more innovative ways to exploit soil bacteria for new antibiotic drug leads. The non-ribosomal peptide antibiotic, teixobactin represents the most recent class of antibiotic discovered in decades (Ling

et al., 2015). Teixobactin was discovered due to the innovative iChip, a device which increases the cultivability of previously unculturable soil bacteria up to 50-fold (Nichols et al., 2010).

Bacteria have evolved to thrive in some of the most extreme environments on the planet. Discovery efforts have shifted from terrestrial soil to exploiting more extreme ecological niches as potential reservoirs for novel antimicrobials. Such environments are prospected because they are usually unexplored, offer immense microbial diversity and new environments can yield new bacterial species for further interrogation. Attractive sites for microbial diversity in the extremes include, acid mine drainage sites, arid terrain, hyperacidic lakes, hot springs, deep-sea hydrothermal vents, sea ice, permafrost, deep-sea anoxic lakes and radioactively contaminated sites (Merino et al., 2019, Rafiq et al., 2019, Tortorella et al., 2018).

1.1.4 Dereplication to avoid re-discovery of previously known antibiotics

Exploiting new environmental niches can increase the likelihood of discovering novel antibiotic drug classes. However, over-mining these niches presents a major and unsurprising obstacle; the re-discovery of previously identified bioactives. Dereplication was first defined as a process of quickly identifying known chemotypes within a sample of interest (Beutler et al., 1990). Implementing a robust dereplication strategy is paramount to ensure that valuable time and resources are not wasted on the discovery of previously identified antimicrobials. This is one of the main reasons why there has been a mass exodus of “big pharma” from the natural product discovery arena.

Extensive libraries comprised of high-quality spectral data can be investigated through publicly available online databases. Commonly utilised databases in the natural product field include: The Dictionary of Natural Products and The Global Natural Product Social Molecular Networking (GNPS) (Mohimani et al., 2018). Such bioinformatic resources are commonly

complemented with genomics to illustrate the potential of an organism to synthesise novel compounds.

1.2 Genomic approaches to antibiotic discovery

The first prokaryotic genome to be sequenced in its entirety was that of *Haemophilus influenzae* via a combination of whole genome random shotgun and Sanger dideoxy chain termination sequencing (Fleischmann et al., 1995). This combination proved more efficient than Sanger sequencing alone which was heavily utilised from the late 1970s; proving pivotal in the early completion of The Human Genome Project, a globally orchestrated, 13-year endeavour with an approximate cost of $3 billion USD (Venter et al., 2015).

The advent of highly parallel, next generation sequencing technologies such as Illumina sequencing, PacBio single molecule real-time sequencing (SMRT) and Nanopore technologies have been fundamental in decoding the prokaryotic genome. These technologies are coupled to powerful bioinformatic tools to reconstruct and annotate genomes. Public, online databases such as NCBI GenBank house thousands of complete bacterial genome sequences. Such databases can be interrogated for gene clusters associated with antibiotic biosynthesis. *In silico* mining of genome sequence data targeting previously uncharacterised natural product gene clusters shows potential as a pathway to discover novel antimicrobial compounds.

1.2.1 Secondary metabolite biosynthetic gene clusters (SMBGCs)

The potential of a bacterial strain to produce secondary metabolites is usually easily identifiable from genome sequence data, presenting as large operons that are co-localised in distinct regions of the genome and termed secondary metabolite biosynthetic gene clusters (SMBGCs). These gene clusters are usually chromosomally encoded, although plasmid associated biosynthetic gene clusters have been reported (Martínez-Bueno et al., 1990). In addition to regulatory genes within the gene cluster that serve to activate, repress and terminate the resulting secondary

metabolite, global regulators outside of the cluster respond to environmental cues to trigger secondary metabolite biosynthesis. Such regulators can control numerous biosynthetic gene clusters throughout the genome (van der Heul et al., 2018). Several bioinformatic tools have been developed to facilitate genomic interrogation of an organism of interest. In 2011, the Antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) was introduced as a tool to specifically identify SMBGCs within sequenced bacterial and fungal genomes. Other platforms such as ClustScan, NP.searcher and structure-based sequence analysis of polyketide synthases (SBSPKS) target the non-ribosomal peptides and polyketide compound classes specifically associated with bacteria (Starcevic et al., 2008, Li et al., 2009b, Anand et al., 2010).

1.2.2 The non-ribosomal peptides

Non-ribosomal peptides are a major group of natural products produced by bacteria and fungi. Non-ribosomal peptides are assembled independently of the ribosome. Production of these valuable secondary metabolites is due to large, multi-domain enzymes known as non-ribosomal peptide synthetases (NRPS) (Fig 1.2). NRPS are comprised of repeated functional domains that select, activate and covalently couple amino acids from a pool of ~500 potential building blocks (Gaudelli et al., 2015). In addition to the 20 common amino acids, NRPS also utilise non-proteinogenic amino acids. Products of these gene clusters usually yield metabolites between 2-18 residues in length, that are often cyclised and extensively crosslinked. This class of natural products includes many clinically relevant antibiotics such as: β-lactam containing penicillins, the polymyxins, teicoplanin, daptomycin and vancomycin.

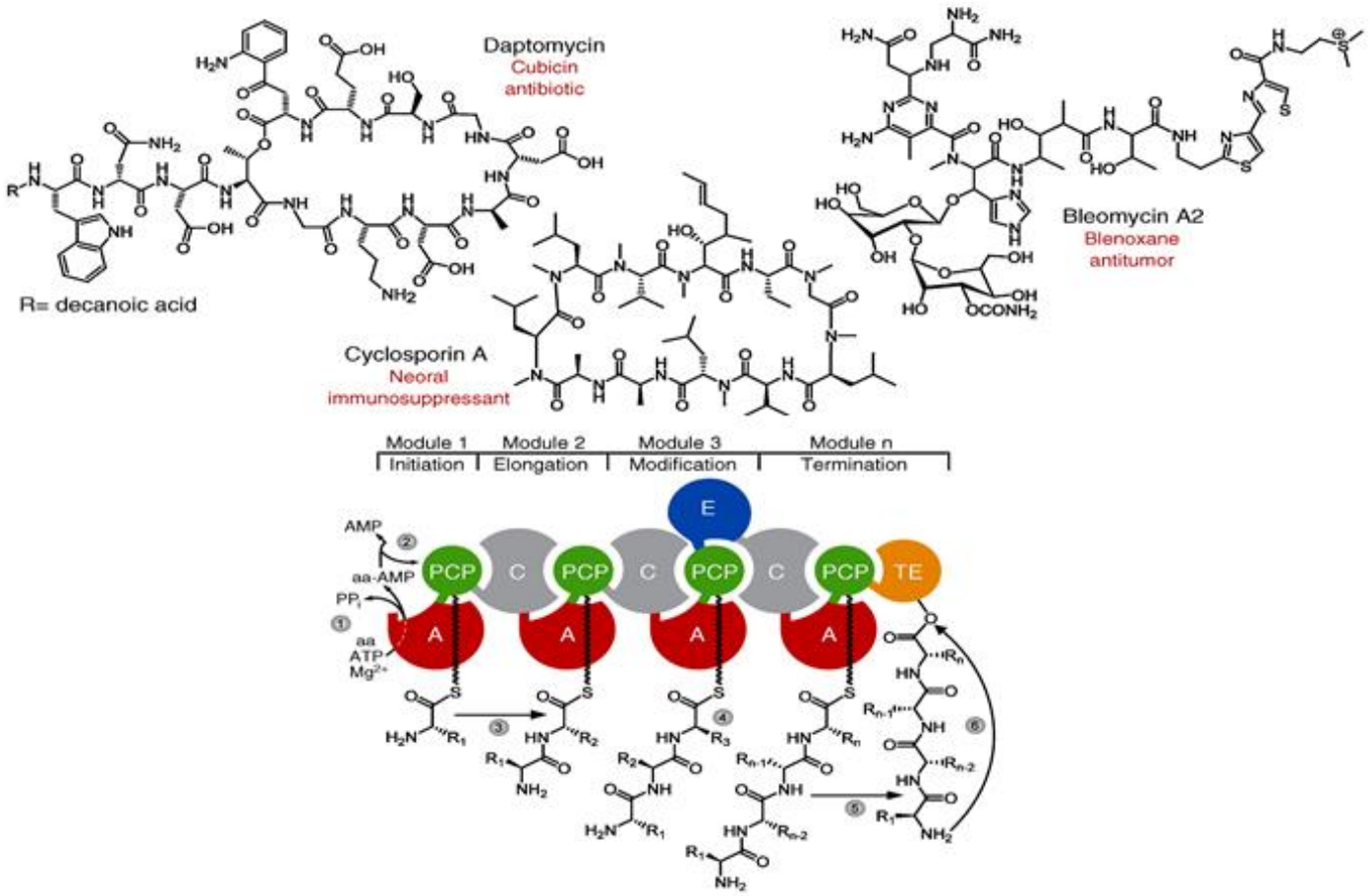


Fig 1.2 Non-ribosomal peptide synthetase assembly line. Therapeutically valuable peptides (daptomycin, cyclosporine A and bleomycin A20). Biosynthesis is initiated by selection from a pool of both standard and non-proteogenic amino acid substrates. Each module within this enzymatic unit contains multiple catalytic domains which serve to incorporate an amino acid to the growing chain. (Strieker et al., 2010)

1.2.3 The polyketides

Polyketides constitute another major class of natural products. They are produced by polyketide synthases (PKS) that catalyse the processive condensation of acetyl-CoA or propionyl-CoA to yield polymers with characteristic beta-keto functional groups (Chan et al., 2009). To date, over 10,000 polyketides have been characterised (Gonzalez-Lergier et al., 2005). PKS are large, multidomain enzymes that synthesise polyketides from malonyl-CoA or methylmalonyl-CoA extender units (Fig 1.3). Polyketide biosynthesis and fatty acid biosynthesis share similarities (Cronan and Thomas, 2009). PKS and fatty acid synthase (FAS) share homologous domains: ketosynthase (KS), acyltransferase (AT), acyl carrier protein (ACP), ketoreductase (KR) and dehydratase (DH) (Santin and Moncalian, 2018) (Marcet-Houben et al., 2008) (Cronan and Thomas, 2009). Actinomycetes and cyanobacteria are prolific producers of polyketides with antibacterial activity (Behal, 2000). Erythromycin, a broad-spectrum, cyclic polyketide antibiotic produced by *Saccharopolyspora erythraea* is assembled based on a on a six-module PKS biosynthetic gene cluster (Cortes et al., 1990). The polyketides include potent immunosuppressive, antiparasitic, and antitumoral agents (Ho et al., 1996, Sun et al., 2002, Sandmann et al., 2004). Clinically important antimicrobial polyketides include: rifamycin, azithromycin and clarithromycin.

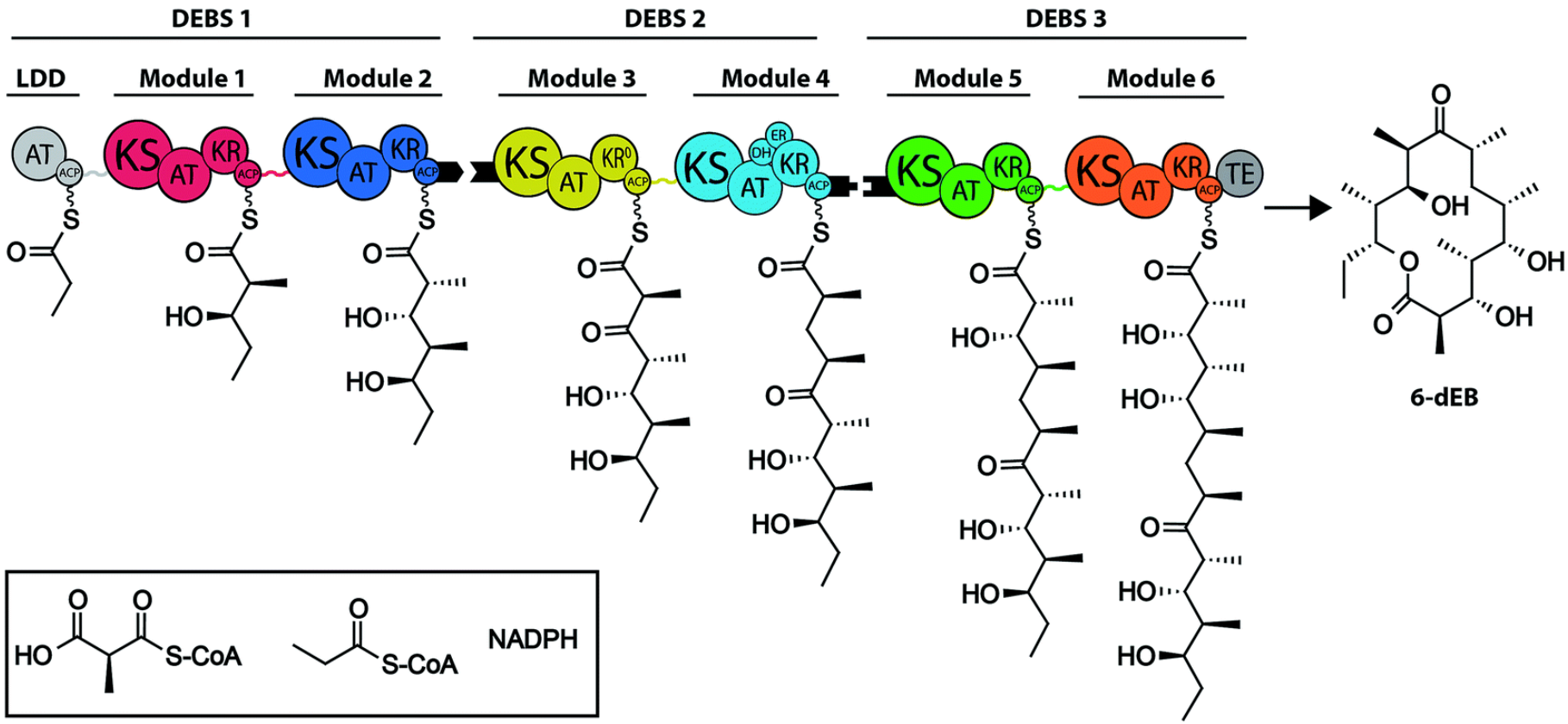


Fig 1.3 Polyketide synthase assembly line. The macrolide ring formation 6-deoxyerythronolide B (6-Deb) a precursor of the antibiotic erythromycin. Polyketides are formed by condensation of acyl-thioesters such as malonyl-CoA and methylmalonyl-CoA yielding metabolites with diverse structures and bioactivity, many of which undergo glycosylation, methylation, hydroxylation and oxidation after cyclisation (Klaus and Grininger, 2018).

1.2.4 The Actinomycetes: A major source of novel antimicrobials

The phylum Actinobacteria represents one of the largest taxonomic units within the domain Bacteria (Barka et al., 2016). Actinomycetes are Gram-positive, predominately aerobic, ubiquitous in terrestrial and aquatic ecosystems and prolific producers of secondary metabolites. The actinomycetes possess impressive repertoires of SMBGCs, which can range from 25-50 clusters within a single strain (Baltz, 2018). These organisms are undoubtedly the most relevant resource for medically important natural products with varied biological activity and structural diversity. Two-thirds of all antibacterial agents utilised clinically are derived from terrestrial soil actinomycetes, predominately from the genus *Streptomyces* (Fair and Tor, 2014).

1.2.5 The Streptomyces

Streptomyces are Gram-positive, branched filamentous bacilli belonging to the phylum Actinobacteria. These aerobic soil saprophytes were originally described by Waksman in 1932 (Waksman and Henrici, 1943). *Streptomyces* are highly regarded and deserving of the title "most prolific producers of secondary metabolites" yielding 17 classes of clinically utilised antibiotics since 1944 (Table 1.1) (Berdy, 2005). In addition to antibacterial agents, products of these gene clusters include many clinically utilised immunosuppressive, antitumorals, antivirals, antifungals and antiparasitic agents (Ackermann et al., 1996, Raveh et al., 2013, Engelhardt et al., 2010, Pimentel-Elardo et al., 2010, Behal, 2000).

Table 1.1 Key antibiotics produced by the genus *Streptomyces* (de Lima Procópio et al., 2012).

Antibiotic	Year discovered	Species
streptomycin	1944	*S. griseus*
cephalosporins	1945	*S. clavuligerus*
chloramphenicol	1949	*S. venezuelae*
neomycin	1949	*S. fradiae*
tetracycline	1950	*S. aureofaciens*
nystatin	1950	*S. noursei*
viomycin	1951	*S. vinaceus* *S. capreolus*
virginiamycin	1952	*S. pristinaespiralis* *S. virginiae*
lyncomycin	1952	*S. lincolnensis*
cycloserine	1955	*S. garyphalus*
vancomycin	1956	*S. orientalis*
noviobiocin	1956	*S. niveus*
kanamycin	1957	*S. kanamyceticus*
fosfomycin	1969	*S. fradiae*
ribostamycin	1970	*S. ribosidificus*
daptomycin	2003	*S. roseosporus*
platensimycin	2006	*S. platensis*

1.2.6 The Nocardia

The *Nocardia* represents a distinct genus of Gram-positive, soil dwelling saprophytic Actinobacteria of which there are 80 species; over 30 of which are opportunistic pathogens (Brown-Elliott et al., 2006). In healthy individuals, infection occurs usually via traumatic cutaneous inoculation (Outhred et al., 2011). However, in immunocompromised hosts the clinical spectrum is variable. Common infection sites are the lungs (progressive pneumonia), central nervous system (meningitis) and the heart (endocarditis) (Menéndez et al., 1997, Ellner and Bennett, 1976, Castelli et al., 2011). Disseminated disease in the kidneys, joints and gastrointestinal tract have also been reported (Torres et al., 2000). *Nocardia* species commonly observed to cause disease in humans include: *N. nova*, *N. abscessus*, *N. transvalensis*, *N. farcinica*, *N. cyriacigeorgica* and *N. brasiliensis* (Brown-Elliott et al., 2006).

1.2.7 Antibiotics produced by Nocardia species

Exploiting human pathogenic bacteria for the production of bioactive secondary metabolites is a relatively unexplored field. Pidot and Coyne highlighted the potential of underexplored, human pathogenic bacteria from proteobacteria and firmicutes phyla as novel sources of antimicrobial agents (Pidot et al., 2014). The potential of some *Nocardia* to cause disease sparked interest in exploring Nocardial genomes to better understand pathogenicity. Komaki *et al.,* sequenced the genome of seven *Nocardia* species commonly associated with human infections. Genomic analysis of these pathogens illustrated that this generally overlooked genus has great potential as an untapped resource for novel antibiotics. This team surmised that *Nocardia* genomes are similar in size and carry just as many SMBGCs as *Streptomyces*, the accepted "superstar" within the field of antibiotic discovery. They noted that these gene clusters varied considerably amongst *Nocardia* species and put forth *Nocardia brasiliensis* as a particularly promising candidate endowed with the highest quantity of SMBGCs overall (Komaki et al., 2014). Utilisation of comparative genomics and bioinformatic resources echoes

these findings showing that the genus *Nocardia* harbours an impressive array of SMBCGs that rival the *Streptomyces.* The *Nocardia* therefore represent a potentially untapped source for novel secondary metabolite discovery.

Tubelactomicin A, a novel lactone antibiotic was discovered following empirical screening of *Nocardia* with activity against multidrug resistant acid-fast organisms (Igarashi et al., 2000). Nargenicin, a macrolide antibiotic with strong anti-MRSA activity was isolated from *Nocardia tenerifensis* (Sohng et al., 2008, Pidot et al.). Nocathiacin (I-III) belonging to the thiazolyl peptide family was isolated from *Nocardia* and exhibited potent activity against MRSA, MDR *Enterococcus faecium* and penicillin resistant *Streptococcus pneumoniae* (Li et al., 2003). N-demethylvancomycin, a novel vancomycin analogue was also isolated from *Nocardia orientalis* (Boeck et al., 1984). The bioactive end products of this genus are not just limited to antimicrobial agents, antifungals, antitumorals and immunosuppressive agents have also been isolated from *Nocardia* species (Higashide et al., 1977, Muroi et al., 1980, Shigemori et al., 1998).

1.2.8 Transcriptionally silent SMBGCs

Although SMBGCs are easily identifiable from sequencing data, correlation of a product with gene cluster of origin remains challenging. Exploiting the biosynthetic potential of "talented" actinomycetes such as *Nocardia* via empirical screening is hampered by the fact that many of these SMBGCs are transcriptionally silent, owing to tight transcriptional and/or translational regulation (Rutledge and Challis, 2015). Moreover, induction of gene expression relies upon a myriad of environmental cues that are difficult to mimic *in vitro*. These cryptic clusters may represent novel sources of antimicrobials with unique mechanisms of action (McLean et al., 2019). *In vitro* SMBGC activation has been a key undertaking within the field of natural product discovery. To date, numerous strategies have been developed to address this challenge.

1.3 Methods for the activation of transcriptionally silent SMBGCs

1.3.1 The one strain many compounds approach (OSMAC)

Media composition can strongly influence the secondary metabolomic profile in bacterial fermentations. Although there is no consensus on the "ideal media", a common strategy is to screen with a range of media from rich, well-defined to minimal aiming to potentially awaken cryptic SMBGCs. Induction of genes encoding secondary metabolites occurs in the nutrient limited, late growth phase (idiophase) (Ruiz et al., 2010). Systematic alteration of culture conditions has proven successful in unlocking the biosynthetic potential of bacteria.

Bode *et al.,* illustrated how minor alterations in cultivation parameters can lead to big differences in the resulting secondary metabolome. This sequence-independent approach involves altering growth conditions of a bacterial culture. This method has been termed the OSMAC (one strain many compounds) approach and has been used extensively by the natural product research community as a means of activating cryptic biosynthetic gene clusters. Common parameters that are easily manipulated include: media composition, aeration, pH, incubation temperature and vessel type. Adapting this approach, the team discovered over 100 compounds from six microorganisms, covering 25 classes, 20 of which were isolated from a single organism (Bode et al., 2002). By simply increasing the incubation temperature from 27°C to 42°C in *Streptomyces venezuelae* cultures Doull *et al.,* induced a heat shock response. This resulted in a significant increase in biosynthesis of the type II polyketide antibiotic jadomycin B which is ordinarily expressed at low levels in *Streptomyces.* Jadomycin exhibits potent activity against Gram-positive bacteria. The team reported a similar response following exposure of the same isolate to 6% ethanol (Doull et al., 1994).

1.3.2 Co-culturing bacteria to induce antibiotic biosynthesis

Another commonly utilised strategy to activate secondary metabolism in bacterial fermentations is to introduce a challenge organism (co-culture) (Onaka et al., 2001, Okada and Seyedsayamdost, 2017, Pettit, 2011). Secondary metabolites play a key role in bacterial communication. Inter/intraspecies "talk" can be exploited by using such metabolites as a physiological trigger to awaken ordinarily silent SMBGCs. The underlying premise is that bacteria in nature, never grow in isolation. They are in constant interaction with macro/microorganisms and their products, which can be competitive or symbiotic in nature. Attempting to mimic these conditions in monocultures is challenging therefore the addition of challenge organisms is a focus of natural product research. Onaka *et al.,* showed that challenging *Streptomyces* sp. with mycolic acid containing bacteria such as *Tsukamurella, Rhodococcus* and *Corynebacterium* activated ordinarily transcriptionally silent SMBGCs (Onaka et al., 2011). This led to the discovery of a novel polycyclic polyketide antibiotic alchivemycin A which exhibited potent antimicrobial activity against the Gram-positive bacterium *Micrococcus luteus* (Igarashi et al., 2010). The macrolide antibiotic chalcomycin was discovered when *Streptomyces* sp. Mg1 was co-cultured with *Bacillus subtilis* (Barger et al., 2012). Chalcomycin facilitated cell lysis and colony degradation in the competing *Bacillus subtilis*.

Exploration of more complex co-culture communities for novel antimicrobials has also been fruitful. Pishchany *et al.*, explored the potential of a complex community of actinomycetes to express antimicrobial natural products that were not ordinarily expressed in axenic equivalents. The team created a nine strain actinomycete culture and utilised whole genome sequencing to deconvolute the community, ultimately identifying a producer-inducer pair: *Amycolatopis* sp./*Streptomyces coelicolor* respectively. This led to the discovery of the polyketide

amycomicin which exhibited potent activity against *Staphylococcus aureus* (Pishchany et al., 2018).

1.3.3 SMBGC activation using chemical elicitors

The addition of small molecules (chemical elicitors) to bacterial cultures is a recognised strategy to induce antibiotic biosynthesis and expand the chemical diversity within a strain of interest. This approach not only serves to activate transcriptionally silent SMBGCs but also to enhance the synthesis of metabolites ordinarily expressed at low levels. The exact mechanisms of gene induction via chemical elicitors is poorly understood in actinomycetes. It is thought that such elicitors can directly influence transcription of SMBGCs or induce transcriptional activators (Abdelmohsen et al., 2015).

Low concentrations of the rare earth elements scandium and lanthanum have been shown to increase expression of SMBGCs up to 12-fold in *Streptomyces coelicolor* (Tanaka et al., 2010). The addition of Dimethyl sulfoxide (DMSO), ethanol and sodium butyrate to *Streptomyces* cultures have also been successful in SMBGC activation (Chen et al., 2000, Pettit, 2011, Moore et al., 2012, Abdelmohsen et al., 2015).

The presence of enzyme inhibitors in bacterial cultures can vastly affect the resulting secondary metabolome. The first synthetic molecule shown to alter secondary metabolism in bacteria was an inhibitor of the enzyme phosphopantetheinyl transferase which is involved in acyl carrier protein activation in fatty acid biosynthesis. The molecule, named 7ae was shown to enhance the expression of the polyketide antibiotic actinorhodin in *Streptomyces coelicolor*. The mechanism of action of 7ae remains elusive (Foley et al., 2009). This approach has also been utilised to induce secondary metabolism in fungi. Shwab *et al.*, treated fungal isolates belonging to the genus *Penicillium* and *Alternaria* with trichostatin A, a histone deacetylase

inhibitor. The team reported a statistically significant increase in the expression of previously unknown secondary metabolites in comparison to axenic fungal cultures (Shwab et al., 2007).

1.3.4 Sub-inhibitory levels of antibiotics to induce antibiotic biosynthesis

The introduction of low dose antibiotics to bacterial fermentations has been successful in activating cryptic gene clusters. This high-dose toxicity, low-dose stimulatory effect is known as hormesis (Okada and Seyedsayamdost, 2017, Davies et al., 2006). The thiazole oligopeptide goadsporin exhibits potent activity against *Streptomyces* sp. However, sub-inhibitory levels of this antibiotic were shown to have a morphogenic effect as well as promoting the production of pigments and antibiotics in 42 randomly selected *Streptomyces* strains (Onaka et al., 2001). The antibiotic trimethoprim was reported to act as a global activator of five SMBGCs when added to cultures of to *Burkholderia thailandensis* at sub-lethal concentrations (Seyedsayamdost, 2014).

The most comprehensive approach to chemical elicitation was conducted by Nodwell and colleagues. This involved screening a library of >30,000 compounds for their ability to induce SMBGCs in *Streptomyces coelicolor*. The team identified a total of 19 compounds which increased the production of the blue pigmented polyketide antibiotic actinorhodin. Four of the 19 compounds belonged to a family of antibiotics referred to as the ARC2 series. These compounds, were structurally related to the antibacterial and antifungal agent triclosan which inhibits the enzyme Enoyl-acyl carrier protein reductase (FabI) in fatty acid biosynthesis. This study highlights that bacterial primary and secondary metabolism are more intrinsically linked that previously thought.

1.3.5 Mutagenic strategies to antibiotic discovery

Great efforts have been made to induce or enhance secondary metabolite biosynthesis. Incorporating random/targeted mutations in a biosynthetic gene cluster or targeting ribosomes and polymerases can enhance gene expression enough to chemically characterise the gene product, assess bioactivity and decide on the commercial potential of the metabolite. Ribosomal engineering has been effective at increasing antibiotic production in *Streptomyces* sp. The creation of ribosomal mutant strains of *Streptomyces lividans* activated a silent gene cluster resulting in the production of the antibiotic actinorhodin (Shima et al., 1996). Gene cluster induction was due to a mutation in the *rspL* gene encoding the ribosomal protein S12, a component of the 30S subunit. This mutation stabilises the ribosomal 70S complex which enhanced secondary metabolite biosynthesis.

In *Streptomyces coelicolor,* deletion of *relA* (a global transcriptional regulator) or a mutation in *relC* (encoding 50S ribosomal protein L11) results in the inability of the strain to synthesise the alarmone guanosine-3',5'-bisdiphosphate (ppGpp). This consequentially diminishes the expression of genes associated with the biosynthesis of the antibiotic actinorhodin. Ochi *et al*., showed how introducing mutations in the rifampicin resistance conferring β-subunit of RNA polymerase (RNAP) forces it to behave stringently in the absence of ppGpp suggesting that modification of antibiotic resistance conferring regions on RNAP may mimic ppGpp binding, increasing antibiotic biosynthesis (Xu et al., 2002, Ochi and Hosaka, 2013).

Goa *et al.,* introduced a reporter guided mutant selection protocol for the targeted activation of cryptic gene clusters in *Streptomyces* sp. Genetic diversity was introduced via genome wide, random UV mutagenesis coupled to a reporter system to select for the desired cluster activated mutants. The team successfully validated this protocol by reactivating the gene cluster responsible for jadomycin biosynthesis in *Streptomyces venezuelae* ISP5230 and went on to

activate a cryptic gene cluster in *Streptomyces* PGA64 yielding two new aminoglycoside antibiotics, gaudimycin D and E (Guo et al., 2015).

1.3.6 Heterologous hosts for antibiotic biosynthesis

Expression of an entire SMBGC in a more amenable heterologous host has been a successful method for gene cluster activation. Commonly utilised surrogate hosts include well characterised *Streptomyces* sp., *Pseudomonas* sp. and *Escherichia coli*. This involves mobilising the gene cluster into a vector and engineering the cluster to be functionally expressed in the heterologous host. Cloning vectors such as plasmids, cosmids, fosmids and bacterial artificial genomes are combined with transformation protocols such as polyethylene glycol or $CaCl_2$ assisted transformation, electroporation, intergeneric conjugation and transduction in a heterologous host (Galm and Shen, 2006). However, transcriptional and translational regulatory signals and codon usage differences in the heterologous host can hinder formation of a functional end product, if any at all (Galm and Shen, 2006). Additionally, this methodology is limited to smaller gene clusters usually those < 40 kb. SMBCGs are much larger, ranging from 20 kb up to 200 kb challenging current recombinant technology (Nah et al., 2017).

1.3.7 Promoter refactoring to activate transcriptionally silent SMBGCs

Promoter refactoring by replacing the natural promoter with a constitutive or easily inducible promoter can circumvent the issue of expressing large gene clusters in a heterologous host. Olano *et al.*, successfully utilised this approach in *Streptomyces albus* J10074. The team successfully activated two transcriptionally silent biosynthetic gene clusters by utilising the constitutive promoter *PermE*. This promoter drives the expression of erythromycin biosynthetic genes in *Saccharopolyspora erythraea.* This resulted in the activation of a cluster responsible for blue pigment production and a second cluster that yielded two novel members

of the macrolactam family, *6-epi*-alteramide A and B (Olano et al., 2014). The same team overexpressed the positive regulatory gene *pimM* which is native to the pimaricin gene cluster in *Streptomyces natalensis.* The product of *pimM* activates the LuxR family of transcriptional regulators. An integrative vector was utilised under the control of *PermE* for chromosomal integration in *Streptomyces albus* J10074 which increased the expression of *6*-epi-alteramide A and B four-fold.

1.3.8 Transformation-associated recombination (TAR)

Extraction and subsequent sequencing of bacterial DNA directly from environmental samples eliminates the need for prior cultivation, shedding light on the biosynthetic potential of the unculturable majority. Combining this metagenomic approach with heterologous expression of SMBGCs cloned from environmental DNA shows promise as a means to discover novel antimicrobials. Reassembly of gene clusters from sections of smaller overlapping clones in cosmid vector systems circumvents the issues that arise from cloning large gene clusters. However, this can be labour intensive.

Kim *et al.*, introduced a less arduous technique which successfully captured large SMBGCs on overlapping environmental DNA cosmid clones. The team selected the type II polyketide synthase biosynthetic gene cluster architype, which are highly conserved and common amongst soil bacteria. Reassembly of these large clusters was possible in *Saccharomyces cerevisiae*, via a technique termed transformation-associated recombination (TAR), which is reliant upon homologous recombination to integrate a target sequence from a gDNA pool (Kim et al., 2010). The target DNA and a capture vector with short homology arms (40bp) specific to the sequence of the flanking region of interest were used. Transformation in *S. cerevisiae* yielded a stable plasmid containing the target DNA sequence, which was then chromosomally integrated using phage integrases. TAR cloning systems have been successfully combined with constitutive

promoter refactoring increasing the likelihood of gene cluster activation. This led to the discovery of the antiproliferative agent lazarimide A and B (Montiel et al., 2015).

1.3.9 CRISPR/Cas9 system

In the late 1980s the sequence of a gene was determined in *Escherichia coli* with no known function (*iap)* (Ishino et al., 1987). It wasn't until 2002 that this sequence was termed: clustered regularly interspaced short palindromic repeats (CRISPR) (Jansen et al., 2002). These repeated sequences were separated by non-repeating DNA sequences termed spacers. It was noted that 40% of sequenced bacterial genomes possessed CRISPR clusters (Mojica et al., 2000). These clusters were shown to lie adjacent to highly conserved genes termed CRISPR associated genes (Cas). The non-repeating spacer sequences were shown to be viral in origin suggesting that CRISPR is a component of a bacterial adaptive immune response to phage attack. RNA transcribed from these spacer units dictate the cleavage specify of the Cas 9 enzyme highlighting its potential as a genomic editing tool. (Pourcel et al., 2005, Adli, 2018).

The programmable capacity of the Cas9 enzyme has been exploited in the natural product discovery field to successfully edit the genomes of *Streptomyces*, *Myxobacteria*, *Bacillus*, *Pseudomonas*, and *Cyanobacteria* (Tong et al., 2018). Zhang *et al.,* successfully utilised the CRISPR-Cas9 system to knock-in constitutive promoters in five *Streptomyces* sp., successfully activating multiple SMBGCs of different classes, including an unidentified type II polyketide in *Streptomyces viridochromeogenes* (Zhang et al., 2017). Similar promoter knock-in approaches were successful in activating a transcriptionally silent gene cluster encoding the type I polyketide auroramycin in *Streptomyces roseosporus*, the product of which exhibited potent anti-MRSA activity (Lim et al., 2018).

1.4 Metabolomics

Primary metabolism in bacteria is energy yielding and essential for cell viability. Conversely, microbial metabolites that are not essential for bacterial growth but do infer a competitive advantage are termed secondary metabolites and include antimicrobials, pigments and toxins. Metabolomics is the study of low molecular weight organic compounds that are end products or intermediates of metabolism in microorganisms, cells, biofluids and tissues. The molecular footprint of an organism is dynamic, ever-changing and heavily influenced by genetic and environmental factors. The concentration of metabolites in a dataset is reflective of the underlying biochemistry facilitating an extremely accurate interoperation of the molecular phenotype in question. Metabolomics is the most recent addition to the "omics" field (genomics, transcriptomics and proteomics). The past two decades has seen an exponential rise in publications focused on metabolomic analysis of biological systems (Rochfort, 2005).

Gene induction of bacterial SMBGCs is poorly understood, tightly regulated, and heavily reliant upon an array of external environmental cues. The metabolomic profile of a bacterial candidate can be compared under a range of culturing parameters, shedding light on optimal growth conditions for the induction of secondary metabolism. Interrogating the potential of a strain through metabolomics involves first culturing the organism in question. This is followed by extracting secondary metabolites directly from the growth media with an organic solvent. Commonly utilised solvents include: ethyl acetate, methanol or *n*-butanol. Extracts from each condition are subsequently spectroscopically interrogated. Analysing the metabolomic profile of biosynthetically talented bacteria allows visualisation of well conserved metabolites. This facilitates strain prioritisation of bacterial species that appear to express a more unique cohort of metabolites, saving time and resources.

1.4.1 High-performance liquid chromatography-mass spectrometry

High-performance liquid chromatography (HPLC) remains the most commonly utilised technique for the separation of chemical entities from crude organic extracts. HPLC is robust, fast and applicable to the separation of all groups of microbial metabolites. HPLC is commonly coupled with mass spectrometry (MS) for compound identification. (Ackermann et al., 1996, Hubert et al., 2015). MS is highly accurate technique that identifies metabolites (if known) based on mass, elemental structure and fragmentation pattern. Additionally, the high sensitivity of MS facilitates dereplication of metabolites present within a sample at extremely low levels. This LC-MS coupling is the go-to technique for streamlining natural product discovery efforts. Nuclear magnetic resonance (NMR) is another predominant technique for structural determination and dereplication of natural products. NMR was initially less sensitive than MS but advances in this area have bridged this gap (Hubert et al., 2015, Marshall and Powers, 2017).

1.5 Molecular networking

Molecular networking is a powerful tool in the field of metabolomics, facilitating the visualisation of complex mass spectrometric data. Molecular networking generates large metabolic datasets by grouping MS fragmented ions into sub-networks that are representative of molecular families which have structural similarities. This generates a comprehensive global picture of the secondary metabolic profile of an organism of interest.

1.5.1 The Global Natural Products Social Molecular Network (GNPS)

The Global Natural Products Social Molecular Networking (GNPS) is the world's largest and most comprehensive natural product mass spectrometry repository offering open-access to community acquired, raw, processed or identified tandem mass spectrometry data (MS/MS or

MS_2). This platform is equipped with an algorithm to comparatively assess structural similarities between fragment ions in mass spectral data. A central component of this platform is an extensive, community aquired reference library of spectral data which facilitates strain priortisation and rapid dereplication of previously known metabolites (Wang et al., 2016).

The natural product dereplicator component of GNPS is a bioinformatic tool for annotaton of known natural products in MS/MS data using an *in silico* fragmentation tree. Initially, this dereplication algorithm was estblished to identify peptiditic natural products only. The introduction of the improved DEREPLICATOR+ algorithm has now been expanded with greater confidence in the annotation of peptidic natural products as well as covering drug classes such as the polyketides, terpenes, benzenoids, alkaloids and flavonoids. This update yields five times more hits than the previous dereplication tool by searching over 200 million tandem mass spectra. DEREPLICATOR+ was further complemented by the addition of VarQuest which facilitates the search for bioactive varients of known peptidic natural products.

1.5.2 Applications of molecular networking

Crüsemann *et al.,* assessed the metabolomic profile of 146 marine *Salinospora* and *Streptomyces* sp. under a range of different growth and extraction protocols. The team then constructed a molecular network and identified 15 distinct molecular families of natural products and their analogues, as well as providing valuable insight into optimal growth and extraction protocols (Crusemann et al., 2017). The novel NRP/PK hybrid peptide myxoprincomide was discovered by comparing the metabolomic profile of wild type and knock- out stains of *Myxococcus xanthus* following rounds of targeted mutagenesis (Cortina et al., 2012). Senges *et al.,* compared the supernatants of *Streptomyces chartreusis* NRRL 3882 grown on a range of different media types. The team, guided by molecular networking, highlighted the abundance of uncharacterised chemistry in *S. chartreusis* (Senges et al., 2018).

Genomic analysis of *S. chartreusis* indicated the presence of 128 SMBGCs. The team identified over 1000 distinct metabolites, only 22 of which have been previously characterised.

1.6 Project rationale and thesis outline

The indiscriminate use of antibiotics and the propensity of bacteria to rapidly evolve resistance has meant that the treatment of infectious disease is reverting to that of a pre-antibiotic era. This sobering fact highlights the dire need for new drug leads within the field of natural product discovery. This project aims to address the current antibiotic crisis and attempt to expand the diminishing antibiotic discovery pipeline by focusing on the biosynthetically talented genus *Nocardia.*

Nocardia species harbour an impressive repertoire of biosynthetic gene clusters and represent an untapped source with the potential to yield novel antibiotics. This project employed a combination of traditional empirical screening methodologies combined with cutting-edge bioengineering and metabolomic profiling to generate new antibiotic leads from this generally overlooked genus.

1.6.1 Chapter three: Empirical screening for novel antibiotics

This chapter draws influence from the OSMAC approach commonly utilised by natural product discovery research groups. The varying parameter in this case was growth media. An empirical screening of 169 pathogenic actinomycetes predominantly from the genus *Nocardia* (156 isolates) was conducted. Each isolate was screened on a total of 19 different media types; attempting to activate SMBGCs within individual strains. The composition of each media type ranged from well-defined to nutrient limited, increasing the probability of antibiotic biosynthetic gene cluster activation.

Each isolate was provisionally screened against two antibiotic sensitive bacteria which were representative of Gram-positive and Gram-negative groups: *Staphylococcus aureus* (Newman)

and *Escherichia coli* (DH10B) respectively. Isolates that inhibited the growth of Newman and/or DH10B were subsequently screened for antibiotic activity against a panel of five, prolific, multidrug resistant pathogens: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae, Acinetobacter baumannii* and *Escherichia coli*.

1.6.2 Chapter four: Gene cluster activation via promoter refactoring

This chapter focused on a yet unnamed *Nocardia* species held at the Peter Doherty Institute, University of Melbourne. Genome sequencing and subsequent bioinformatic analysis indicated that the genome of this isolate harbours 20 putative secondary metabolite biosynthetic gene clusters. One of these clusters comprised a 24 kb non-ribosomal peptide synthetase locus that appeared to be transcriptionally silent. This chapter utilised recombinant DNA technology, attempting to activate this gene cluster by refactoring the native promoter with *PermE,* a strong constitutive promoter. Cluster activity was subsequently assessed by inhibition screens and HPLC analysis.

1.6.3 Chapter five: Metabolomics and molecular networking

The secondary metabolomic profile of ten pathogenic *Nocardia* isolates of unknown potential was investigated. Complete genome sequences were obtained using hybrid PacBio and Illumina assemblies. All isolates were cultured on five of the best performing media types that facilitated secondary metabolite biosynthetic gene induction (Chapter 3). The secondary metabolomic footprint of each isolate was interrogated via LC-MS/MS analysis and subsequent dereplication through the Global Natural Products Social Molecular Networking (GNPS) platform. A comprehensive spectral network was constructed using the open source bioinformatic platform Cytoscape. The molecular network gave a comprehensive picture of the molecular diversity of each strain as well as highlighting core secondary metabolites readily expressed by all isolates and distinct molecular families present within the extracts. This

analysis streamlined the process of strain prioritisation for those isolates showing greater potential as a source of novel antimicrobials; as well as shedding light on the ideal media type for secondary metabolite gene induction in *Nocardia* species.

Chapter 2: Materials and methods

2.1 Bacterial overlays

All pathogenic actinomycete isolates for empirical screening were sourced from the Microbiological Diagnostic Unit, Peter Doherty Institute, University of Melbourne. A total of 169 pathogenic actinomycetes were screened for antimicrobial activity, each isolated from patients in Victoria between 2002 and 2016 (appendix 1). Initial phenotypic identification and subsequent 16S rRNA sequencing was conducted at the Peter Doherty Institute prior to the commencement of this project.

2.1.1 Isolate storage

All actinomycete isolates were stored at the Microbiological Diagnostic Unit, Peter Doherty Institute, University of Melbourne at -80°C in glycerol (20%) prior to the commencement of this project.

2.1.2 Actinomycete stocks

Each actinomycete isolate was subcultured twice prior to screening. Isolates were streaked onto brain heart infusion agar (BHI) and incubated at 30°C for seven days. A single colony of each isolate was inoculated into 10 ml BHI broth and incubated at 30°C for an additional seven days with constant agitation (200 rpm). Bacterial aggregations indicative of actinomycete growth in broth were disrupted with a 10 ml syringe prior to glycerol stock preparation. Isolates were subsequently stored long-term at -80°C.

2.1.3 Antibiotic sensitive and multidrug resistant stocks

All antibiotic sensitive and multidrug resistant bacteria were streaked onto Luria agar (LA) and incubated at 37°C for 48 hours. After this period, a single colony of each isolate was inoculated into 10 ml Luria broth (LB) broth and incubated at 37°C for an additional 24 hours with

constant agitation (200 rpm). Glycerol stocks (20%) were then prepared for each isolate and stored long-term at -80°C.

2.1.4 Solid-agar plate preparation

All solid growth media *(n =19)* was prepared as described in appendix 2.1 and autoclaved for 20 minutes at 121°C. Approximately 70 ml of each media type was aseptically poured into large 120 x 120 x 17 mm vented square plates (Greiner) and stored at 4°C prior to overlay.

2.1.5 Antibiotic producing candidates

Each actinomycete isolate was inoculated onto 19 different media types for a 12-day period to establish individual colonies and to facilitate secondary metabolite biosynthesis. All plates were subsequently overlaid with *Staphylococcus aureus* Newman and *Escherichia coli* DH10B (antibiotic sensitive strains) (Fig 2.1). Any isolates observed to inhibit the growth of the antibiotic sensitive panel were then screened against a panel of multidrug resistant pathogens (*Escherichia coli, Klebsiella pneumoniae, Staphylococcus aureus, Enterococcus faecium* and *Acinetobacter baumannii*) on the media that best facilitated antibiotic production (Fig 2.2). Full details on the strain designation of the multidrug resistant isolates utilised in this study is available in appendix 3. Actinomycete isolates that presented a zone of inhibition against the primary antibiotic sensitive and multidrug resistant panel were rescreened on individual plates, in triplicate, to confirm reproducibility.

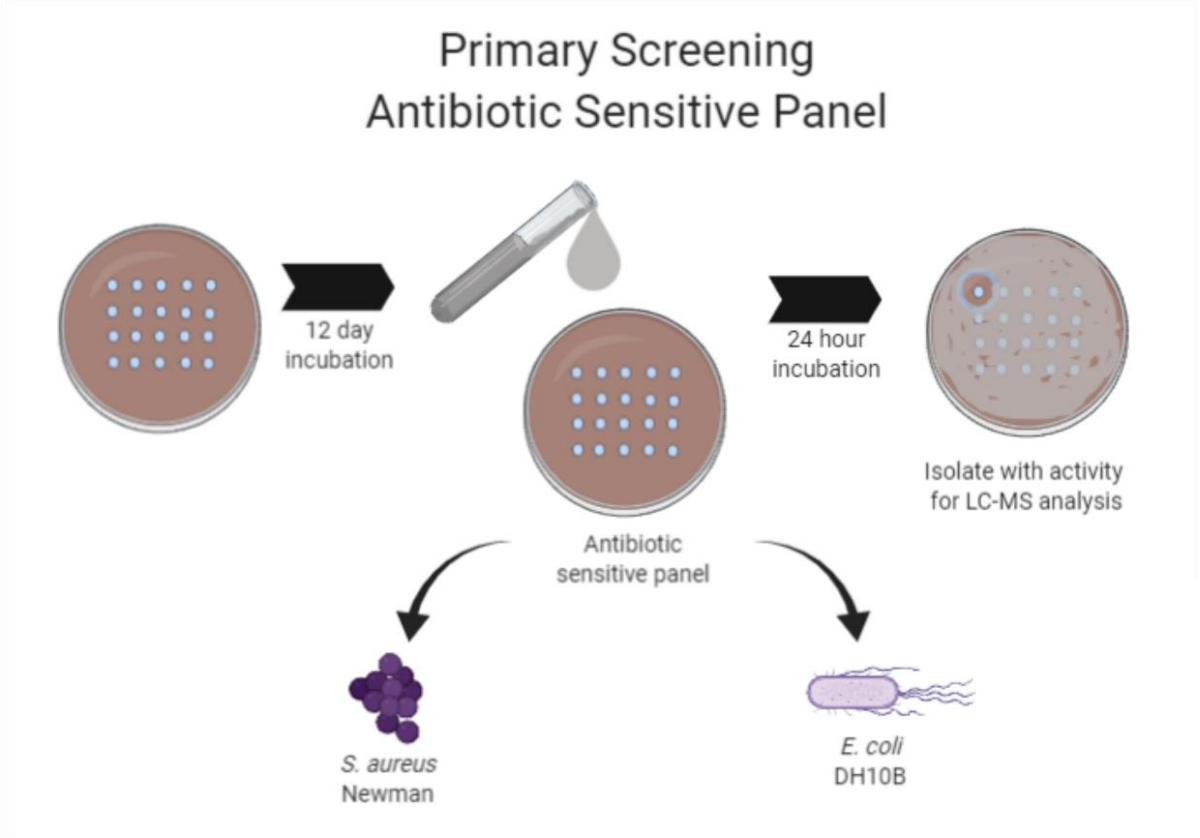


Fig 2.1 Primary antibiotic sensitive screening methodology. Each isolate was screened on 19 distinct media types and subsequently overlaid with an antibiotic sensitive organism *Escherichia coli* DH10B and *Staphylococcus aureus* Newman. Antibiotic producing isolates were prioritised and the media type that facilitated activity was noted. Those isolates were subsequently screened against a secondary multidrug resistant panel of bacterial pathogens. Image created using BioRender software (biorender.com).

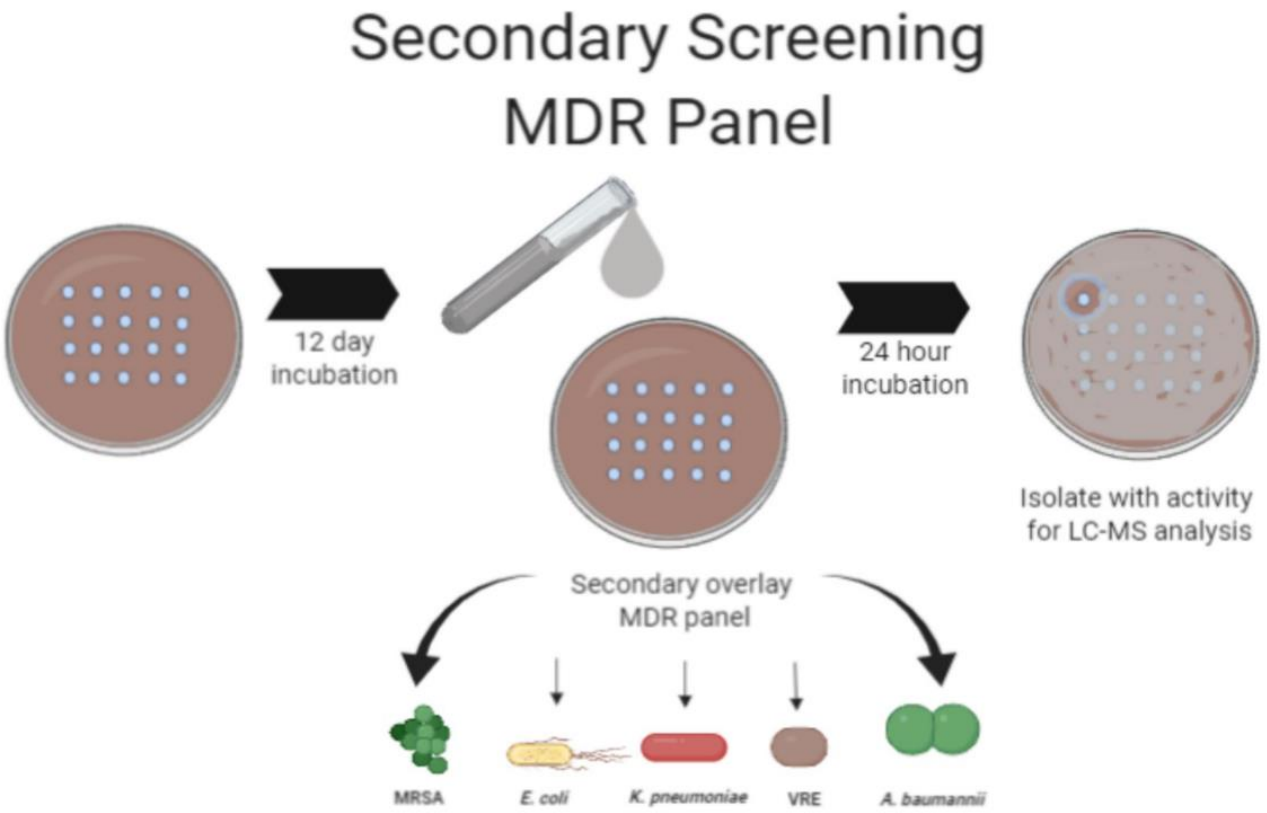


Fig 2.2 Secondary multidrug resistant screening methodology. Isolates that exhibited antibacterial activity against a primary antibiotic sensitive panel were subsequently screened against a secondary multidrug resistant panel of bacterial pathogens. Media selection for secondary screening was selected based upon the media type that best facilitated antibiotic production in the antibiotic sensitive screening phase of this project. Image created using BioRender software (biorender.com).

2.1.6 Overlay plate inoculation

A total volume of 200 µl from each actinomycete glycerol stock was aliquoted into a designated well within a 96-well plate (Sigma-Aldrich). An empty space was left between each well and the pattern of inoculation was staggered. Each 96-well plate accommodated a total of 45 actinomycete isolates and three positive controls. A 96-well replicator was used to aseptically inoculate all isolates and corresponding controls onto previously poured 120 x 120 x 17 mm vented square agar overlay plates. All plates were left for a period of ten minutes to facilitate the absorbance of the inoculum into the agar. Plates were then inverted and incubated at 28°C for a period of 12 days to facilitate the growth of each actinomycete isolate. After this period, plates were subjected to UV light for 20 minutes to kill all actinomycete isolates and halt further antibiotic production prior to overlay.

2.1.7 Antibiotic sensitive and multidrug resistant overlays

A total volume of 100 µl of glycerol stock from each antibiotic sensitive and multidrug resistant strain was inoculated into 50 ml of nutrient broth and incubated overnight at 37°C with constant agitation (200 rpm). Nutrient agar (0.6%) was autoclaved for 20 minutes at 121°C. A final volume of 5 ml of nutrient broth/ bacterial culture was added to the nutrient agar and the mixture shaken lightly. Approximately 40 ml of nutrient agar/bacterial culture was then poured over each of the corresponding media plates. Plates were left to solidify for ten minutes before being inverted and stored at 37°C overnight. All Plates were subsequently inspected and photographed to record any observable zones of inhibition indicative of antimicrobial activity.

2.2 Secondary metabolite extraction *Streptomyces* isolates

2.2.1 Plate inoculation

Each antibiotic producing candidate was streaked onto agar plates containing 20 ml of the growth media that facilitated antibiotic production. A total of 20 plates were streaked for each of the producing isolates ensuring there were enough bacteria to produce a confluent lawn of growth in each case. Five uninoculated plates of each media type were used as negative controls. All plates were incubated for a period of 12 days at 28°C.

2.2.2 Solid/liquid secondary metabolite extraction

All agar from the antibiotic producing strains and negative controls were cut into 1 cm square pieces and placed into a 1000 ml screw cap glass bottle. Ethyl acetate (500 ml) was added and each bottle was shaken vigorously for ten seconds and left to sit overnight at room temperature. Ethyl acetate from each bottle was poured into individual 1 L Erlenmeyer flasks and ~10 g of sodium sulphate (Alfa Aesar) was added to absorb residual water. The ethyl acetate/sodium sulphate solution was then filtered into a 1 L round-bottom flask and the ethyl acetate was evaporated in a rotary evaporator (Buchi). The remaining residue was then resuspended in 50 µl of HPLC grade methanol and stored at 4°C. LC-MS analysis was carried out by collaborators in the Bode laboratory, Goethe University, Frankfurt, Germany.

2.3 Whole genome sequencing isolate AUSMDU00041311

2.3.1 gDNA extraction for Nanopore sequencing

High molecular weight gDNA was extracted for long-read sequencing via “gentle” bacterial cell lysis and enzymatic digestion coupled to column purification using genomic tips (Qiagen).

BHI broth (10 ml) was inoculated with a single colony of antibiotic producing isolate AUSMDU00041311 and incubated for seven days with constant agitation (200 rpm). Approximately 3 ml of bacterial culture was harvested by centrifugation at 8000 x g for five

minutes. The bacterial cell pellet was resuspended in 1 ml of buffer B1 (Qiagen). A total of 20 µl of RNAse (10 mg/ml) and 20 µl of Lysozyme (100 mg/ml) was added. The mixture was pulse vortexed for 15 seconds and placed in a horizontal shaker at 37°C for 60 minutes (80 rpm). Proteinase K 200 µl (20 mg/ml) was added and the tube gently inverted prior to incubation in a horizontal shaker at 37°C for 60 minutes (80 rpm). Denaturing Buffer 350 µl (Qiagen) was added followed by gentle inversion of the tube prior to being placed in a horizontal shaker overnight at 50°C (80 rpm).

The lysate was vortexed for five seconds until it appeared clear and centrifuged at 3700 x g for ten minutes. A 20/G column (Qiagen) was inserted into a 15 ml falcon tube and washed with 1 ml of QBT buffer (Qiagen). The lysate was transferred to the 20/G column and allowed to flow into the falcon tube. The column was washed three times with QC buffer (Qiagen) and allowed to flow into the falcon tube. The falcon tube was discarded and the 20/G column was placed over a 2 ml tube (Eppendorf).

gDNA was eluted by two washes with 1 ml of QF buffer (Qiagen). The eluted gDNA was dried down to ~500 µl (SpeedVac) for two hours (run time) with two hours (heating) at 43°C. DNA was purified using AMPure XP beads (Beckman Coulter Life Sciences) as per manufacturer's guidelines and eluted with 50 µl of molecular grade water. gDNA was stored at 4°C prior to Nanopore sequencing.

2.3.2 Nanopore sequencing

Nanopore sequencing was peformed on a MinION™ DNA sequencer (Oxford Nanopore Technologies) by Dr. Koen Vandelannoote. The following long-read assembly approaches were utilised: Flye, Unicycler-hybrid, Canu and Wtdbg2. Trim adapters and barcodes missed by Guppy (Porechop). Long-reads reduced to a 100x subset containing the longest reads of highest quality (Filtlong). Reduced set assembly with one long-read assembler. Two rounds

of long-read polishing the draft assembly using the entire long-read dataset (Racon). Long-read polishing using the entire long-read dataset with an algorithm that understands the Guppy base-calling error model (Medaka). Three rounds of short-read polishing (Pilon), Trimming, circularising (Berokka) and Annotation (Prokka).

2.4 Gene cluster activation via promoter refactoring

Promotor refactoring was carried out on *Nocardia* isolate AUSMDU00041268 that could not be speciated following 16S rRNA sequencing within the Peter Doherty Institute, University of Melbourne.

2.4.1 Base plasmids

Promoter refactoring was carried out by utilisation of a two-plasmid workflow namely: pNH95 and pDLS. The former possessing the constitutive promoter *PermE*, an apramycin resistance cassette for positive selection of successful clones and an ColE1 origin of replication. The latter, a non-replicative homologous recombination vector equipped with a *sacB* gene for counter selection against single-crossover mutants.

2.4.2 Plasmid pTPS2460 assembly

A region of genomic DNA from isolate AUSMDU00041268 ~1.2 kb in size immediately upstream of the proposed promoter insertion site was PCR amplified. This fragment was inserted between restriction sites *EcoRI* and *KpnI* of plasmid pNH95. A second region of genomic DNA ~1.2 kb in size was PCR amplified immediately downstream of the proposed promoter insertion site. This fragment was inserted between restriction sites *PstI* and *HindIII* of plasmid pNH95. This resulted in the plasmid pTPS2460 with homology arms flanking the *PermE* promoter region (Fig 2.3).

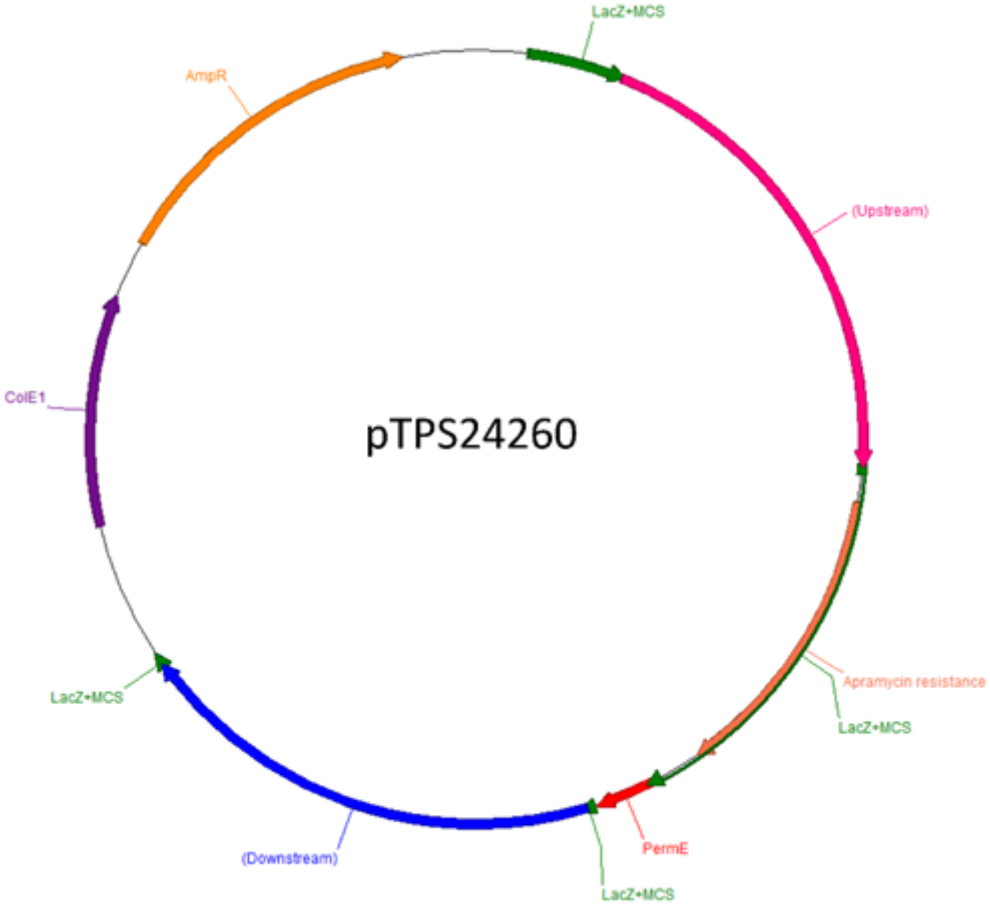


Fig 2.3. pTPS2460 plasmid assembly. Illustrating 1.2 kb fragment upstream (pink) and downstream (blue) of the proposed *PermE* promoter (red) insertion site. This plasmid possessed an apramycin resistance cassette (orange) and a ColE1 origin of replication for plasmid amplification (purple).

2.4.3 Plasmid pTPS10019 assembly

A 3.77 kb region (upstream gDNA fragment→Apra→*PermE*→downstream gDNA fragment) from plasmid pTPS2460 was digested with restriction endonuclease *SbfI* and subsequently ligated into plasmid pDLS. This resulted in the non-replicative, homologous recombination plasmid pTPS10019 that was utilised for transformation by electroporation in *Nocardia* isolate AUSMDU00041268 (Fig 2.4).

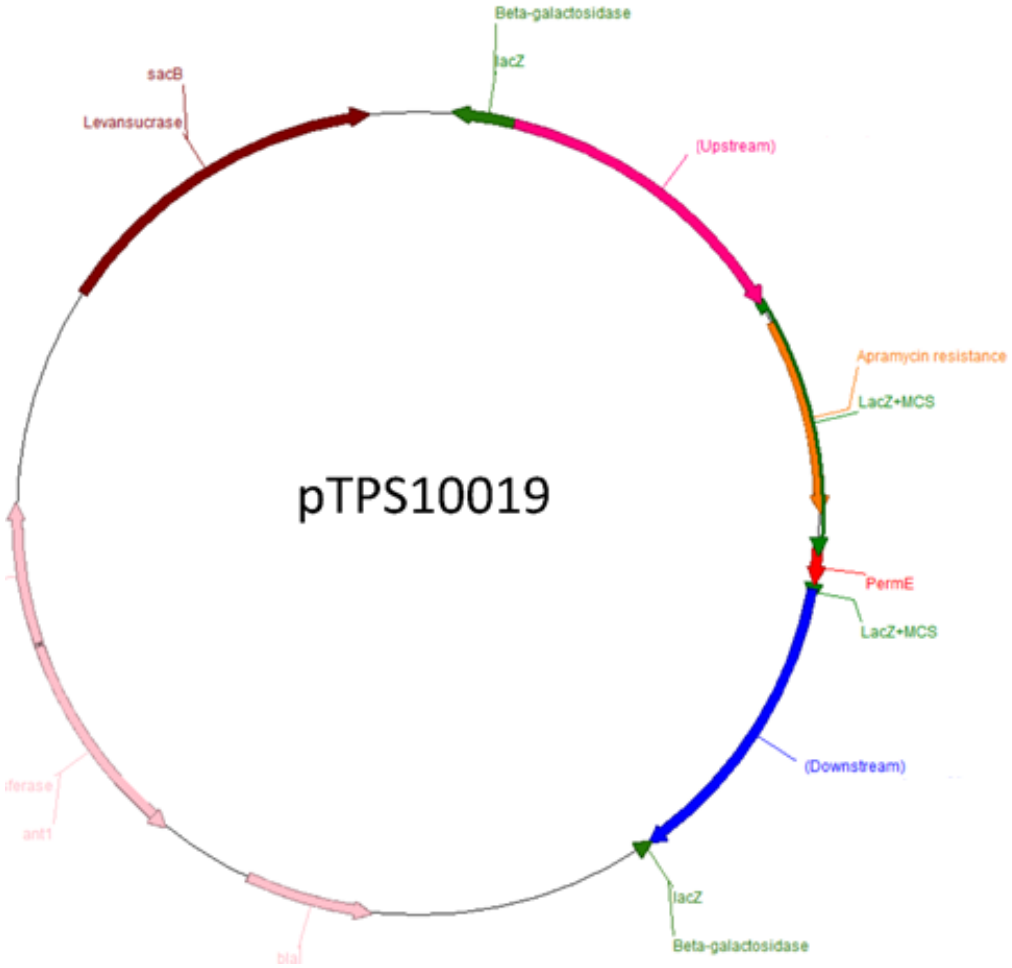


Fig 2.4 pTPS10019 plasmid assembly. Illustrating the 3.77 kb region (upstream gDNA fragment (pink), apramycin resistance (orange), *PermE* (red), downstream gDNA fragment (blue) assembly. This plasmid possessed a *sacB* cassette for counter selection against single-crossover mutants.

2.4.4 Polymerase chain reaction (PCR)

The synthetic oligonucleotides utilised within this study were obtained from Integrated DNA Technologies (Table 2.1). The PCR reaction included: 10 µl Phusion GC buffer, 5 µl forward and reverse primer (20 mM), 1 µl deoxynucleotide triphosphates (20 mM), 5 µl genomic DNA template, 0.4 µl Phusion high-fidelity DNA polymerase and 23.6 µl nuclease free water.

The final 50 µl PCR reaction mixture was run on a 2720 thermal cycler (Applied Biosystems) under the following settings: **Stage 1**: one cycle - 95°C (5 min), **Stage 2**: 35 cycles - 95°C (13 sec), 58°C (10 sec), 72°C (1 min), **Stage 3**: one cycle - 75°C (7 min).

Table 2.1 Synthetic oligonucleotides and associated plasmids.

Primer Name	**Sequence**	**Plasmid**
Cluster 12 (5'F)	aattGAATTCggcggttgtttgaggggcg	pNH95
Cluster 12 (5'R)	aattGGTACCttcgtcgtgacgaagcggtg	pNH95
Cluster12 (3'F)	aattCTGCAGatgaatccgcgatcggccac	pNH95
Cluster 12 (3'R)	aattAAGCTTtgtgcagttcgagttcgtcc	pNH95
Cluster 12 (5’F)	aattCCTGCAGGggcggttgtttgaggggcg	pDLS
Cluster 12 (3’R)	aattCCTGCAGGggcggttgtttgaggggcg	pDLS

2.4.5 Agarose gel electrophoresis

Successful PCR amplification of gDNA was confirmed by agarose gel electrophoresis. All samples were loaded into a 1% (w/v) agarose gel buffered in tris-acetate-EDTA (TAE), (0.4 M Tris, 50 mM EDTA) supplemented with 2.5 µl SYBR safe. Loading dye (2 µl) was added to each sample. Run parameters were set at 110V for 45 minutes. A 1 kb DNA ladder was used to determine fragment size.

2.4.6 Restriction Digest

Restriction digests were carried out in CutSmart buffer (5 µl). One unit of each high-fidelity restriction enzyme was added to the reaction and the final volume was brought to 50 µl with nuclease free water. Digests were incubated in a water bath for at 37°C for one hour. Successful digestion was confirmed by agarose gel electrophoresis prior to ligation reaction.

2.4.7 Ligation reaction

Ligation reactions were carried out with a 1:3 and 1:6 insert to vector molar ratio. Reaction mix was comprised of T4 ligation buffer (1 µl), T4 DNA ligase (one unit) and the final volume was brought to 10 µl with nuclease free water.

2.5 DNA extraction

2.5.1 DNA extraction from agarose gel

Following purification by gel electrophoresis, all PCR amplified cloning fragments were visualised by UV radiation and excised from the agarose gel using a scalpel blade. DNA/agarose mass was transferred to a 1.5 ml tube and DNA was extracted using a QIAquick gel extraction kit (Qiagen) in accordance with manufacturer's guidelines.

2.5.2 DNA concentration

The concentration of PCR amplified gDNA fragments was determined with a Qubit 2.0 fluorometer and Qubit dsDNA BR assay kit (Thermo Fisher Scientific). The assay is highly selective for double-stranded DNA over RNA and accurate for concentrations ranging from 100 pg/µl - 1,000 ng/µl.

2.6 Transformation *Escherichia coli*

2.6.1 Transformation by heat shock (ultra-competent Escherichia coli)

Preparation of optimally competent *Escherichia coli* was achieved by utilising the Inoue method (Im, 2011). Competent cells were stored at -80°C prior to transformation by heat shock.

Transforming DNA (up to 25 ng/50 µl of cells) was added to competent cells and left to sit on ice for 30 minutes. Transformation mix was placed in a heat block (42°C) for a period of 90 seconds and immediately transferred into an ice bath for a period of two minutes. Super optimal broth (800 µl) with catabolite repression (SOC) was added to each transformation mix and

incubated at 37°C for five minutes. The transformation mix was then placed in a 37°C shaking incubator to facilitate the expression of the apramycin resistance cassette. Transformation mix (200 µl) was plated onto super optimal agar (SOA) supplemented with 50 µg/ml apramycin (in triplicate for each molar ratio). Plates were then inverted and incubated overnight at 37°C.

2.6.2 Positive selection (Escherichia coli)

Following transformation by heat shock successful *E. coli* clones were streaked onto SOB agar supplemented with 50 µg/ml apramycin and incubated overnight at 37°C. This was done to ensure that the subsequent confirmatory PCR reaction did not amplify the ligation mix resulting in a false positive.

2.6.3 Colony PCR (Escherichia coli)

Individual colonies of positive clones were lifted from the agar using a sterile toothpick and added directly to the colony PCR reaction.

The PCR reaction included: 4 µl Phusion GC buffer, 2 µl forward and reverse primer (20 mM), 0.4 µl deoxynucleotide triphosphates (20 mM), 0.2 µl Phusion high-fidelity DNA polymerase and 11.4 µl nuclease free water.

The final 20 µl PCR reaction mixture was run on a 2720 thermal cycler (Applied Biosystems) under the following settings: **Stage 1**: one cycle - 95°C (5 min), **Stage 2**: 35 cycles - 95°C (13 sec), 58°C (10 sec), 72°C (1 min), **Stage 3**: one cycle - 75°C (7 min).

The PCR extension time varied depending on final amplicon size. One minute was added for every kilobase amplified. Successful PCR amplification of plasmid inserts were confirmed by agarose gel electrophoresis.

2.6.4 Plasmid extraction

PCR confirmed *E. coli* clones were lifted from culture plates using a sterile loop and added to 10 ml LB supplemented with apramycin (50 µg/ml) and incubated at 37°C overnight with constant agitation (200 rpm) prior to plasmid extraction. Plasmids were extracted using a plasmid mini-kit (Qiagen) as per manufacturer's instructions and stored at -20°C. Plasmid concentration was determined using a Qubit 2.0 fluorometer and Qubit dsDNA BR assay kit (Thermo Fisher Scientific).

2.6.5 Illumina sequencing

Correct assembly of plasmid pTPS10019 was confirmed by in-house Illumina sequencing at the Peter Doherty Institute, University of Melbourne. DNA libraries were created using Nextera XT DNA preparation kit (Illumina). Sequencing was performed on the NextSeq platform (Illumina) with 2 x 150 bp paired end chemistry. A sequencing depth of > 50x was targeted for each sample. Sequence reads were assembled with SPAdes (v 3.10.1) (Bankevich et al., 2012) and annotated with Prokka v 1.12 (Seemann, 2014).

2.7 Transformation Nocardia

2.7.1 Transformation by electroporation (Nocardia)

Nocardia wild type isolate AUSMDU00041268 glycerol sock (500 µl) comprised of ~1.21 $x10^8$ CFU/ml was added to 50 µl of BHI broth and incubated for 12 days at 28°C. Bacterial cells were then harvested and washed twice with 25 ml of ice-cold glycerol (10%). The suspension was transferred into a chilled electroporation cuvette (2 mm), 3 µg of plasmid pTPS10019 was added and mixed gently by pipetting. Each cuvette was pulsed (2225V/ 0150Ω/ 0025 µF). BHI broth (900 µl) was added to each cuvette and mixed by gentle pipetting. The transformation mixture was transferred to a 1.5 ml tube and incubated at 37°C without

agitation for two hours. The transformation mix was plated onto BHI agar supplemented with apramycin (50 µg/ml). Plates were inverted and incubated at 30°C for seven days.

2.7.2 Positive selection (double-crossover mutants)

All visible colonies were streaked onto BHI agar supplemented with apramycin (50 µg/ml) and sucrose (2%) to counter select for double-crossover mutants.

2.7.3 gDNA extraction ΩPermE mutants

A total of two AUSMDU00041268 Ω*PermE* mutants were generated from two independent transformations. gDNA was extracted from both mutants utilising a rapid phenol/chloroform extraction methodology as reported by Cheng and Jiang (Cheng and Jiang, 2006). Illumina sequencing by synthesis was utilised to confirm the generation of successful double-crossover mutant strains.

2.8 Secondary metabolite extraction isolate AUSMDU00041268

2.8.1 Culture preparation for HPLC analysis

Glycerol stocks were made for each Ω*PermE* mutant strain and stored at -80°C. A total of 100 µl of each Ω*PermE* mutant glycerol stock was added to 50 ml of ISP2 broth prior to HPLC analysis. Mutant cultures were supplemented with 50 µg/ml apramycin and incubated for a period of 12 days at 30°C (200 rpm). Corresponding AUSMDU00041268 wild type cultures were included.

2.8.2 Liquid/liquid secondary metabolite extraction

AUSMDU00041268 Ω*PermE* mutants and corresponding wild type cultures were poured into individual 200 ml screw cap glass bottles and 50 ml of ethyl acetate was added. The mixture was shaken vigorously and left to sit overnight at room temperature. Each culture/solvent mix was then poured into liquid/liquid extraction glassware and allowed to settle for five minutes.

The aqueous layer was removed and discarded. The higher density ethyl acetate layer was collected in a 250 ml glass beaker and ~10g sodium sulphate was added. The ethyl acetate/sodium sulphate solution was filtered into a 500 ml round-bottom flask and the ethyl acetate was evaporated in a rotary evaporator. The remaining residue was then resuspended in 50 µl of HPLC grade methanol and stored at 4°C prior to HPLC analysis.

2.8.3 Reversed-phase high-performance liquid chromatography

RP-HPLC was employed to determine whether the presence of *PermE* in both mutant strains activated the targeted non-ribosomal peptide synthetase locus. All samples were analysed using a 150 mm x 4.6 mm column. The mobile phase was comprised of solution A: Millipore ultra-pure water (0.1% (v/v) trifluoroacetic acid) and solution B: (0.1% trifluoroacetic acid (v/v) acetonitrile). Solvent flow rate was set at 0.5 ml/min, total run time for each sample was 25 minutes.

2.8.4 Bioactivity screen Ω*PermE mutants*

A bioactivity screen was conducted to assess whether promoter refactoring successfully activated the targeted non-ribosomal peptide synthetase locus. A total of 5 µl from each AUSMDU00041268 Ω*PermE* mutant and corresponding wild type glycerol stocks was pipetted onto Bennett's agar. Each plate was left for ten minutes to allow the inoculum to absorb into the agar. Plates were inverted and incubated at 28°C for 12 days. After this time plates were overlaid with five MDR pathogens: *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus* and *Enterococcus faecium* and *Acinetobacter baumannii*) and comparatively assessed for antibiotic activity.

2.9 Molecular networking

The secondary metabolome of ten pathogenic *Nocardia* isolates was interrogated via molecular networking. Each *Nocardia* isolate was cultured on five distinct media types and their

secondary metabolome was comparatively analysed. Culture media selection was based on the best performing media during the primary screening phase (Chapter 3) of this project. Glycerol stocks were created for each *Nocardia* isolate as described in materials and methods (2.1.2). All solid growth media *(n = 5)* was prepared as described in appendix (2.3). Information on individual *Nocardia* species screened is available in appendix (1).

2.9.1 Plate inoculation

A total of 800 µl from each media type was pipetted into a 96-well plate and allowed to set for ten minutes. This was followed by the addition of 5 µl from each *Nocardia* glycerol stock to the 96-well plate. Plates were incubated at 28°C for a period of 12 days.

2.9.2 Secondary metabolite extraction

A plug of agar from each well was placed into a 1.5 ml tube (Eppendorf). Each agar plug was broken up using a sterile toothpick. Ethyl acetate (1 ml) was added to each tube and vortexed for 30 seconds. Samples were left to sit overnight at room temperature. All samples were centrifuged at 17000 x g for five minutes to allow the solid agar to pellet. The solvent layer was removed and placed into a sterile 1.5 ml tube (Eppendorf). Solvent evaporation was carried out in a SpeedVac: two hours (run time), two hours (heating) at 43°C.

2.9.3 High-performance liquid chromatography-mass spectrometry

All *Nocardia* samples for metabolomic investigation were forwarded to the Bode laboratory, Frankfurt, Germany for LC-MS/MS analysis. UHPLC-ESI-HRMS/MS analyses was performed using an UltiMate 3000 Liquid chromatography system linked to a Bruker Impact II Q-TOF mass spectrometer. Runs were performed using a flow rate of 0.4 ml/min and gradient of MeCN 0.1% formic acid in H_2O (5:95% to 95:5% over 15 minutes).

2.9.4 Molecular network generation

Data were converted to the. mzXML format using DataAnalysis v4.3 (Bruker). All MS/MS data files were first converted from Agilent MassHunter data files (.d) to mzXML format and uploaded to the GNPS server (gnps.ucsd.edu). A comprehensive molecular network was generated using an integrated GNPS spectral clustering algorithm. A metadata file (.txt) comprised of sample attributes was generated (MDU number, media type, species ID, negative control, sample and biological replicate number) and uploaded to GNPS. Parameters included: precursor ion mass tolerance (0.02 Daltons), fragment ion mass tolerance (0.5 Daltons), cosine (0.5) and minimum number of matched peaks (4). These parameters were selected based upon a previous publication investigating secondary metabolism in *Streptomyces* sp. (Crusemann et al., 2017).

Nodes associated with negative controls were removed from the network analysis. All nodes with ions present that appeared in both the negative controls and samples were removed. All nodes with ions present that appeared in only one of the two biological replicates were removed. The spectral network was then uploaded to Cytoscape (3.7.1) and visualised with a class default setting prior to analysis.

Chapter 3: Empirical screening for novel antibiotics

3.1 Introduction

A key advantage when striving to discover novel antimicrobials is to utilise an original, previously unexamined source of bacteria. The Peter Doherty Institute, University of Melbourne possesses an extensive and previously unexamined collection of pathogenic actinomycetes isolated from patients within Victoria. *Streptomyces* are highly regarded and deserving of the title most prolific producers of bioactive secondary metabolites with antimicrobial activity (Berdy, 2005). However, utilisation of in-house whole genome sequencing, comparative genomics and bioinformatic resources at the Peter Doherty Institute indicates that the genus *Nocardia* harbours an impressive and in some cases, unique repertoire of secondary metabolite biosynthetic gene clusters (SMBGCs). These underexamined gene clusters represent a potentially untapped source for novel antimicrobials.

This chapter investigated the use of a high-throughput, empirical screening approach; utilising a soft-agar overlay assay to determine antimicrobial activity in 169 pathogenic actinomycetes, predominantly from the genus *Nocardia*. Isolates were screened on various growth media (*n = 19)* ranging from rich, well-defined to minimal, aiming to awaken cryptic antibiotic SMBGCs of unknown potential. All isolates were overlaid with an initial primary antibiotic sensitive panel consisting of *Escherichia coli* and *Staphylococcus aureus*. Isolates that presented an observable zone of inhibition indicative of the production of antimicrobial natural products were subsequently screened against a secondary MDR panel of prominent bacterial pathogens. Schematics of screening methodology are available in materials and methods (Fig 2.1 and 2.2).

This chapter aimed to address the current antibiotic crisis by finding new lead compounds for the development of novel antimicrobials. Specifically, those compounds with activity against some of the world's most prolific MDR bacterial pathogens. This was achieved by exploiting

an unexamined collection of actinomycetes held at the Peter Doherty Institute, University of Melbourne.

3.2 Chapter aims

i. High-throughput, empirical screening of 169 pathogenic actinomycetes for novel antimicrobial natural products
ii. To identify any bioactive secondary metabolites exhibiting activity against a panel of prominent MDR pathogens
iii. To utilise liquid chromatography-mass spectrometry and bioinformatic resources to determine whether such bioactive metabolites have potential for further development.

All pathogenic actinomycetes within this study were isolated from a cohort of patients within Victoria, Australia between the years 2002 and 2016. Due to the lack of access to patient records following treatment, patient pathology could not be wholly attributed to the presence of this group of bacteria. Therefore, infection and clinical presentation by pathogenic actinomycetes in such patients is probable but not definitive.

3.3 Antibiotic sensitive and multidrug resistant panel

3.3.1 Escherichia coli DH10B

Escherichia coli DH10B is genetically modified from a progenitor strain K-12. The DH10B strain is desirable for its high transformation efficiency and maintenance of large plasmids (Durfee et al., 2008). This genotype is also useful as a screening tool as it exhibits resistance to the aminoglycoside streptomycin only. Streptomycin resistance in DH10B further streamlined primary screening by eliminating any hits from *Streptomyces* isolates that produce this commonly encountered antibiotic. In addition, the high sensitivity of DH10B to other classes of antibiotics allowed for selection and prioritisation of antibiotic producing isolates for further investigation.

3.3.2 Staphylococcus aureus Newman

This strain was originally isolated from a male presenting with secondarily infected tubercular osteomyelitis (Duthie and Lorenz, 1952). This genotype is extensively utilised in animal models of staphylococcal disease due to its robust virulence properties and lack of any known antibiotic resistance determinants (Baba et al., 2008). This deficit of resistance genes makes this strain an ideal antibiotic sensitive candidate for the primary screening phase of this project.

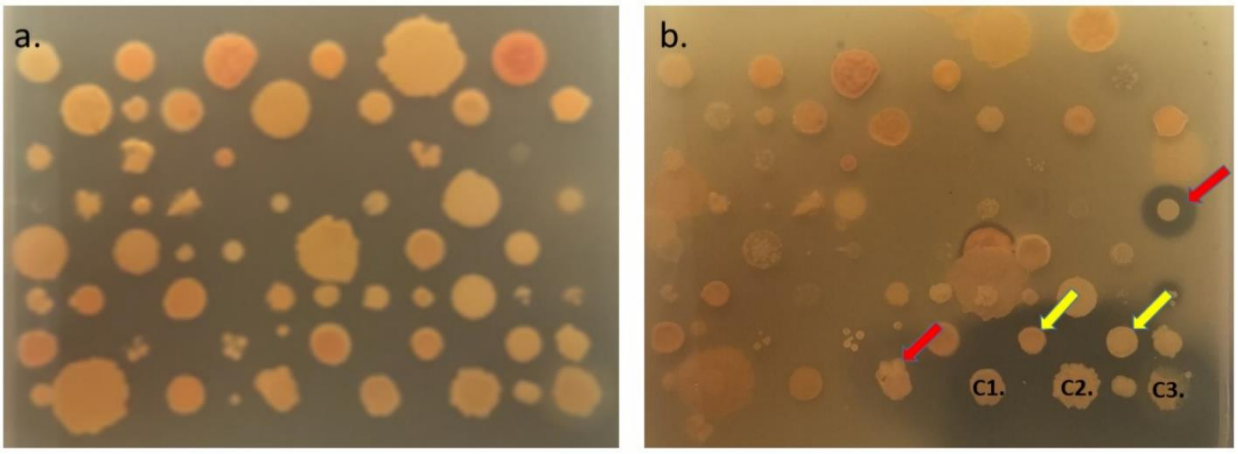


Fig 3.1 Results of a typical screen of 45 individual actinomycete isolates. Isolates were grown on asukamycin media after a 12-day incubation period. Images show before (LHS) and after (RHS) overlay with antibiotic sensitive *Staphylococcus aureus* Newman. Isolates exhibiting antimicrobial activity (red arrows), positive controls C1 (*Nocardia terpenica*), C2 (*Nocardia arthritidis*) and C3 (*Streptomyces* sp.) previously found to reliably produce antibiotics on a range of media types and under a variety of conditions. In cases where the zone of inhibition derived from the positive controls engulfs isolates in surrounding proximity isolates are recorded as potential antibiotic producers (yellow arrows).

3.3.3 Multidrug resistant panel

All isolates with antibacterial activity against a primary antibiotic sensitive panel were then screened against a panel of MDR pathogens. The five pathogens that were selected for the secondary screening phase of this chapter included: *Enterococcus faecium*, *Escherichia coli*, *Acinetobacter baumannii*, *Klebsiella pneumoniae* and *Staphylococcus aureus*. Isolates with activity against this multidrug resistant panel were then prioritised for further investigation.

3.4 Results

A total of 169 actinomycetes were screened for antimicrobial activity within this chapter. Of these isolates, 156 belonged to the genus *Nocardia* of which there were 13 different species. The remaining actinomycetes screened were from the following genera: *Streptomyces* (8), *Dietzia* (3), *Rhodococcus* (1) and *Auritidibacter* (1).

3.4.1 Antibiotic producing isolates

Of the 169 actinomycetes screened, 43% exhibited antimicrobial activity against *Staphylococcus aureus* Newman (herein referred to as "Newman"). Antimicrobial activity against *Escherichia coli* DH10B (herein referred to as "DH10B") was observed in 22% of all actinomycetes screened. Dual activity against both Newman and DH10B was observed in 20% of isolates screened (Fig 3.2).

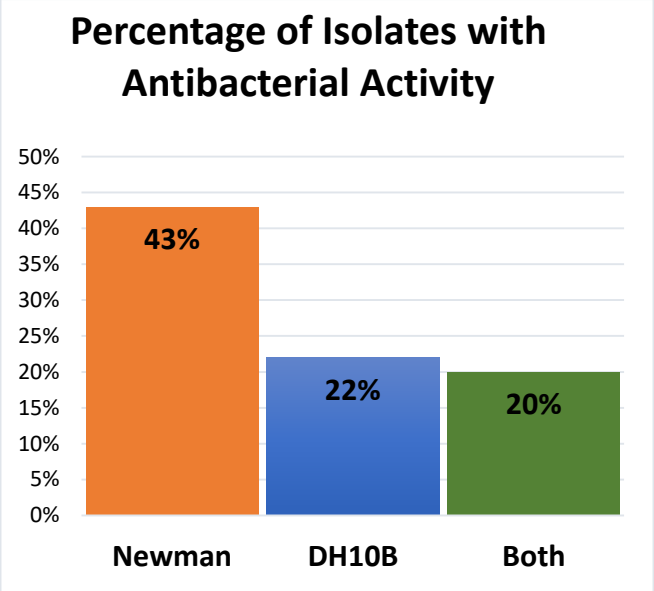


Fig 3.2 Antibiotic activity of all actinomycetes screened. Isolates with activity against *Staphylococcus aureus* Newman (orange), *Escherichia coli* DH10B (blue) and activity against both (green).

3.4.2 Media performance

The media that best facilitated antibiotic production overall in *Nocardia* against both Newman and DH10B was Bennett's (17 isolates) followed by, M1, MYM (16 isolates) and minimal media (14 isolates).

3.4.3 Nocardia isolates active against Staphylococcus aureus Newman

Primary screening of all 169 actinomycete isolates yielded 72 strains capable of inhibiting the growth of Newman (Fig 3.3). Of these 72 strains, 61 were from the genus *Nocardia* and 11 were from other genera screened. Of the 13 different *Nocardia* species screened, ten inhibited the growth of Newman on one or more media type. However, not all isolates from each *Nocardia* species produced a zone of inhibition. *N. brasiliensis* was one exception to this, where all 11 isolates screened inhibited the growth of Newman. *N. otitidiscaviarum*, *N. puris* and *N. vaccinia* had no observable activity against Newman on any of the 19 media types. Of the three *Nocardia* isolates that could not be speciated, two exhibited activity against Newman.

3.4.4 Production of anti-Newman antibiotics by media type

Of the 19 media types assessed, the production of antibiotics active against Newman appeared to occur most frequently in *Nocardia* species when they were grown on M1 agar (13 isolates). Other media that appeared to induce the production of antibiotics active against Newman from a range of *Nocardia* strains included: Bennett's (12 isolates), Bennett's modified and minimal media (11 isolates). In contrast, media that performed poorly regarding facilitation of antibiotic production in *Nocardia* included: AMM and V22 (two isolates), ATCC2, ISP2 and SM3 media (four isolates) (Fig 3.4).

Of the four other genera screened: *Streptomyces* (8), *Dietzia* (3), *Rhodococcus* (1) and *Auritidibacter* (1); all exhibited activity against Newman on a range of media types with the exception of one *Auritidibacter* and one *Dietzia* isolate. All eight *Streptomyces* isolates were active against Newman this was observed on 13 of the 19 media types. The media that best facilitated antibiotic production amongst this group included: MYM (8 isolates), M1 (7 isolates) and AMM (5 isolates).

3.4.5 Nocardia isolates active against Escherichia coli DH10B

Primary screening of 169 actinomycete isolates yielded 38 strains capable of inhibiting DH10B (Fig 3.3). Of these 38 strains, 30 were from the genus *Nocardia*. Of the 13 *Nocardia* species screened seven exhibited antibiotic activity against DH10B this included: *N. cyriacigeorgica, N. nova, N. veterana, N. paucivorans, N. farcinica, N. brasiliensis* and *N. abscessus.* Not all isolates of each individual *Nocardia* species inhibited the growth of DH10B. None of the three unspeciated *Nocardia* isolates exhibited activity against DH10B.

3.4.6 Production of anti-DH10B antibiotics by media type

Of the 19 media types assessed, the production of antibiotics active against DH10B appeared to occur most frequently in *Nocardia* sp. when they were grown on MYM agar (six isolates). Other media that appeared to induce antibiotic production from a range of strains included Bennett's (five isolates), AMM, rare 3 and V22 (four isolates) (Fig 3.4). In contrast, media that did not facilitate antibiotic production in *Nocardia* whatsoever when screened against DH10B included: ISP1, A3M, GYEA, Bennett's modified and asukamycin media.

All eight *Streptomyces* isolates exhibited antimicrobial activity against DH10B. This activity was observed on six of the 19 media types. No activity was observed from the three other genera (*Dietzia*, *Rhodococcus* and *Auritidibacter*). Antibiotic production was observed in *Streptomyces* on six out of 19 media types. The media that best facilitated antibiotic production in *Streptomyces* was AMM (four isolates).

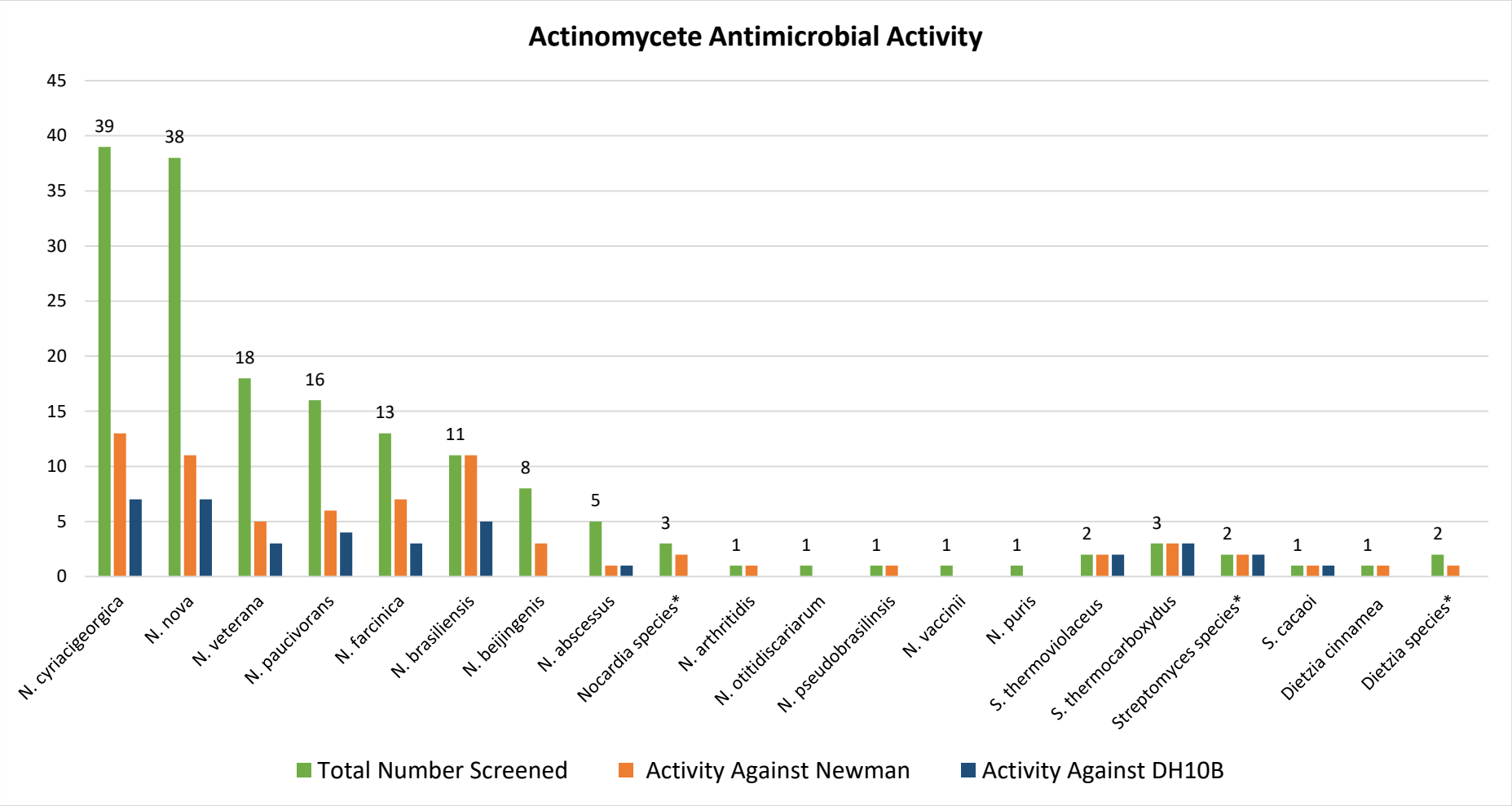


Fig 3.3 Breakdown of the number of antibiotic producing isolates. Individual species tested with activity against Newman and DH10B.

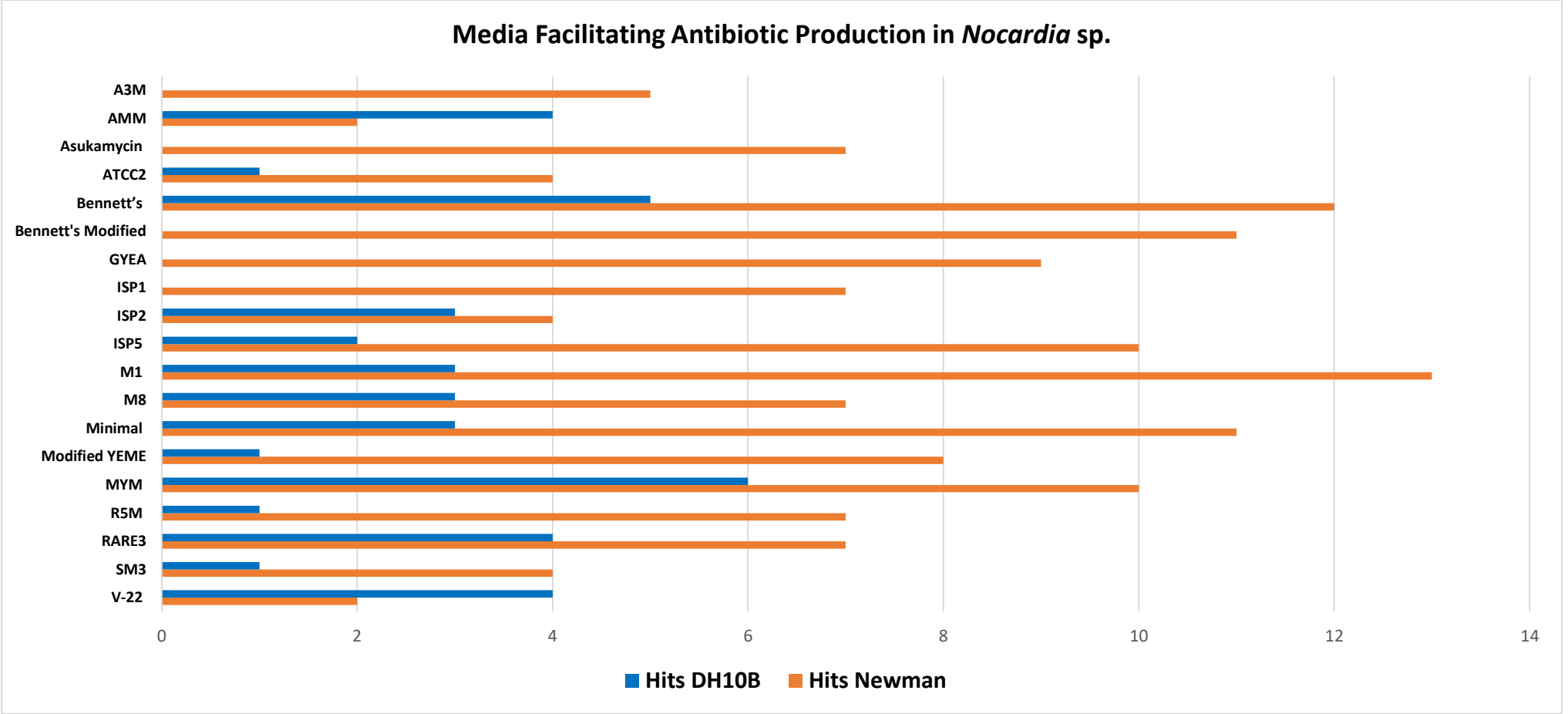


Fig 3.4 Media facilitating antibiotic production in *Nocardia*. Isolates that exhibited antibiotic production when screened against antibiotic sensitive strains: Newman (orange) and DH10B (blue).

3.5 Multidrug resistant screening

Any isolates that exhibited a zone of inhibition against a primary antibiotic sensitive panel were subsequently screened against a secondary multidrug resistant panel. A total of 66 *Nocardia* and 11 isolates from other genera qualified to be screened against the MDR panel. This selection was based upon the observation of zones of inhibition following overlay with Newman, DH10B (or both).

Of the 19 media types utilised in the primary screening phase 12 were selected for screening against the MDR panel. This selection was based upon the media which facilitated antibiotic production in the primary screening phase and included: ISP1, ISP2, rare 3, M1, minimal, Bennett's, asukamycin, A3M, MYM, modified YEME, GYEA and AMM.

Of the 66 *Nocardia* isolates screened, none exhibited antibiotic activity against the MDR panel. However, two *Streptomyces* isolates (AUSMDU00041311 and AUSMDU00041317) exhibited antimicrobial activity against two multidrug resistant strains.

3.5.1 Streptomyces isolate AUSMDU00041311

Following overlay with five multidrug resistant pathogens, isolate AUSMDU00041311 exhibited subtle antimicrobial activity when grown on M1 agar and challenged with multidrug resistant *Escherichia coli* (Fig 3.5). No antimicrobial activity was observed when this isolate was screened against the remaining four MDR pathogens. Phenotypic identification and subsequent 16S rRNA sequencing identified isolate AUSMDU00041311 as *Streptomyces cacaoi*.

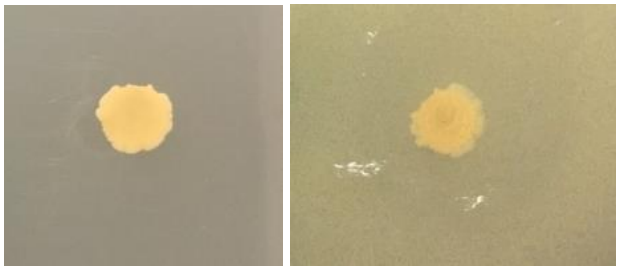


Fig 3.5 *Streptomyces cacaoi* isolate AUSMDU 00041311. Exhibiting subtle activity against MDR *Escherichia coli* on M1 media (RHS) with corresponding negative control (LHS).

3.5.2 Streptomyces isolate AUSMDU00041317

Following overlay against five multidrug resistant pathogens, Isolate AUSMDU00041317 exhibited antimicrobial activity against MDR *Acinetobacter baumannii*. This activity was observed on ISP2 and rare 3 media types. The diameter of the zone of inhibition was larger when isolate AUSMDU00041317 was grown on rare 3 media (Fig 3.6). No antimicrobial activity was observed when isolate AUSMDU00041317 was screened against the remaining four MDR pathogens. Phenotypic testing determined that isolate AUSMDU00041317 belongs to the genus *Streptomyces*. Subsequent 16S rRNA sequencing could not fully speciate, with a closest match to: *S. albus, S. rangoonensis* and *S. gibsonii* reported.

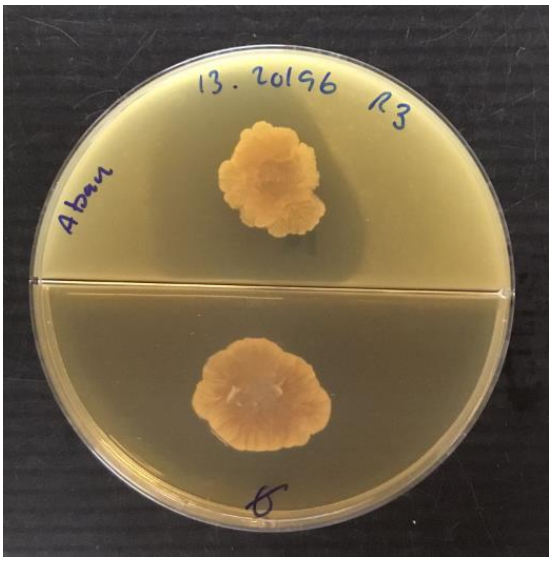


Fig 3.6 *Streptomyces* sp. isolate AUSMDU00041317. Exhibiting antimicrobial activity against MDR *Acinetobacter baumannii* on rare 3 media (top) and corresponding negative control (bottom).

3.6 Whole genome sequencing

3.6.1 Streptomyces isolate AUSMDU00041311

Nanopore whole genome sequencing of *Streptomyces cacaoi* isolate AUSMDU00041311 was carried out to gain insight into the secondary metabolite biosynthetic potential of the strain. antiSMASH genome annotation software was utilised to identify the probable biosynthetic origin of the antibiotic observed to inhibit the growth of MDR *Escherichia coli* in this chapter.

3.6.2 Streptomyces isolate AUSMDU00041317

Nanopore whole genome sequencing and antiSMASH analysis was not carried out for *Streptomyces* isolate AUSMDU00041317 which exhibited antimicrobial activity against multidrug resistant *Acinetobacter baumannii*. This was due to time limitations within this project.

3.6.3 antiSMASH analysis isolate AUSMDU00041311

Following whole genome Nanopore sequencing, isolate AUSMDU00041311 was bioinformatically investigated via the Antibiotics and Secondary Metabolite Analysis Shell (antiSMASH). This isolate possesses a total of 30 SMBGCs which included: hybrid NRPS/PKS (8), NRPS (9), terpene (5), siderophores (2), ectoine (2), lanthipeptide (1), nucleoside (1), bacteriocin (1) and T1-PKS (1) (Fig 3.7).

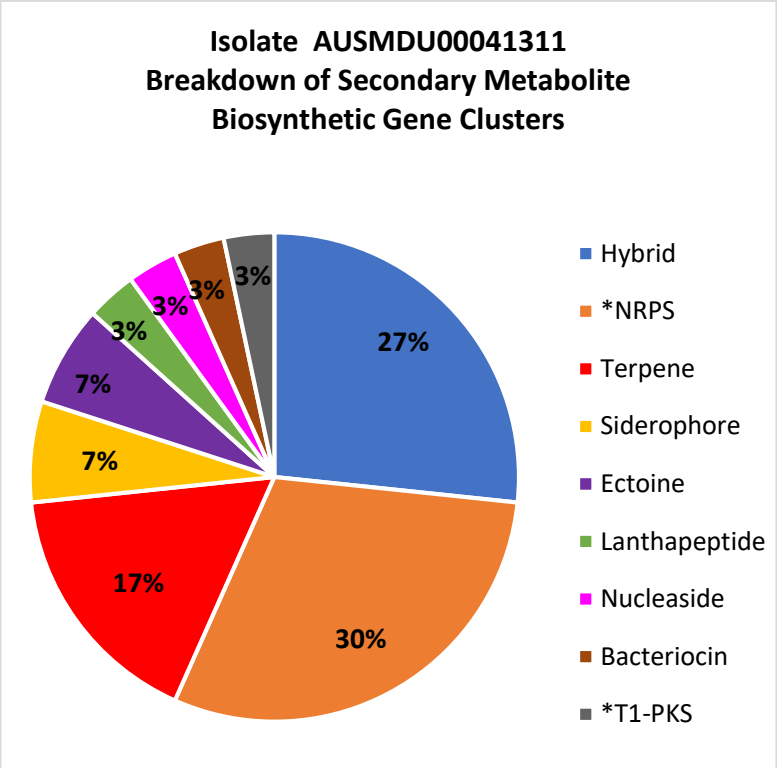


Fig 3.7 Secondary metabolite biosynthetic gene cluster breakdown isolate AUSMDU00041311. *NRPS (non-ribosomal peptide synthetase) *T1PKS (Type I polyketide synthase).

Of the 30 SMBGCs present end product identification of three clusters was obtained (100% match) this included: cluster 19 (desferrioxamine), cluster 10 (ectoine) and cluster 26 (actinonin). Four gene clusters had an unknown end product (0% match) and included: cluster 2 (bacteriocin), cluster 11 (lanthipeptide), cluster 20 (bacteriocin) and cluster 24 (betalactone) (Table 3.2).

Table 3.2 AntiSMASH interrogation of isolate AUSMDU00041311. Listing all 30 SMBGCs including cluster number, gene cluster type and similarity match to known compounds (%). The three gene clusters with 100% end product identification (desferrioxamine, ectoine and actinonin) are highlighted in green and the nucleoside puromycin with 87% identity shown in blue. The T1PKS/NRPS-like, lanthipeptide, bacteriocin and betalactone biosynthetic loci with unknown end products are highlighted in red.

Cluster Number	Gene Cluster Type	Closest Match	Similarity
1	NRPS, betalactone, T3PKS, T1PKS	totopensamide	64%
2	**T1PKS/NRPS-like**	**unknown**	**unknown**
3	T1PKS, NRPS-like, CDPS	candicidin	76%
4	NRPS	vazabitide A	8%
5	NRPS	surugamide A surugamide D	9%
6	terpene	carotenoid	54%
7	CDPS	streptomycin	2%
8	NRPS	PM100117 PM100118	52%
9	**siderophore**	**desferrioxamine**	**100%**
10	**ectoine**	**ectoine**	**100%**
11	**lanthipeptide**	**unknown**	**unknown**
12	terpene	daptomycin	3%
13	other, NRPS	svaricin	25%
14	**nucleoside**	**puromycin**	**87%**
15	T2PKS, NRPS	pristinamycin	20%

16	ectoine	kosinostatin	13%
17	NRPS-like	calcium-dependent antibiotic	12%
18	siderophore	ficellomycin	3%
19	NRPS	herboxidiene	2%
20	**bacteriocin**	**unknown**	**unknown**
21	NRPS, T1PKS	kanamycin	61%
22	terpene	abyssomicins M–X	9%
23	NRPS	surugamide A surugamide D	57%
24	**betalactone**	**unknown**	**unknown**
25	terpene	hopene	69%
26	**NRPS**	**actinonin**	**100 %**
27	CDPS	laspartomycin	4%
28	T1PKS	herbimycin	6%
29	T1PKS, NRPS	coelibactin	45%
30	NRPS, transAT-PKS	phthoxazolin	18%

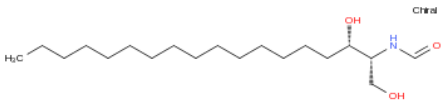
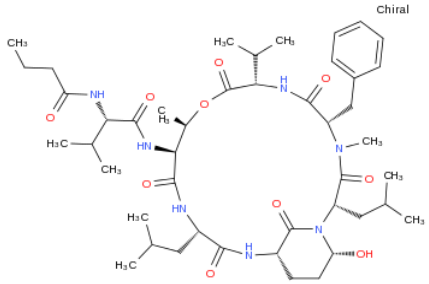
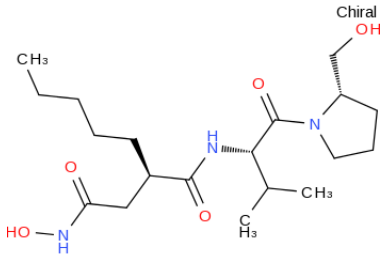
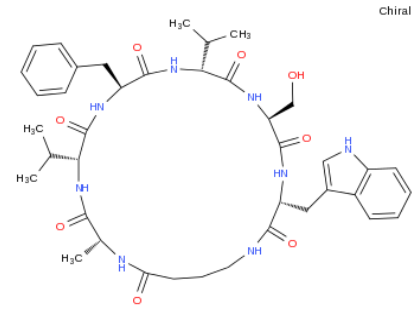
3.6.4 DEREPLICATOR+ analysis

To establish whether the antibiotics produced by *Streptomyces cacaoi* AUSMDU00041311 and *Streptomyces* sp. AUSMDU00041317 were novel a solid/liquid secondary metabolite extraction was carried out on both isolates. The resulting extracts were forwarded to the Bode laboratory, Goethe University, Frankfurt, Germany for subsequent LC-MS/MS analysis. Spectral data files were uploaded to the Global Natural Products Social Molecular Networking (GNPS) platform. The DEREPLICATOR+ algorithm is an extension of the GNPS platform and provides natural product researchers with a dereplication strategy. (Mohimani et al., 2018).

3.6.5 DEREPLICATOR+ analysis isolate AUSMDU00041311

The MS/MS data files associated with isolate *AUSMDU00041311* grown on M1 media which facilitated antibiotic activity against multidrug resistant *Escherichia coli* was uploaded to the GNPS server. The DEREPLICATOR+ algorithm was utilised to determine if there were any previously identified secondary metabolites within the media extract. This yielded a total of four hits of known metabolites (Table 3.3).

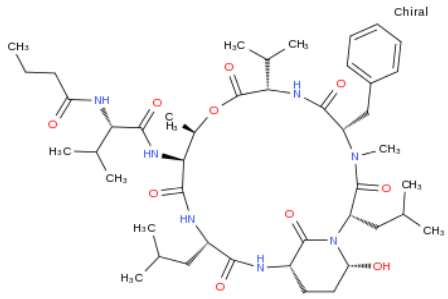
Table 3.3 DEREPLICATOR+ analysis of isolate AUSMDU00041311. The secondary metabolites identified in crude extracts included: cedefingol, tasipeptin A, actinonin and unguisin C.

DEREPLICATOR+ Predicted Core Structure	**Compound ID**	**MetMass**
	cedefingol	329.29
	tasipeptin_A	869.53
	actinonin	385.26
	unguisin C	774.41

3.6.6 DEREPLICATOR+ analysis of isolate AUSMDU00041317

The MS/MS data files associated with *Streptomyces* isolate AUSMDU00041317 grown on rare 3 media which facilitated antibiotic activity against multidrug resistant *Acinetobacter baumannii* was uploaded to the GNPS server. The DEREPLICATOR+ algorithm was utilised to determine if there were any previously identified secondary metabolites within the media extract. This yielded a single hit for tasipeptin A (Table 3.4).

Table 3.4 DEREPLICATOR+ analysis of isolate AUSMDU00041317. A single hit for the known secondary metabolite tasipeptin A was obtained.

DEREPLICATOR+ Proposed structure	**Compound ID**	**MetMass**
	tasipeptin A	869.53

3.7 Discussion

The traditional empirical screening carried out within this chapter was inspired by the one strain many compounds (OSMAC) approach demonstrated by Bode and colleagues. The varying parameters selected were media composition and challenge organism (Newman and DH10B). Despite being the predominant genus screened, none of the *Nocardia* isolates were capable of inhibiting the growth of any of the MDR panel. However, when activity was examined after challenge with Newman, *Nocardia* showed great potential as a source of new drug scaffolds for further investigation; with 61 out of 156 isolates readily expressing observable

antimicrobial metabolites. When *Nocardia* isolates were screened against DH10B a total of 30 isolates exhibited inhibitory activity. This decreased number of hits in response to Gram-negative challenge is not unexpected and likely due to outer membrane permeability issues and numerous drug efflux mechanisms that are ubiquitous in Gram-negative bacteria. Considering the urgency for new drug leads with activity against Gram-negatives, it is recommended that the narrow spectrum antibiotics observed with efficacy against DH10B are prioritised over those with activity against Newman alone. Candidates exhibiting broad spectrum activity against both Newman and DH10B should also be a prioritised for further investigation.

Of the 13 *Nocardia* species screened only seven were present in adequate numbers to assess species performance. These included: *N. cyriacigeorgica* (39 isolates), *N. nova* (38 isolates), *N. veterana* (18 isolates), *N. paucivorans* (16 isolates) *N. farcinica* (13 isolates) and *N. brasiliensis* (8 isolates). The remaining *Nocardia* species were not present in significant enough numbers to evaluate secondary metabolite production. All *N. brasiliensis* isolates exhibited inhibitory activity against Newman. This was not observed in any of the other *Nocardia* species screened indicating that that *N. brasiliensis* is an excellent candidate for further investigation. This observation echoes findings by The Gonoi laboratory in which they comparatively assessed NRPS and PKS biosynthetic loci amongst seven *Nocardia* strains: *N. asteroides* NBRC 15531^{T}, *N. otitidiscaviarum* IFM 11049, *N. farcinica* IFM 10152, *N. cyriacigeorgica* GUH-2, *N. brasiliensis* HUJEG-1, *N. brasiliensis* NBRC 14402^{T} and *N. brasiliensis* IFM 10847). The team showed that *N. brasiliensis* isolates possessed the highest quantity of NRPS and PKS biosynthetic gene clusters as well as showing a correlation between gene cluster quantity and a larger genome size (Komaki et al., 2014)

To add an additional dynamic to this chapter, media type was evaluated for its ability to illicit secondary metabolite biosynthesis in different species. No individual species exhibited antibiotic activity on all 19 media types which suggests that there is no such thing as an “ideal

media" when screening actinomycetes. However, some media types facilitated antibiotic production better than others. Of the 19 media types assessed complex media elicited SMBGC induction to a greater extent in comparison to minimal media parallels. This highlights the importance of a rich media base; which appears to act as a physiological trigger for secondary metabolite biosynthesis in *Nocardia*.

The best performing media overall was Bennett's. This media type was assessed for any component that could be attributed to increasing the likelihood of induction of antibiotic biosynthetic genes. It was noted that this media type contained a relatively high amount of glucose (10 g/L). Therefore, glucose concentration was assessed within all media types to determine if its presence aids the induction of antibiotic biosynthetic genes. This does not appear to be the case, as other media types contained identical concentrations of glucose (asukamycin, GYEA and SM3) and performed relatively poorly in comparison to Bennett's. It is recommended that Bennett's media is the best choice for facilitating antibiotic biosynthesis in *Nocardia* species against Gram-positive and Gram-negative challenge and it would be advisable to utilise this media in future screening efforts. However, it should be noted that media which did not perform particularly well overall should not be completely disregarded. This was demonstrated with isolate AUSMDU00018987 (*N. brasiliensis)* which produced a significant (4cm) zone of inhibition against Newman. Interestingly, this zone was only observed when this isolate was grown on three of the 19 media types (A3M, asukamycin and rare 3) none of which ranked within the top performing media overall. This suggests that "badly performing" media has the potential to significantly activate cryptic gene clusters in certain isolates.

These observations suggest that media suitability is highly variable depending on genus and species type. Additionally, secondary metabolite biosynthetic gene clusters of which are

numerous in actinomycetes appear to have varying requirements for induction that differ from one cluster to the next adding an additional layer of complexity to gene activation.

Of the other genera tested, *Streptomyces* performed well, with every isolate screened producing a zone of inhibition against both Newman and DH10B on a variety of media types. It is unsurprising that the two candidates which exhibited activity against MDR *Escherichia coli* and *Acinetobacter baumannii* were from this genus. This underscores why natural product researchers remain heavily focused upon the *Streptomyces* as a source of novel antimicrobials.

To determine whether the antimicrobial metabolites observed were in fact novel, DEREPLICATOR+ analysis was carried out on both producing *Streptomyces* isolates. This yielded a total of four hits for known natural products. Encouragingly, none of these natural product hits have been reported to inhibit the growth of multidrug resistant Gram-negative pathogens as was observed in this chapter. This suggests that both antibiotics could be novel and warrant further investigation.

An additional antiSMASH analysis was performed for *Streptomyces cacaoi* isolate AUSMDU00041311 which exhibited activity against multidrug resistant *Escherichia coli*. This provided further insight into the significant biosynthetic potential of this strain. Isolate AUSMDU00041311 is endowed with 30 SMBGCs of these, only three were identified with 100% certainty and included: desferrioxamine, ectoine and actinonin. Of the remaining 27 clusters, four encoded unknown end products. The remaining gene clusters are low identity matches. It should be noted that that antiSMASH similarity scores, in some instances, can be extremely low; as the algorithm is designed to identify any homologous sequences against an entire database of biosynthetic gene clusters. Low similarity scores indicate minor similarities in biosynthetic gene cluster sequences. Generally, a threshold of above 85% indicates an end product that would have similar bioactivity to the predicted named metabolite.

When DEREPLICATOR+ analysis was complemented by antiSMASH analysis for isolate AUSMDU00041311 only one secondary metabolite was present in both cases, the non-ribosomal peptide actinonin. This antibacterial peptide is a potent peptide deformylase inhibitor (essential for peptide maturation in bacteria) and is currently in phase II clinical trials. Actinonin exhibits activity against Gram-positive bacteria indicating that this antimicrobial is not responsible for the anti-Gram-negative activity observed in this chapter. The biosynthetic gene cluster responsible for actinonin production in *Streptomyces* sp. ATCC 14903 was recently elucidated by gene deletion and heterologous pathway expression (Wolf et al., 2018). Here we report the production of actinonin in media extracts following LC-MS/MS analysis which was complemented by antiSMASH interrogation in *Streptoyces cacaoi* AUSMDU00041311 which to our knowledge has not been previously reported.

Streptomyces cacaoi is known to produce polyoxin which was the first discovered broad-spectrum antifungal nucleoside antibiotic (Suzuki et al., 1965). This phytopathogenic agent is commonly utilised as a non-toxic agricultural fungicide. Lysocellin is a polyether antibiotic commonly isolated from *Streptomyces cacaoi* (Ebata et al., 1975). This exhibits activity against Gram-positive bacteria including drug resistant *Staphylococcus aureus* (resistance to Streptomycin, erythromycin, chloramphenicol and penicillin). Interestingly, neither polyoxin or lysocellin were identified in media extracts or following bioinformatic investigation of *Streptomyces cacaoi* isolate AUSMDU00041311 suggesting that the SMBCG endowment of a species can vary greatly. To our knowledge there has not been any published literature regarding the production of antimicrobials by *Streptomyces cacaoi* with activity against multidrug resistant *Escherichia coli*. Therefore, it is recommended that this antimicrobial should be prioritised for further investigation and molecular structure elucidation.

Chapter 4: Biosynthetic gene cluster activation via promoter refactoring

4.1 Introduction

The advent of next generation sequencing has vastly increased the acquisition rate of genomic data. This coupled to ever decreasing whole genome sequencing costs permits great insight into the full biosynthetic potential of "talented" secondary metabolite producers such as the actinomycetes. Traditional bioactivity screens are useful for identifying visually observable antimicrobial zones of inhibition from readily expressed secondary metabolites. However, downfalls with this approach include the re-discovery of known natural products and the inability to visualise or identify zones of inhibition from antimicrobials expressed at extremely low levels. In addition, many of the secondary metabolite biosynthetic gene clusters (SMBGCs) responsible for antibiotic production in bacteria are transcriptionally silent (cryptic). Utilisation of genome mining and bioinformatics to identify previously uncharacterised SMBGCs in bacteria coupled to genetic engineering to activate cryptic gene clusters represents a more contemporary approach to natural product discovery.

In-house whole genome sequencing data from a collection of 100 pathogenic *Nocardia* isolates held at the Peter Doherty Institute, University of Melbourne were uploaded to the Antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) bioinformatic platform to identify unique SMBGCs for further investigation. antiSMASH is a software-pipeline that facilitates identification, annotation and cross-genome analysis of secondary metabolite biosynthetic loci from all known classes of secondary metabolites (Medema et al., 2011). This comparative genomic approach facilitated the ranking of biosynthetic gene clusters amongst *Nocardia* isolates in order of uniqueness. Unique clusters often yield unique metabolites therefore utilising this pipeline increased the likelihood of discovering novel antimicrobial natural products for future development.

This workflow identified a *Nocardia* isolate (AUSMDU00041268) that possessed a SMBGC that appears to be unique amongst the 100 *Nocardia* genomes assessed. Furthermore, prior to the commencement of this project, isolate AUSMDU00041268 was screened for antibiotic production on a range of media types and no antimicrobial activity was observed. Isolate AUSMDU00041268 was therefore selected for promoter refactoring as the targeted biosynthetic gene cluster is likely transcriptionally silent and could encode a novel antimicrobial for further investigation.

The constitutive promoter P*ermE* drives expression of erythromycin biosynthetic genes in *Saccharopolyspora erythraea.* P*ermE* refactoring has been successfully utilised to activate and enhance secondary metabolite biosynthesis in a range of Actinobacteria (Li et al., 2009a, Luo et al., 2015). Considering this, *PermE* was selected for the promoter refactoring component of this chapter. Secondary metabolite extracts from mutant and wild type strains were subjected to HPLC analysis and bioactivity assays to evaluate the success of promoter refactoring and gene cluster activation.

This chapter aims to address the stark disconnect between the biosynthetic potential of *Nocardia* sp. and the resulting expression of the encoded chemistry through utilisation of genomics, genetic engineering, bioinformatics and bioactivity led assays.

4.2 Chapter aims

I. Activation of a transcriptionally silent secondary metabolite biosynthetic gene cluster via constitutive promoter refactoring
II. To examine the activity of the resulting mutant strains via reversed-phase high-performance liquid chromatography
III. To screen mutants for antibiotic activity against a range of multidrug resistant pathogens

4.3 Results

4.3.1 Nocardia isolate AUSMDU00041268

Phenotypic identification and subsequent 16S rRNA sequencing was carried out on isolate AUSMDU00041268. The results indicated that this isolate belongs to the genus *Nocardia* but could not be speciated. The genome sequence of this isolate was compared to that of 100 other *Nocardia* previously sequenced at the Peter Doherty Institute, University of Melbourne. antiSMASH analysis indicated that isolate AUSMDU00041268 possesses a total of 20 SMBGCs. A 24 kb non-ribosomal peptide synthetase locus was identified which appears to be unique to this isolate (Fig 4.1).

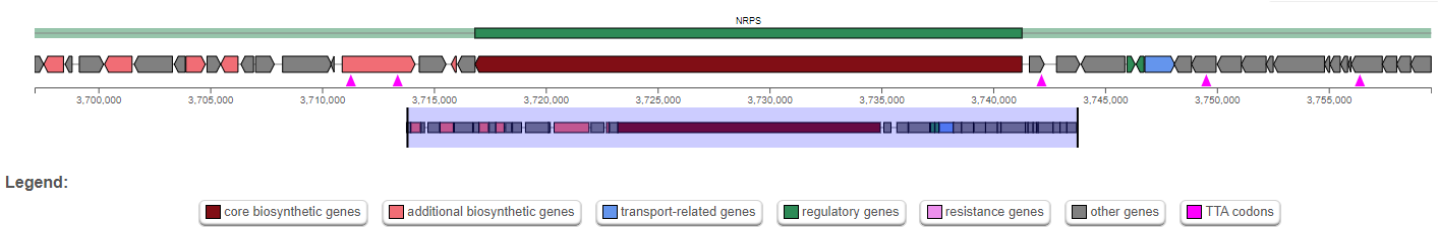


Fig 4.1 antiSMASH depiction of the 24 kb non-ribosomal peptide locus selected for promoter refactoring. Gene sequences are colour coded per the functionality of their predicted protein sequences.

4.3.2 Promoter refactoring

Promoter refactoring was carried out by utilisation of a two-plasmid workflow namely: pNH95 and pDLS. The former possesses the constitutive promoter *PermE*, an apramycin resistance cassette for positive selection of successful clones and a ColE1 origin of replication for plasmid amplification. A region of gDNA ~ 1.2 kb upstream (5') and downstream (3') of the proposed promoter insertion site was cloned into vector pNH95 yielding construct pTPS2460. The 5'→Apra→*PermE*→3' region of construct pTPS2460 was excised and subsequently cloned into the non-replicative homologous recombination vector pDLS yielding construct pTPS10019. This construct possessed a *sacB* gene which encodes the secreted enzyme levansucrase involved in the hydrolysis of sucrose and the synthesis of levans (high-molecular weight fructose polymers). In the presence of sucrose, *sacB* is toxic to bacteria through a

mechanism that is not well understood. Therefore, *sacB* was utilised as a counter selection marker against single-crossover mutants following growth on sucrose supplemented agar. pTPS10019 was introduced to *Nocardia* sp. AUSMDU00041268 via electroporation. Positive clones were selected for on BHI agar supplemented with apramycin 50 µg/ml and 2% sucrose.

Correct assembly of homologous recombination vector pTPS10019 was determined by Illumina sequencing prior to transformation by electroporation. Post-transformation, successful double-crossover mutants were also confirmed by Illumina sequencing. Two independent transformations were carried out yielding two Ω*PermE* mutant strains for further investigation hereby referred to as: Ω*PermE1* and Ω*PermE2*.

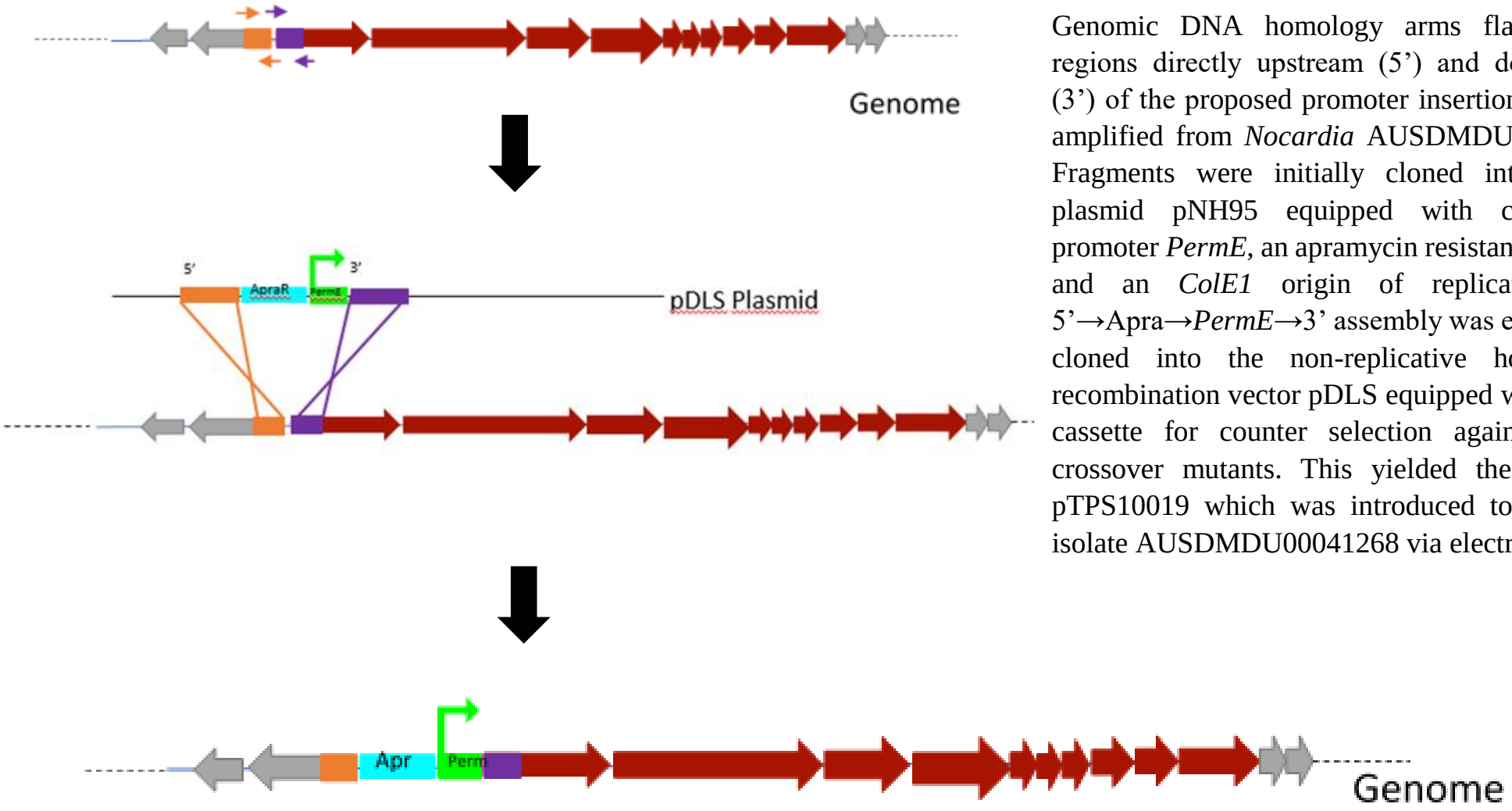


Fig 4.2 Workflow to generate Ω*PermE* mutant.

Genomic DNA homology arms flanking the regions directly upstream (5') and downstream (3') of the proposed promoter insertion site were amplified from *Nocardia* AUSDMDU00041268. Fragments were initially cloned into a base plasmid pNH95 equipped with constitutive promoter *PermE*, an apramycin resistance cassette and an *ColE1* origin of replication. The 5'→Apra→*PermE*→3' assembly was excised and cloned into the non-replicative homologous recombination vector pDLS equipped with a *sacB* cassette for counter selection against single-crossover mutants. This yielded the construct pTPS10019 which was introduced to *Nocardia* isolate AUSDMDU00041268 via electroporation.

4.3.3 RP-HPLC analysis

To ascertain whether promoter refactoring activated the targeted NRPS biosynthetic gene cluster, *ΩPermE1, ΩPermE2* and wild type strains were cultured for a period of 12 days on ISP2 media. HPLC analysis of ethyl acetate extracts derived from mutant and wild type strains was conducted. HPLC trace analysis showed consistency in the number of peaks present in both mutant and wild type strains. This indicated that the targeted biosynthetic gene cluster did not appear to be activated in either of the mutant strains (Fig 4.3).

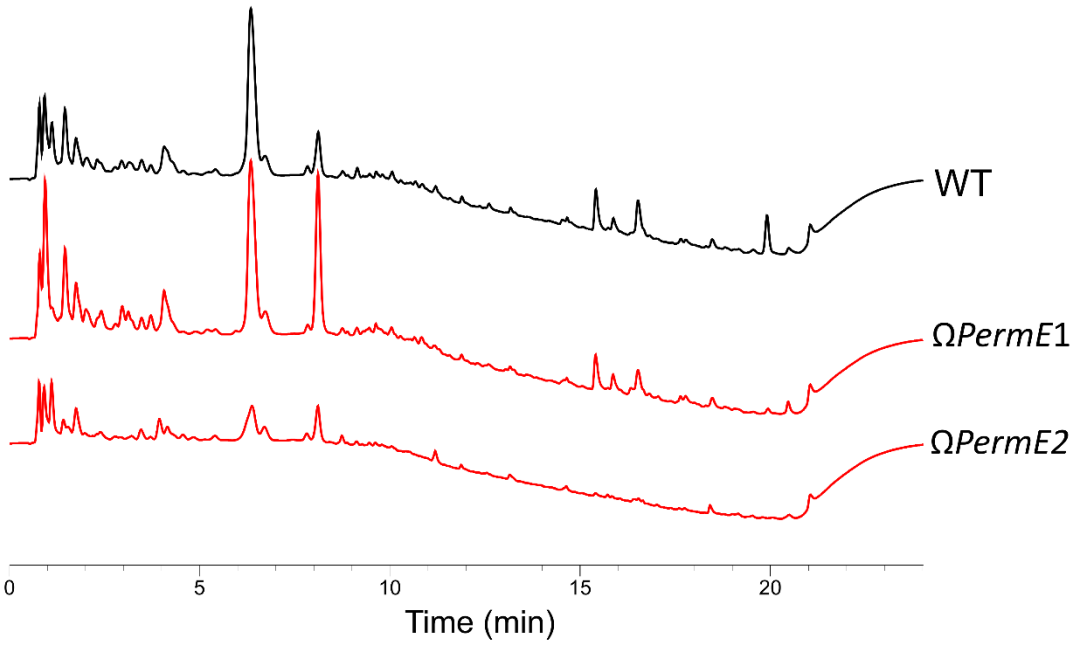


Fig 4.3 HPLC trace analysis of AUSMDU00041268 Ω*PermE*1, Ω*PermE*2 and wild type strains. Wild type (black), Ω*PermE*1 and Ω*PermE*2 mutants (red). Each isolate was cultured on ISP2 media for a period of 12 days prior to extraction and subsequent HPLC analysis.

4.3.4 Bioactivity screening

To determine whether there was any antimicrobial activity suggestive of gene cluster activation in *Nocardia* isolate AUSMDU00041268, a bioactivity screen was performed on Ω*PermE1* mutant and corresponding wild type strain. The media that was selected for bioactivity screening was Bennett's as this media type best facilitated antibiotic production overall in *Nocardia* species during the empirical screening component of this project (Chapter 3). Ω*PermE1* mutant and wild type strains were screened against a panel of five multidrug resistant pathogens: *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Enterococcus faecium* and *Acinetobacter baumannii*). Antimicrobial activity was observed in response to overlay with multidrug resistant *Acinetobacter baumannii* in both mutant and wild type strains (Fig 4.4). No activity was observed in response to the remainder of the multidrug resistant panel.

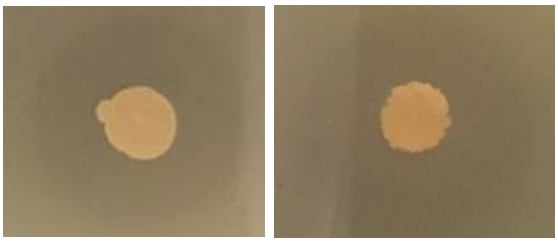


Fig 4.4 Bioactivity screen AUSMDU00041268 Ω*PermE1* mutant and wild type. AUSMDU00041268 Ω*PermE1* mutant (LHS) and wild type (RHS) both exhibiting antimicrobial activity against multidrug resistant *Acinetobacter baumannii* on Bennett's media.

4.3.5 DEREPLICATOR+ analysis

To determine if the anti-*Acinetobacter baumannii* antibiotic produced by both AUSMDU00041268 Ω*PermE1* mutant and wild type strains was novel and had potential for further investigation, a solid/liquid secondary metabolite extraction was performed. Both strains were grown on Bennett's media which facilitated antibiotic activity against MDR *Acinetobacter baumannii*. Ethyl acetate extracts were forwarded to collaborators at the Bode laboratory, Goethe University, Frankfurt, Germany for Liquid chromatography-Mass spectrometry (LC-MS) analysis. The MS/MS data files were uploaded to the Global Natural Products Social Molecular Networking (GNPS) platform for dereplication analysis via DEREPLICATOR+. The results indicated that the antibiotic observed to inhibit the growth of *Acinetobacter baumannii* in this chapter had no spectral match within the DEREPLICATOR+ secondary metabolite database indicating that the anti-*Acinetobacter baumannii* antibiotic could be novel.

4.3.6 antiSMASH analysis

The secondary metabolite biosynthetic potential of *Nocardia* wild type isolate AUSMDU00041268 was bioinformatically investigated via the antiSMASH platform. The results indicate that this isolate possesses a total of 20 SMBGCs which included: hybrids (5), NRPS (7), terpene (5), ectoine (1), and PKS (2) (Fig 4.5). Of these 20 SMBGCs, the end product of two clusters was identified (100% match) this included: cluster 3 (nocobactin NA) and

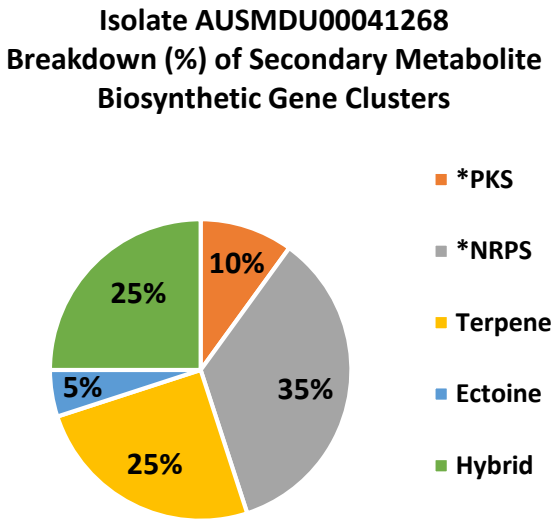


Fig 4.5 SMBCG breakdown *Nocardia* isolate AUSMDU00041268. *NRPS (non-ribosomal peptide synthetase) *PKS (polyketide synthase).

cluster 16 (ectoine). Eight gene clusters had an unknown end product (0% match) and included: cluster four (NRPS), cluster 10 (terpene), cluster 11 (terpene), cluster 14 (NRPS), cluster 15 (NRPS), cluster 17 (terpene), cluster 19 (NRPS/T1PKS hybrid) and cluster 20 (NRPS/bacteriocin hybrid) (Table 4.1).

Table 4.1 antiSMASH interrogation of *Nocardia* isolate AUSMDU00041268. Listing all 20 SMBGCs including cluster number, gene cluster type and similarity match to known compounds (%). The two gene clusters with 100% end product identification (nocobactin NA and ectoine) are highlighted in green. All eight gene clusters which produce an unknown secondary metabolite are highlighted in red. Cluster 13, the NRPS locus which was selected for promoter refactoring is shown in blue.

Cluster Number	Gene Cluster Type	Closest Match	Similarity
1	T1PKS	selvamicin	11%
2	NRPS-like	isatropolone A isotropolone B isotropolone C	9%
3	**NRPS, terpene**	**nocobactin NA**	**100%**
4	**betalactone, NRPS**	**unknown**	**Unknown**
5	T1PKS, NRPS	heterobactin	72%
6	terpene	carotenoid	27%
7	arylpolyene	streptomycin	8%
8	NRPS	leinamycin	2%
9	terpene	isorenieratene	25%
10	**terpene**	**unknown**	**Unknown**
11	**betalactone**	**unknown**	**Unknown**
12	NRPS	pentalenolactone	15%
13	**NRPS**	**herboxidiene**	**9%**
14	**NRPS-like**	**unknown**	**Unknown**
15	**NRPS**	**unknown**	**Unknown**
16	**ectoine**	**ectoine**	**100%**
17	**terpene**	**unknown**	**Unknown**
18	NRPS	calicheamicin	2%
19	**NRPS, T1PKS**	**unknown**	**Unknown**
20	**NRPS, bacteriocin**	**unknown**	**Unknown**

4.4 Discussion

There are only a limited number of constitutive promoters shown to be functional for heterologous gene expression in actinomycetes (Li et al., 2015). This chapter aimed to induce the biosynthesis of an unknown secondary metabolite in *Nocardia* from a unique, likely transcriptionally silent, biosynthetic gene cluster by utilisation of the constitutive promoter *PermE*. The selection of this promoter was based upon pervious publications conveying the success of *PermE* in activating SMBGCs in *Streptomyces* sp. (Li et al., 2009a, Li et al., 2015, Baltz, 2010). Next generation sequencing was employed to confirm a successful double-crossover mutant strain incorporating the constitutive promoter *PermE* upstream of the target cluster. This was followed by HPLC analysis of AUSMDU00041268 Ω*PermE1* and Ω*PermE2* mutant and wild type strains after 12 days of culture on ISP2 media. HPLC profiles displayed good consistency between mutant and wild type strains indicating that *PermE,* in this case, did not induce gene expression of the targeted biosynthetic gene cluster. To fully determine whether the promoter refactoring efforts in this chapter were unsuccessful, it is recommended that the HPLC analysis of mutants Ω*PermE1* and Ω*PermE2* are complemented with RT-qPCR to detect RNA transcripts indicating increased expression of the targeted NRPS gene cluster. RT-qPCR analysis was beyond the parameters of this project.

To date there have been no publications regarding utilisation of *PermE* in *Nocardia* sp. Separate experiments in our laboratory to induce expression of SMBGCs in other *Nocardia* sp. utilising *PermE* have also been unsuccessful. This indicates that the *PermE* promoter does not function in *Nocardia* sp. It is suggested that promoter refactoring of the targeted NPRS biosynthetic gene cluster in a more amenable heterologous host could lead to gene cluster activation. Previous publications utilising *S. coelicolor*, *S. lividans* and *S. albus* details successful biosynthetic gene cluster activation following *PermE* refactoring (Li et al., 2009a) (Olano et al., 2014).

The NRPS biosynthetic gene cluster targeted within this chapter is adjacent to two co-localised transcriptional regulators belonging to the ArsR (arsenic-response repressor) family namely: NmtR and KmtR. The ArsR family are metal sensing transcriptional regulators that bind promoter-operator sequences to repress gene expression when metal ion concentration is low or non-existent and de-repress in response to metal ion influx. This family of transcriptional regulators are widely expressed in bacteria. For example, *S. coelicolor* possesses 23 putative ArsR regulators (Gao et al., 2012). Each member of the ArsR family responds to an individual metal ion or a combination of metal ions. Both transcriptional regulators NmtR and KmtR have been previously identified and characterised in *M. tuberculosis* serving to de-repress gene expression in response to the metals nickel and cobalt (Campbell et al., 2007). It is possible that the transcriptional regulators NmtR and KmtR could regulate transcription of the NRPS gene cluster targeted in this chapter.

Interestingly, when isolate AUSMDU00041268 was interrogated via antiSMASH for the presence of transcriptional regulators NmtR and KmtR in any other regions of the genome a second NRPS locus presented with both transcriptional regulators adjacent to the core biosynthetic region. This suggests that transcriptional regulators NmtR and KmtR may play a role in activating NRPS biosynthetic gene clusters in *Nocardia* sp. One could hypothesise that the activation of the targeted NRPS biosynthetic gene cluster in this chapter could be dependent upon the presence of both nickel and cobalt. Exploration of this observation was beyond the parameters of this project. The success of such an approach has not been documented in *Nocardia* species to date. However, the addition of trace elements to bacterial fermentations has been explored with success in *Streptomyces* sp. Haferburg *et al.,* investigated secondary metabolite production in *S. purpurascens* grown on minimal and complex media in comparison to nickel spiked parallels. The team reported intense antibiosis against *Escherichia, Mycobacterium, Streptomyces* and *Candidia* sp. in nickel spiked extracts (Haferburg et al.,

2008). This suggests that the presence of metal ions in bacterial fermentations has great potential to activate transcriptionally silent SMBGCs. Morgenstern *et al.* showed that the addition of cobalt to *S. coelicolor* cultures increased the production of volatile compounds and novel secondary metabolites in addition to tailored versions of previously known metabolites from well-studied pathways. This highlights that the addition of metals to bacterial fermentations can not only induce secondary metabolism leading to novel secondary metabolites but yield new versions of previously known compounds (Morgenstern et al., 2015). It is suggested that future work on the NRPS targeted gene cluster be carried out by repeating the experimental steps within this chapter and supplementing growth media with varying concentrations of cobalt and nickel prior to HPLC analysis.

A bioactivity screen against five multidrug resistant pathogens was performed on AUSMDU00041268 Ω*PermE1* mutant and corresponding wild type strain. A significant zone of inhibition (2cm) was observed in both Ω*PermE1* mutant and wild type strains when grown on Bennett's agar and subsequently screened against multidrug resistant *Acinetobacter baumannii*. The induction of secondary metabolite biosynthetic genes that yielded this antibiotic cannot be attributed to the promoter refactoring efforts within this chapter as the zone of inhibition was observed in both Ω*PermE1* mutant and wild type strains. The biosynthesis of the anti-*Acinetobacter* antibiotic observed in this chapter was facilitated by components Bennett's media further highlighting the suitability of this media type to induce SMBCGs in *Nocardia* sp. To determine whether the anti-*Acinetobacter* antibiotic had been previously identified, ethyl acetate extracts of AUSMDU00041268 wild type strain were uploaded to the DEREPLICATOR+ component of the Global Natural Products Social Molecular Networking (GNPS) platform. No hits for known secondary metabolites were obtained from the database suggesting that anti-*Acinetobacter baumannii* antibiotic identified within this chapter could be novel and warrants further characterisation.

AntiSMASH analysis indicates that isolate AUSMDU00041268 is well endowed with a total of 20 SMBGCs, only two of which were identified with 100% certainty: cluster 3 (ectoine) and cluster 16 (nocobactin NA), neither of which exhibit the anti-Gram-negative activity observed during the bioactivity screening component of this chapter. Additionally, eight of the 20 biosynthetic gene clusters had no known similarity with any other gene clusters in their comprehensive database. It is possible that the gene cluster responsible for the biosynthesis of the anti-*Acinetobacter* antibiotic observed within this chapter falls into one of these eight unidentified gene clusters. This provides further evidence that this antibiotic is yet undiscovered; highlighting its potential as a future drug lead against multidrug resistant *Acinetobacter baumannii*.

Chapter 5: Molecular network analysis

5.1 Introduction

Liquid chromatography-mass spectrometry is an efficient method for the high-throughput analysis of microbial metabolites based on their mass, chromatographic and spectroscopic properties. This technique can be used to assist in defining structural relationships between metabolites. However, this is most efficiently done by comparing raw spectral data from unknown molecules with a database of known, structurally defined compounds. Until recent years, the field of metabolomics was reliant upon commercial spectral databases that were often proprietary and predominantly centered around primary metabolism (METLIN , XCMS). In the natural product discovery field, there is a distinct need for a unified, collaborative, open access database comprising raw, processed and identified secondary metabolite spectral data. Such databases are based upon data exchange that is contributed by the research community itself. In addition, a unified repository for spectral data aids natural product discovery researchers in the rapid identification of previously known metabolites, thus ensuring time and resources are not wasted persuing previously discovered secondary metabolites.

The Global Natural Products Social Molecular Networking (GNPS) platform is a widely-used, curated mass spectral database with computational tools to analyse and interpret complex mass spectral data, including molecular network analysis. Molecular networking is a powerful tool for assessing large metabolomic datasets, allowing researchers to visualise the chemical space present in tandem mass spectrometry (MS/MS) experiments. The process begins by collection of MS^1 spectra which ionises the analyte and separates molecules by their mass-to-charge ratio (*m/z*). Ions of a particular *m/z* are selected from the MS^1 pool and split into smaller fragment ions (MS/MS or MS^2) and subsequently processed by a GNPS spectral networking algorithm. This algorithm converts MS/MS data into vectors, this output can then be visualised as a molecular network. Within a molecular network, fragmented ion masses are grouped into

clusters that represent molecular families (molecules with structural similarities), even when the spectra do not match any known compounds. A central component of this platform is an extensive, community aquired reference library of spectral data, which facilitates rapid dereplication of previously identified metabolites (DEREPLICATOR+) and allows researchers to prioritise bacterial isolates in a time and resourse sensitive manner (Wang et al., 2016, Mohimani et al., 2018).

Analysing the secondary metabolomes of a range of closely related bacterial species from crude fermentation extracts provides opportunities to identify secondary metabolites that are genus and species-specific. Such spectral datasets can be comparatively assessed under a range of different culture conditions such as incubation time, extraction protocols and media type shedding light on the most favourable conditions to facilitate the induction of SMBGCs, therefore optimising overall secondary metabolite production.

This chapter surveyed the biosynthetic potential of a range of *Nocardia* isolates *(n = 10)* which were cultured on five alternate media types (Bennett's, rare 3, M1, minimal and ISP2). Each isolate was representative of a distinct *Nocardia* species: *N. terpenica, N. ignorata, N. abscessus, N. cerradoensis, N. arthritidis, N. pneumoniae, N. brasiliensis, N. beijingensis, N. nova* and a single *Nocardia* isolate that could not be speciated. All isolates had been previously sequenced at the Peter Doherty Institute, University of Melbourne and all possess a diverse range of SMBGCs of unknown potential.

Molecular networking was employed to systematically evaluate the molecular footprint of each *Nocardia* isolate to highlight common molecular families, isolate specific, rare secondary metabolites, and assess the most suitable culture media for secondary metabolite discovery in *Nocardia*.

5.2 Chapter aims

I. To compare the secondary metabolomes of ten *Nocardia* species grown on a range of culture media by utilisation of LC-MS/MS and molecular networking

II. To prioritise *Nocardia* strains and optimal media type for increased secondary metabolite recovery

All *Nocardia* isolates were cultured on five distinct media types and their secondary metabolomes comparatively analysed. Culture media were selected based on the media types that best facilitated antibiotic production during the primary screening phase of this project (Chapter 3). Isolates were inoculated and cultured in a 96-well format as outlined in materials and methods (Chap 2.9). Following a 12 day incubation period, secondary metaboltes were extracted and analysed by LC-MS/MS (Fig 5.1). The resulting spectral data were uploaded to the GNPS platfom for dereplication and analysis.

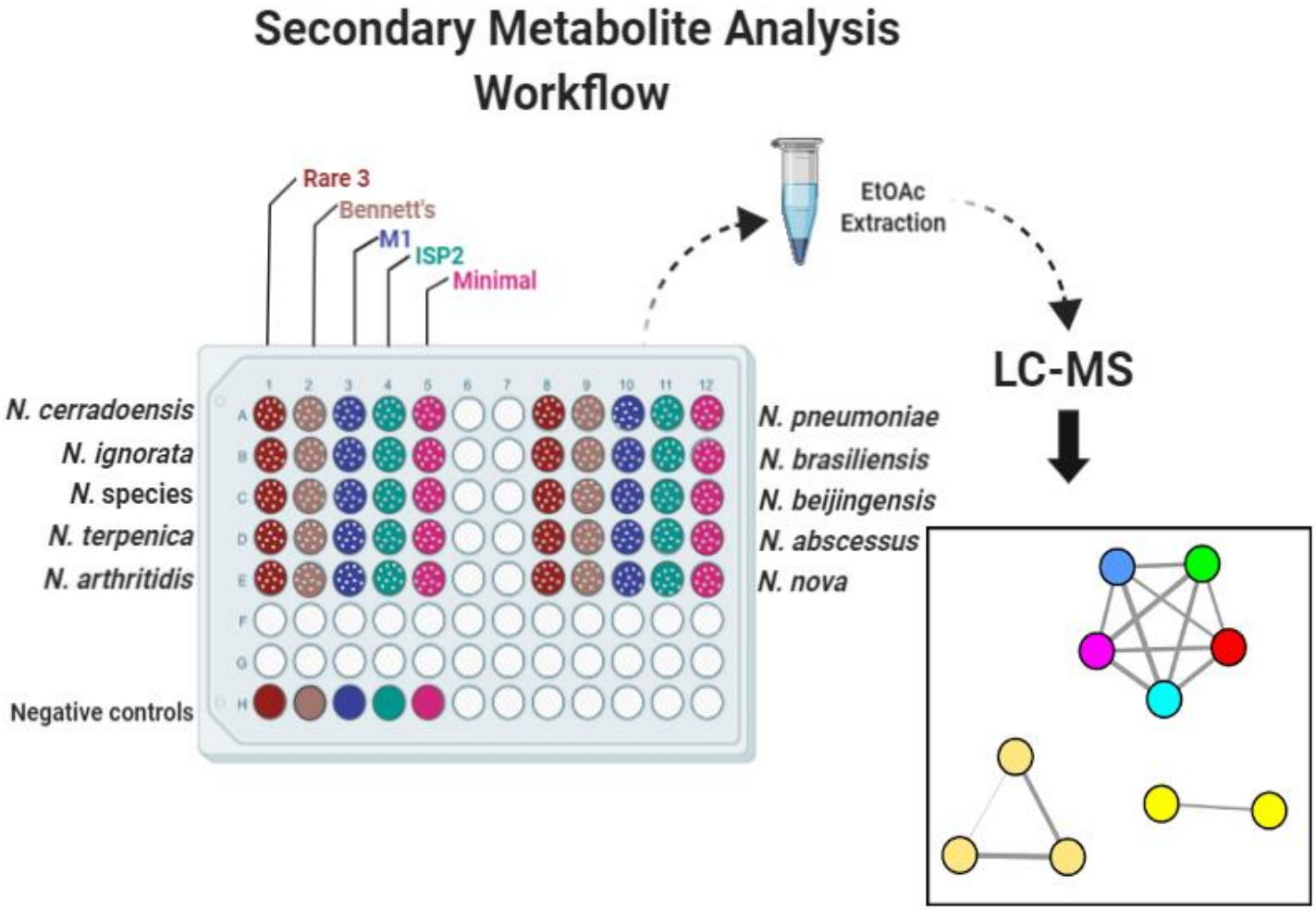


Fig 5.1 Experimental workflow for the analysis of secondary metabolites from a range of *Nocardia* species. Each isolate was grown on five distinct media types. After a period of 12 days, a plug from each well was taken and ethyl acetate extracts were created. All extracts were forwarded to the Bode laboratory, Goethe University, Frankfurt, Germany for LC-MS/MS analysis. The resulting spectral files were subsequently analysed via molecular networking by utilisation of the GNPS platform. Image created using BioRender software (biorender.com).

5.3 Results

5.3.1 Principal coordinates analysis of biological replicates

Each *Nocardia* isolate was incubated for a period of 12 days prior to ethyl acetate extraction. Two biological replicates were created for each isolate grown on all five media types. A principal coordinates analysis (PCoA) was conducted on each biological replicate to ensure reproducibility among replicate plates (Fig 5.2 a-e). Of the ten *Nocardia* species analysed, scattering of the mass spectral data for six isolates occurred in a similar region within the plots, showing that biological replicates were reproducible and representative of the true chemical diversity of each isolate. The remaining four isolates (*N. nova*, *N. cerradoensis*, *N. arthritidis* and *N. beijingensis*) showed variability in the spectral scattering on three of the five media types (M1, ISP2 and rare 3) indicating differences in the resulting secondary metabolite profile of each biological replicate.

5.3.2 Principal coordinates analysis of secondary metabolite diversity

Secondary metabolite diversity was assessed amongst plots based on media type. Plots associated with *Nocardia* grown on Bennett's media showed that all isolates were closely clustered within the chemical space indicating limited secondary metabolite diversity (Fig 5.2a). Secondary metabolite diversity for *Nocardia* isolates that were grown on minimal and M1 media were grouped into two distinct regions of each plot indicating an increased level of secondary metabolite diversity in comparison to that of Bennett's media (Fig 5.2c and e). *Nocardia* isolates grown on rare 3 and ISP2 media exhibited a scattered plot distribution that occupied multiple regions of the chemical space indicating variety and diversity of secondary metabolites produced on both media types (Fig 5.2b and d).

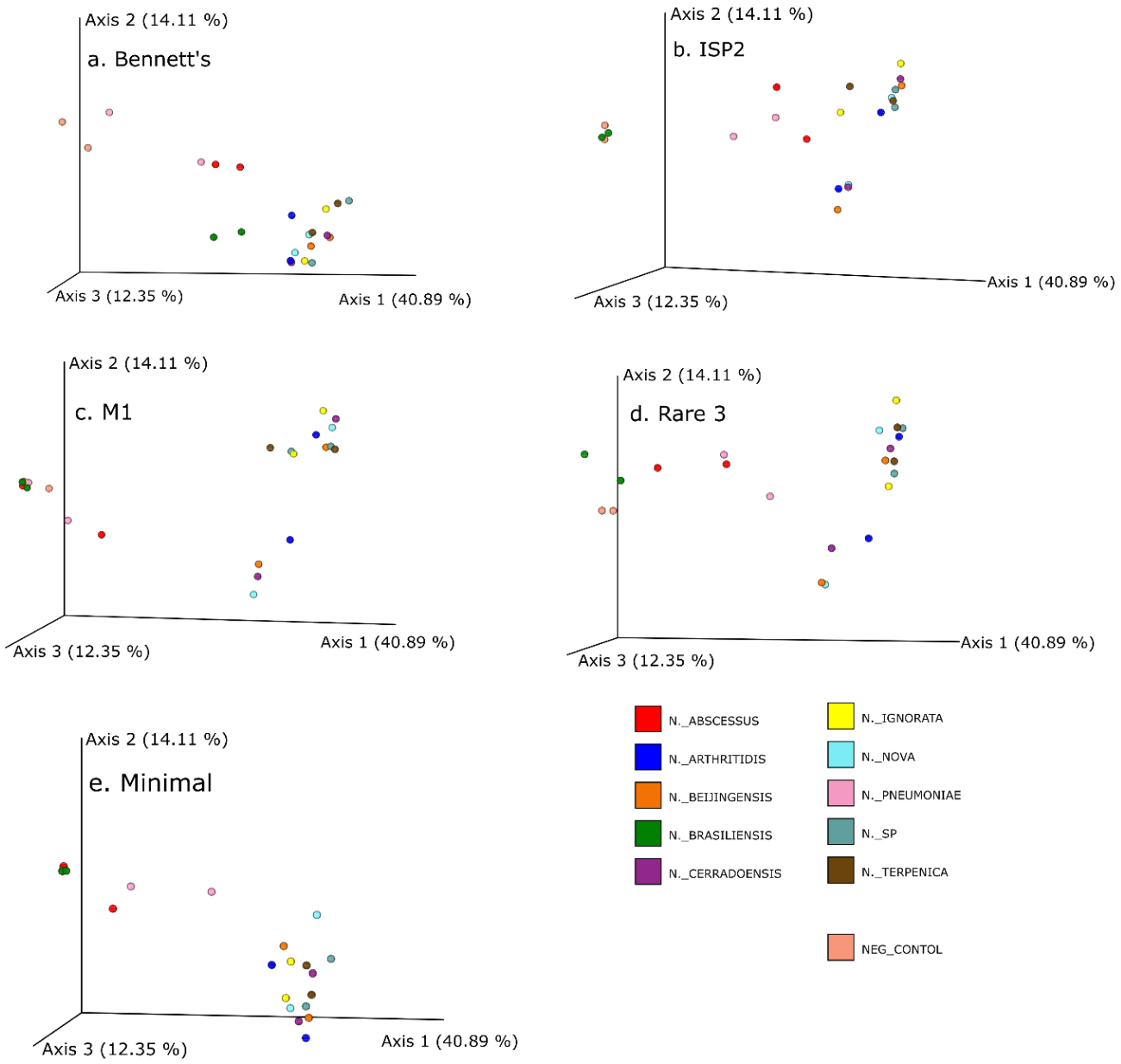


Fig 5.2 Principal coordinates analysis (PCoA) plots. Two biological replicates were included for each *Nocardia* isolate grown on five media types. Each coloured circle indicates an individual *Nocardia* isolate. There are two coloured circles associated with each isolate (biological replicate 1 and 2). Negative controls (uninoculated media) are also shown within each plot.

5.4 Complete molecular network analysis

By utilisation of the GNPS platform, a complete molecular network comprised of all secondary metabolites expressed by all ten *Nocardia* isolates, grown on five different media types was generated (Fig 5.3). Secondary metabolites with structural similarities are comprised of similar parent ion fragmentation patterns. This similarity between parent ion fragmentation patterns can be translated into a cosine score; a score of one meaning identical fragmentation spectra and zero meaning different parent ions. Within a molecular network, parent ions are visualised as nodes which are connected by edges (cosine score). The wider the edges the closer the chemical architecture of the metabolites. The length of each edge, in the context of molecular networking, has no meaning in terms of relationship between nodes.

5.4.1 Molecular network breakdown

A total of 110 individual experimental samples were used to construct the molecular network, producing 33,830 spectra overall. The complete network was comprised of 3013 nodes prior to clean-up. Nodes associated with negative controls were removed in addition to nodes that did not appear in both biological replicates. Post clean-up, the network was comprised of 2287 nodes and 2198 edges. Of the 2287 nodes, 52% were grouped into 171 individual sub-networks, which were defined as having two or more related nodes connected by edges. Each sub-network is representative of a distinct molecular family of secondary metabolites with structural similarities. The remaining 48% of nodes that comprised the network were sufficiently unique that they formed no linkages with any of the remaining spectra. It should be mentioned, that the number of nodes within a sub-network is not equal to the number of metabolites present. Different charge states and/or adducts of the same metabolite can result in the visualisation of multiple nodes of the same chemical species.

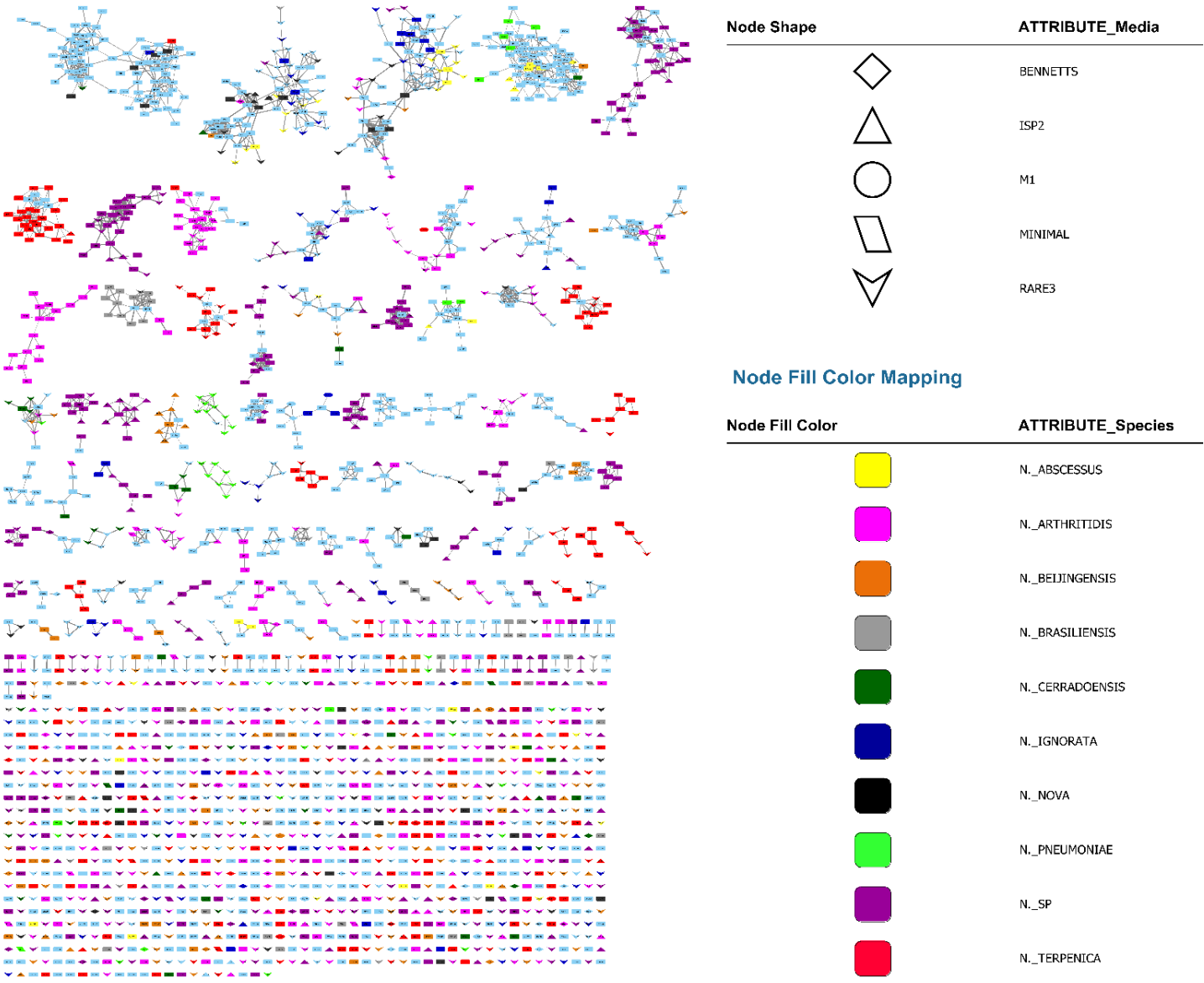


Fig 5.3 Complete molecular network. Comprised of ten *Nocardia* species grown on five distinct media types. The network shows the grouping of 171 sub-networks (two or more connected nodes) representing distinct molecular families of secondary metabolites. The colour of each node corresponds to an individual *Nocardia* species that exclusively expresses a specific metabolite. Rectangular nodes coloured in light blue highlight metabolites produced by multiple *Nocardia* species on a range of media types. The shape of each node is correlated to the media type in which the metabolite was produced.

5.4.2 Isolate specific nodes (ISNs) by species

Of the ten *Nocardia* isolates examined, the species that produced the greatest number of ISNs (metabolites only expressed by that isolate) was a *Nocardia* isolate that could not be speciated (360 nodes). In addition, *N. arthritidis* and *N. terpenica* produced significantly more ISNs than the remaining cohort (289 and 229 nodes respectively). Isolates that had the least number of unique nodes were *N. abscessus* (42 nodes) and *N. cerradoensis* (45 nodes).

5.4.3 Species-specific optimal media type

Of the media types assessed (Bennett's, ISP2, M1, minimal and rare 3), the number of ISNs were greater when isolates were grown on rare 3 media (663 nodes) followed by ISP2 (173 nodes), Bennett's (71 nodes) minimal (29 nodes) and M1 (19 nodes).

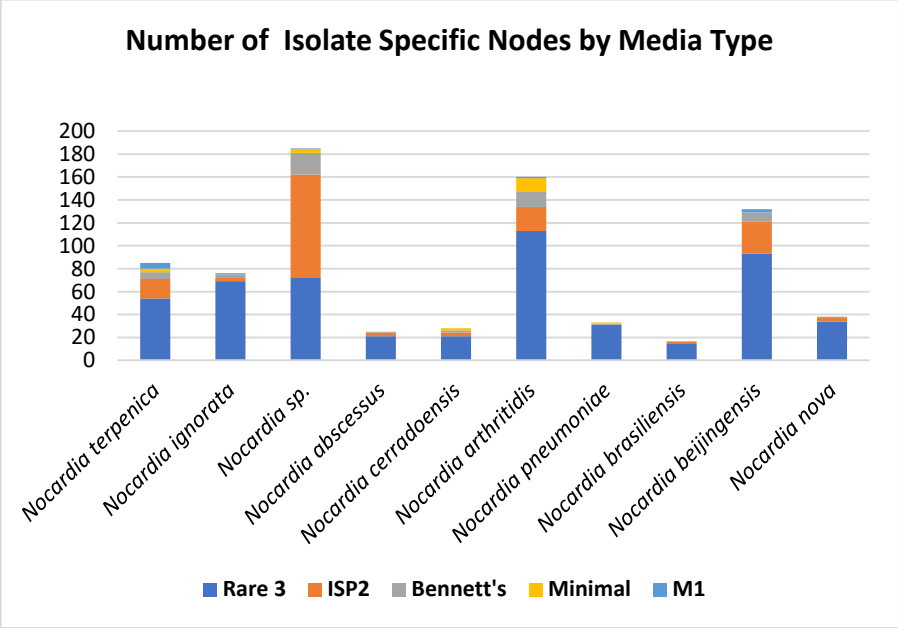


Fig 5.4 Number of isolate specific nodes by species and media type.

The most suitable media type to facilitate maximum secondary metabolite output was assessed in each individual *Nocardia* species. Rare 3 media was associated with higher numbers of nodes within the complete network for all *Nocardia* isolates (with the exception of a single *Nocardia* isolate that could not be speciated). In this case, growth on ISP2 resulted in slightly higher numbers of nodes in comparison to rare 3 (90 nodes and 72 nodes respectively) (Fig 5.4).

5.4.4 Nocardia core metabolites

Of all ten species of *Nocardia* examined only five nodes were common to all *Nocardia* isolates on all media types. These nodes spanned two sub-networks indicative of two distinct molecular families of metabolites. DEREPLICATOR+ analysis on each of these sub-networks yielded no reference metabolite within the database that shared the same parent masses. Therefore, it was not possible to assign a structure/molecular family designation to each of the five conserved nodes.

5.5 DEREPLICATOR+ analysis of complete molecular network

MS/MS data from all ten *Nocardia* isolates grown on five distinct media types was uploaded to the GNPS platform for dereplication analysis. A total of five previously known metabolites and their analogues were identified by exact masses expressed by multiple species of *Nocardia* following DEREPLICATOR+ analysis (Table 5.1). Secondary metabolite hits were associated with four sub-networks of distinct molecular families within the complete molecular network. Identified secondary metabolites included: peptidolipin (NA and L-Val (6)), brasilibactin A, nocardimicin (A, B and G), xenoamicin (A and B) and detoxin D1 (Fig 5.5).

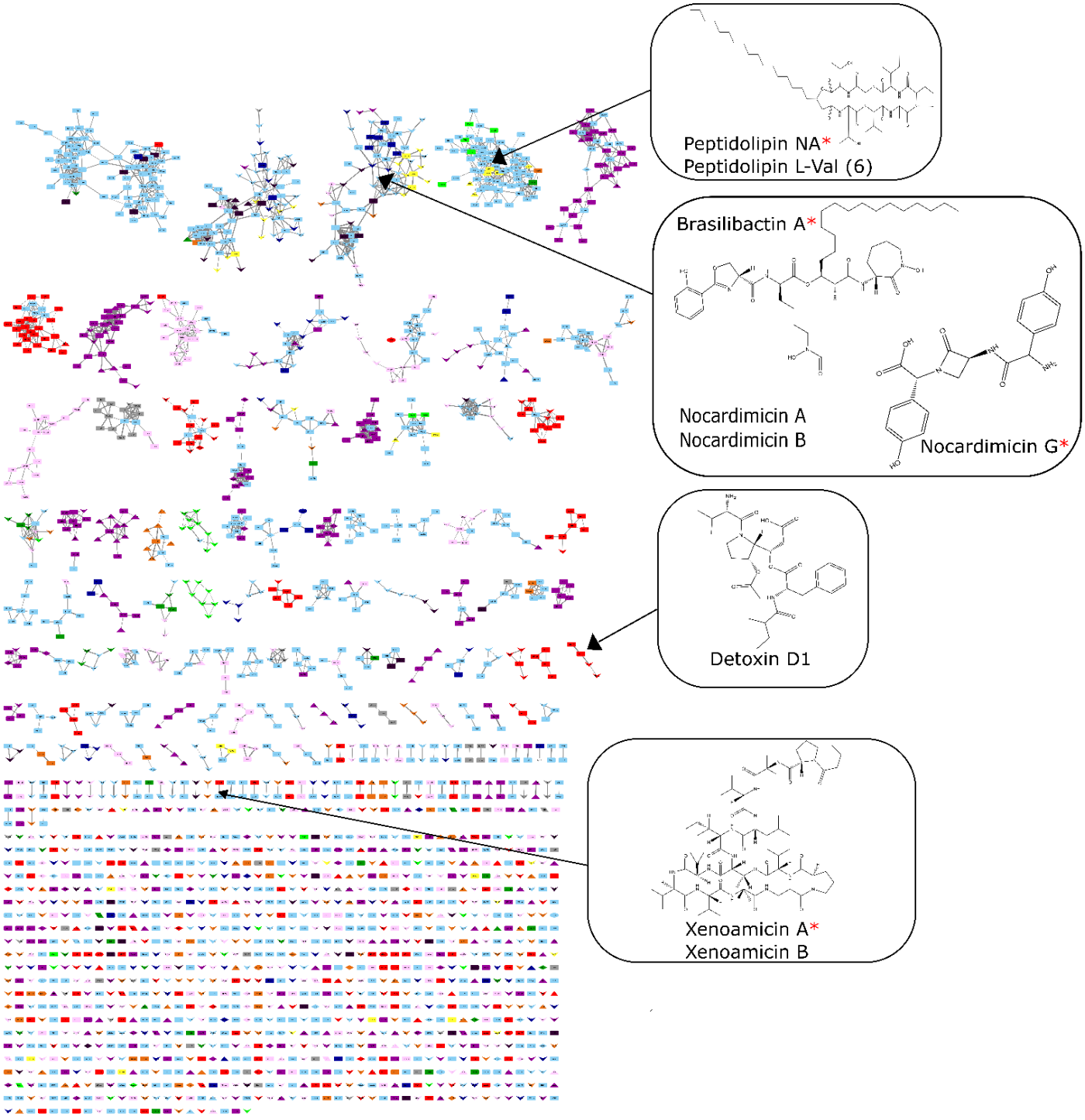


Fig 5.5 Molecular network and DEREPLICATOR+ hits of known metabolites. DEREPLICATOR+ hits are shown and the molecular family in which they are grouped highlighted. * Indicates molecular structure illustrated.

Table 5.1 DEREPLICATOR+ hits and their analogues. *Nocardia* species in which each metabolite was detected and the media types that facilitated biosynthesis. *(a) Bennett's, (b) ISP2, (c) rare 3, (d) M1 and (e) minimal.

16S rRNA sequencing ID	**Peptidolipin NA**	**Peptidolipin L-Val (6)**	**Nocardimicin A**	**Nocardimicin B**	**Nocardimicin G**	**Brasilibactin A**	**Detoxin D1**	**Xenoamicin A**	**Xenoamicin B**
Media type inducing expression*	**(a/ b/c/d/e)**	**(a/b/c/d/e)**	**(a/c)**	**(a/b/c/e)**	**(a/b/c)**	**(a/b/c)**	**(a/b/c/d)**	**(a/b/c/e)**	**(a/e)**
N. terpenica							✓	✓	
N. ignorata			✓	✓				✓	✓
N. sp.			✓	✓				✓	✓
N. abscessus	✓	✓	✓						
N. cerradoensis		✓			✓	✓		✓	✓
N. arthritidis					✓	✓		✓	
N. pneumoniae	✓	✓				✓			
N. brasiliensis					✓	✓			
N. beijingensis	✓	✓			✓	✓		✓	✓
N. nova					✓	✓		✓	✓

5.5.1 Peptidolipin NA and L-Val (6)

The macrocyclic lactone peptidolipin NA and its analogue L-Val (6) was first isolated from *N. asteroides* (Guinand and Michel, 1966a). The antifungal activity of peptdolipin NA has been previously documented (Maget-Dana et al., 1985). Peptidolipin NA was expressed by 30% of the *Nocardia* isolates tested (*N. abscessus*, *N. pneumoniae* and *N. beijingensis*) and its analogue peptidolipin L-Val (6) was produced by 40% of isolates (*N. abscessus*, *N. cerradoensis*, *N. pneumoniae* and *N. beijingensis*). The peptidolipins were readily produced on all five media types selected for screening (Table 5.1).

5.5.2 The peptidolipin sub-network

Due to the clinical relevance of the lipopeptide family as a source of potent antimicrobials (daptomycin) an additional analysis of the peptidolipin sub-network was conducted. This sub-network represented the fourth largest present within the complete molecular network, comprised of 72 nodes in total. Of these 72 nodes, two were identified via DEREPLICATOR+ analysis by exact parent mass (peptidolipin NA and peptidolipin L-Val (6)) (Fig 5.6). The peptidolipin family represent a rare class of secondary metabolites characterised by peptide cyclisation via an ester to a lipophilic tail. Interrogation of this sub-network for the positive identification of known metabolites within the GNPS database yielded no further hits. To date, five additional members of the peptidolipin family have been identified (peptidolipin B-F) (Wyche et al., 2012). Interrogation of the peptidolipin sub-network for masses associated with peptidolipin B-F yielded no further hits. Other structurally related lipopeptides include surfactin, iturin A, and bacillomycin D which are ordinarily produced by *Bacillus subtilis*. The peptidolipin sub-network was interrogated for masses that matched surfactin, iturin A, and bacillomycin D but yielded no additional hits.

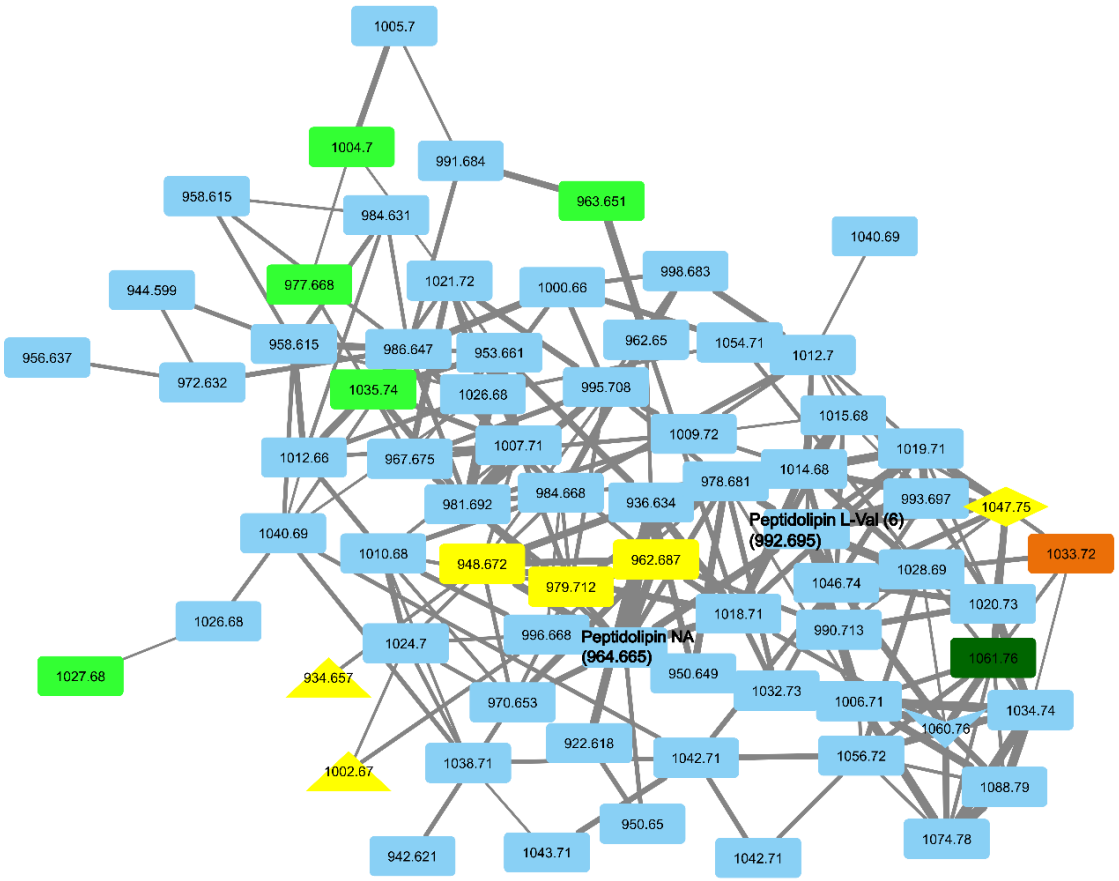


Fig 5.6 The lipopeptide sub-network. Comprised of 72 nodes, DEREPLICATOR+ hits peptidolipin NA and peptidolipin L-Val (6) are highlighted. Light blue rectangular nodes indicate metabolites produced by a range of *Nocardia* species on multiple media types. Edge thickness indicates cosine score between nodes. Yellow nodes indicate metabolites exclusively produced by *N. abscessus*, dark green (*N. cerradoensis)*, orange (*N. beijingensis*) and bright green (*N. pneumoniae*). Rectangular shaped nodes indicate metabolites exclusively produced on ISP2 media, diamond (Bennett's) and arrow (rare 3).

5.5.3 Brasilibactin A and nocardimicin A, B and G

The brasilibactins and nocardimicins represent bioactive siderophores which to date have only been shown to be produced by *Nocardia* sp. Both siderophores and their analogues were matched by exact masses and grouped into a distinct sub-network of structurally similar secondary metabolites within the molecular network.

Brasilibactin A was first isolated from *N. brasiliensis* IFM 0995 (Tsuda et al., 2005). This siderophore exhibits antibacterial activity against *M. luteus* and *S. aureus*. In addition, brasilibactin A displays potent cytotoxicity in murine leukaemia L1210 and human epidermoid carcinoma KB cell lines (Tsuda et al., 2005). Brasilibactin A was detected in 60% of isolates (*N. cerradoensis, N. arthritidis, N. pneumoniae, N. brasiliensis, N. beijingensis* and *N. nova).* This was observed on three of the five media types (Bennett's, ISP2 and rare 3) (Table 5.1).

To date there have been nine nocardimicin analogues identified (nocardimicin A-I). The biological activity of these *Nocardia* associated siderophores was discovered during screening for muscarinic M3 receptor inhibiting compounds. M3 receptors control the contraction of smooth muscle therefore the antagonistic effect of the nocardimicins on M3 receptors show promise in the treatment of respiratory, gastrointestinal and urinary tract disease. The nocardimicins (A and B) to date, have only been associated with *Nocardia* sp. TP-A0674. These results show the production of Nocardimicin A in 30% of isolates (*N. ignorata, N. abscessus* and a *Nocardia* isolate that could not be speciated) grown on Bennett's and rare 3 media. Nocardimicin B was produced by 20% of isolates (*N. ignorata* and a *Nocardia* isolate that could not be speciated) on Bennett's, ISP2, rare 3 and minimal media. Nocardimicin G was produced by 50% of *Nocardia* isolates (*N. cerradoensis, N. arthritidis, N. brasiliensis, N. beijingensis* and *N. nova*) on Bennett's, ISP2 and rare 3 media (Table 5.1).

5.5.4 Xenoamicin A and B

Xenoamicins are acylated tridecadepsipeptides produced by the bacteria *Xenorhabdus doucetiae* DSM17909 and *X. mauleonii* DSM17908. The non-ribosomal peptide biosynthetic gene cluster responsible for xenoamicin biosynthesis (XabABCD) was elucidated in *X. doucetiae* and its involvement in xenoamicin biosynthesis was confirmed by mutagenesis (Zhou et al., 2013). Bioactivity assays revealed that the xenoamicins have activity against *Plasmodium falciparum* highlighting potential as an antimalarial treatment. Xenoamicin A was expressed by 70% of isolates (*N. terpenica*, *N. ignorata*, *N. cerradoensis*, *N. arthritidis*, *N. beijingensis*, *N. nova* and a *Nocardia* isolate that could not be speciated) grown on Bennett's, ISP2, rare 3 and minimal media types. Xenoamicin B was expressed by 50% of isolates (*N. ignorata*, *N. cerradoensis*, *N. beijingensis*, *N. nova* and a *Nocardia* isolate that could not be speciated) grown on Bennett's and minimal media types (Table 5.1).

5.5.5 Detoxin D1

Detoxin D1 is a depsipeptide isolated from *Streptomyces caespitosus var. detoxicus* 7072 GC1. Detoxin D1 exhibits unusual biological activity in the detoxification of the nucleoside antibiotic blasticidin S (produced by *Streptomyces griseochromogenes*) in plant, animal and bacterial cells (Yonehara et al., 1968). Blasticidin S inhibits protein synthesis in both prokaryotic and eukaryotic cells. Detoxin D1 was expressed by a single *Nocardia* isolate (*N. terpenica*) on Bennett's, ISP2, rare 3 and M1 media types (Table 5.1).

5.6 Correlation of genomic data with identified secondary metabolites

5.6.1 Brasilibactin A and nocardimicin A, B and G

Multiple mycobactin-like siderophores have been previously isolated from *Nocardia* species including brasilibactin, nocardimicin, and nocobactin (Tsuda et al., 2005, Ikeda et al., 2005b, Ratledge and Snow, 1974). These salicylate-derived siderophores all have similar structures, differing in the position of methyl and hydroxyl groups, and the length of the attached lipid chains. Brasilibactin A and nocardimicin A, B and G was detected in the crude extracts of multiple *Nocardia* species following LC-MS/MS analysis and dereplication via GNPS. AntiSMASH analysis of the genome sequences of each producing isolate indicated the presence of a nocobactin-like biosynthetic gene cluster in all cases. The biosynthetic gene cluster responsible for nocobactin NA biosynthesis has been previously identified in *N. farcinica* IFM 10152 (Hoshino et al., 2011). The brasilibactins and nocardimicins detected in this chapter share remarkable structural similarities to nocobactin NA indicating a common biosynthetic origin (Fig 5.7). A comparison of the known nocobactin NA biosynthetic gene cluster to the nocobactin-like gene cluster identified in this chapter indicated identical domain

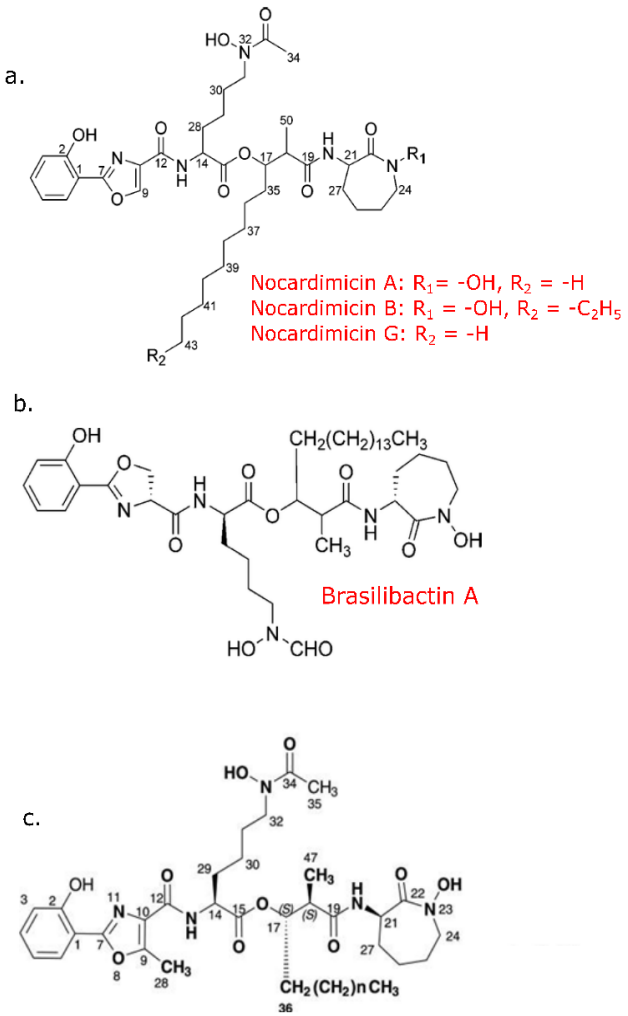


Fig 5.7 Siderophores produced by *Nocardia* species (a) nocardimicin A, B and G (b) brasilibactin A and (c) nocobactin NA

arrangements with the exception of an aryl carrier protein, cyclisation, adenylation and peptidyl carrier protein domain that was only present in the known nocobactin NA biosynthetic gene cluster. This suggests that the nocobactin-like gene cluster identified in multiple *Nocardia* species within this chapter could be responsible for the biosynthesis of the nocardimicins and brasilibactins. However, an extensive comparative investigation of these gene clusters was beyond the scope of this project.

5.6.2 Xenoamicin A and B

Xenoamicin A and B was detected in the crude extracts of multiple *Nocardia* species within this chapter. At present, the biosynthesis of the xenoamicins has only be associated with the bacterial genus *Xenorhabdus.* Here, xenoamicin biosynthesis has been reported in multiple *Nocardia* species. AntiSMASH interrogation of all xenoamicin producing isolates failed to identify a common xenoamicin-like gene cluster common to all producing isolates. However, a NRPS locus was observed in *N. ignorata* that shared a similar domain structure with the exception of two terminal thioesterase domains when compared to the known xenoamicin biosynthetic gene cluster. In addition, the proposed gene cluster possessed 13 adenylation domains indicating the incorporation of 13 amino acid residues concordant with xenoamicin biosynthesis.

5.6.3 Detoxin D1

The biosynthetic origin of detoxin D1 has previously been determined as a hybrid PKS-NRPS encoding gene cluster in *Streptomyces morbarensis* NRRL B-3729(McClure et al., 2016). Detoxin D1 was detected in the crude extracts of a single *N. terpenica* isolate on four of the five media types assessed (Bennett's ISP2, rare 3 and M1). Whole genome sequencing utilising a combination

of PacBio and Illumina DNA sequencing was employed to assemble the complete *N. terpenica* genome to correlate detoxin D1 production with a putative biosynthetic locus.

AntiSMASH analysis indicated that *N. terpenica* possesses a total of 48 secondary metabolite biosynthetic gene clusters. A 13 kb NRPS/PKS hybrid biosynthetic gene cluster was identified that had the same arrangement of biosynthetic genes and encoded domains as previously found for detoxin D1 biosynthesis. The NRPS/PKS locus contained two genes, encoding three adenylation domains indicating the incorporation of three amino acid residues concordant with detoxin D1 biosynthesis (Goering et al., 2016). Polyketide derived domains of the known detoxin D1 gene cluster (ketosynthase (KS) and ketoreductase (KR)) matched in domain order and location with the identified *N. terpenica* biosynthetic gene cluster.

The amino acid composition of Detoxin D1 is known (isoleucine, proline and phenylalanine). Adenylation domains in *N. terpenica* are predicted to select for isoleucine, proline and valine. AntiSMASH analysis of the adenylation domains in known detoxin D1 producer *S. morbarensis* NRRL B-3729 has the same predicted amino acid specificity as *N. terpenica,* incorrectly predicting the monomer valine instead of phenylalanine. This provides further evidence that the *N. terpenica* cluster is the correct candidate for detoxin D1 production.

BLAST analysis of the core NRPS (locus tag Nter_03870) and PKS/NRPS (locus tag Nter_03869) proteins encoded within the *N. terpenica* gene cluster showed 58% and 65% conservation between their corresponding orthologues in the detoxin D1 biosynthetic gene cluster respectively. In addition, a putative dioxygenase (Nter_03866), likely involved in oxidation at the C3 position of the proline, showed 69% conservation with a taurine dioxygenase encoded within the known detoxin biosynthetic gene cluster (Fig 5.6).

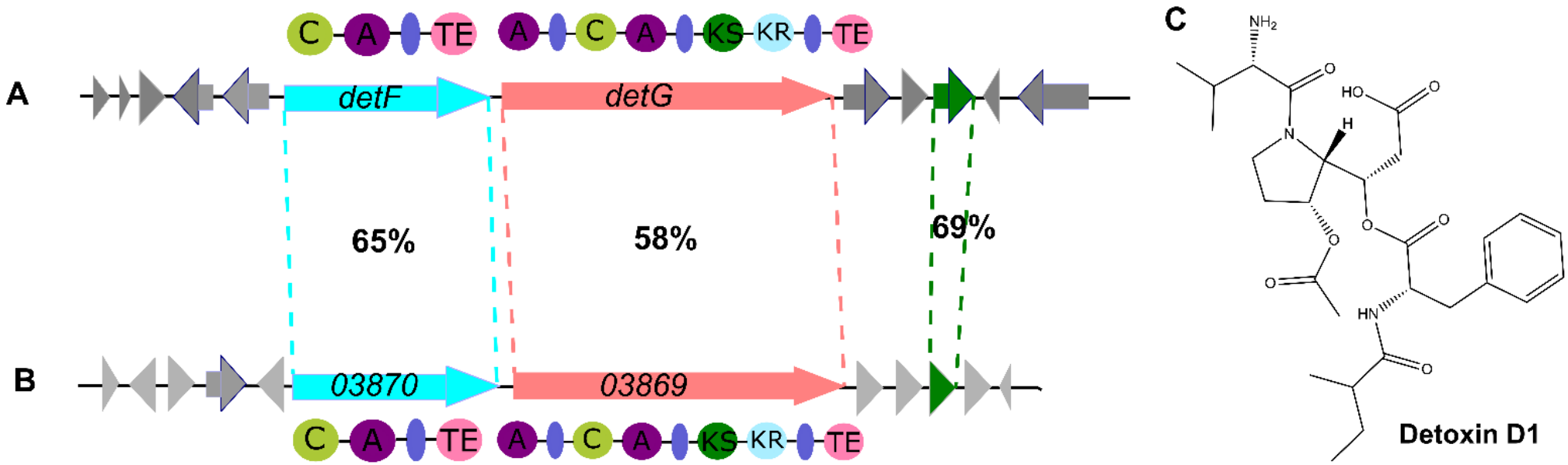


Fig 5.8 Detoxin D1 biosynthetic gene cluster comparison. (A) Known biosynthetic gene cluster of detoxin D1. **(B)** Detoxin-like biosynthetic gene cluster in *N. terpenica.* Highlighting the identical domain organisation of detoxin D1 and detoxin-like biosynthetic gene clusters **(C)** The molecular structure of detoxin D1. Genes depicted in cyan (NRPS), pink (NRPS/PKS) and green (dioxygenase). Homologous open reading frames (grey) and percentage sequence identity between known detoxin biosynthetic gene cluster and detoxin-like gene cluster present in *N. terpenica*.

5.7 Discussion

The objective of this chapter was to systematically investigate the biosynthetic potential of ten previously sequenced *Nocardia* species grown on five distinct media types and to compare strains based on the variety and novelty of their secondary metabolites.

This chapter highlighted the chemical diversity present in the under-investigated genus *Nocardia*; with the additional dynamic of media type influencing the scope and diversity of secondary metabolite production. Here, an easily executed, time and resource efficient protocol was devised for the extraction of secondary metabolites from solid agar. This appears to be the first such study conducted in this fashion, focusing on the *Nocardia*. This observation perhaps explains why there were a limited number of DEREPLICATOR+ hits obtained for previously identified metabolites, highlighting the significant untapped chemical diversity of the *Nocardia.* This finding also underscores the importance of spectral data contributed directly from the research community to platforms like GNPS to aid in metabolite identification.

This chapter successfully identified five distinct families (and their analogues) of metabolites that appear to be produced by various *Nocardia* species. This highlights the potential of the *Nocardia* to yield therapeutically valuable metabolites that are not limited to the realm of antibacterials: (peptidolipin/antifungal, brasilibactin/cytotoxic, xenoamicin/antimalarial, nocardimicin/ muscarinic M3 receptor antagonist and detoxin D1/detoxification of blasticidin S). These results are further accentuated by the importance of media selection which this chapter has shown can vastly affect the resulting chemical diversity present within a molecular network. Rare 3 media proved to be the best facilitator of secondary metabolite production overall differing from its counterparts by the presence of magnesium chloride (2 g/L) and glycerol 100% (5 g/L). The importance of magnesium has been observed in energy production by binding and activating

adenosine triphosphate (ATP) as well as playing a vital role in genomic stability and serving as an essential cofactor of enzymes involved in DNA processing (Ko et al., 1999, Hartwig, 2001). Although trace amounts of Mg^{2+} are expected to be present in all media types assessed, the addition of magnesium chloride at a concentration of 2 g/L (or above) may better facilitate secondary metabolite production in *Nocardia*. The presence of glycerol in bacterial fermentations can serve as an alternative carbon source and also positively impact cell viability and productivity in recombinant protein production (Kopp et al., 2017). Assessing the effect of magnesium and glycerol on increased secondary metabolite production was beyond the parameters of this project. However, it is speculated that the presence of either component (alone or in combination) may contribute to increased secondary metabolite biosynthetic gene cluster induction in *Nocardia* species.

Interestingly, when the molecular network was analysed for nodes present in all species of *Nocardia* across all media types (core *Nocardia* metabolites) only five nodes were detected. This highlights the highly variable metabolome of the *Nocardia* where acquisition of biosynthetic genes (and the induction of such genes) varies greatly amongst each individual species. Additionally, this emphasises the importance of species-specific growth media matching to unlock the full potential of selected isolates for bioprospecting. Considering this, three *Nocardia* isolates could be prioritised for further investigation alongside species compatible media to maximise secondary metabolite gene induction: *N.* sp. (ISP2), *N. arthritidis* (rare 3) and *N. terpenica* (rare 3).

By utilisation of the DEREPLICATOR+ pipeline of GNPS, this chapter has identified previously known metabolites and associated the production of these metabolites to multiple *Nocardia* species. To date, the production of Peptidolipin NA and its analogue peptidolipin L-Val (6) has only been reported in *N. asteroides* (Guinand and Michel, 1966b). This chapter served to identify

four additional *Nocardia* species in which the peptidolipin family of secondary metabolites were detected by exact masses within the molecular network generated. A closer examination of the peptidolipin sub-network yielded no additional secondary metabolite hits via the GNPS spectral database. In addition, the sub-network was interrogated for spectral matches to five additional members of the peptidolipin (peptidolipin B-F) as well as three structurally related cyclic lipopeptides (surfactin, iturin and bacillomycin D) but yielded no further hits. It is speculated that this sub-network and its significant size has the potential to contain additional, yet undiscovered and uncharacterised bioactive analogues of peptidolipin.

Current literature states the production of the nocardimicins is credited to *N.* sp. TP-A0674 and *N. nova* JCM 6044 (Ikeda et al., 2005b, Ikeda et al., 2005a). This chapter detected nocardimicin production in *N. nova* in agreement with current literature, as well as seven additional *Nocardia* species. To date, brasilibactin A production has exclusively been associated with *N. brasiliensis* IFM 0995 (Tsuda et al., 2005). This chapter has also detected brasilibactin A production in *N. brasiliensis* in addition to five other distinct *Nocardia* sp. Furthermore, genomic interrogation of all nocardimicin and brasilibactin producing isolates indicates a nocobactin-like biosynthetic gene cluster present in all cases. As the chemical architecture of nocobactin is similar to that of nocardimicin and brasilibactin and considering nocobactin was not detected in any of the media extracts, it is likely that the biosynthesis of the nocardimicins and brasilibactins is associated with the nocobactin like gene cluster identified in this chapter.

Interestingly, the production of detoxin D1 and the xenoamicins has only been associated with bacterial species outside of the *Nocardia* (*Streptomyces caespitosus* var. *detoxicus* and *Xenorhabdus mauleonii* respectively) (Yonehara et al., 1968, Zhou et al., 2013). Here we report Detoxin D1 production in *N. terpenica* and xenoamicin production in a significant proportion

(70%) of the *Nocardia* cohort assessed within this chapter. As the detoxin D1 biosynthetic gene cluster has been previously defined in *Streptomyces*, antiSMASH interrogation was carried out on the genome of *N. terpenica* which was the only detoxin D1 producer in the *Nocardia* cohort assessed. This resulted in the identification of a hybrid PKS-NRPS biosynthetic gene cluster that is proposed to encode the biosynthetic machinery for detoxin D1 biosynthesis. This gene cluster had identical domain arrangement, number and type of amino acid monomers as the known detoxin D1 biosynthetic gene cluster. At present, this constitutes a highly likely correlation between the known detoxin D1 biosynthetic gene cluster and the gene cluster identified in *N. terpenica.* However, conclusive evidence for the production of detoxin D1 by the identified gene cluster in *N. terpenica* can only be confirmed by additional gene knockout experiments or heterologous expression which were beyond the parameters of this project.

Determining the biosynthetic origin of the xenoamicins in *Nocardia* was convoluted by the fact that there was no common gene cluster present in all producing isolates. Despite this, a NRPS locus was identified in a producing *N. terpenica* isolate that shared a similar domain organisation; 13 adenylation domains indicating an end product with 13 amino acid residues concordant with the xenoamicins.

The relatively unexplored metabolite repertoire of the *Nocardia* is reflected within the molecular network built in this research, highlighting the potential of this underexamined genera to yield novel, and potentially therapeutically valuable natural products. It is expected that reanalysis of the mass spectra contributing to the molecular network will yield further metabolite identification hits as spectral databases improve. However, the speed of such progress will depend upon the level of contributions from the natural product research community, including ongoing annotation and

dereplication of deposited spectral data. All data generated within this chapter was subsequently uploaded to the publicly available GNPS platform to aid future research endeavours.

Chapter 6: Final discussion

The discovery of antimicrobial secondary metabolites with varied activity and structural diversity from actinomycetes (particularly *Streptomyces*) has been the focus of intense research. The global onset of multidrug resistant bacteria coupled to the stark decline of novel antimicrobials from the once fruitful antibiotic discovery pipeline has caused a shift in focus towards other bacterial genera as new sources of antimicrobial agents with clinical significance.

This project was undertaken to investigate the potential of a group of underexplored pathogenic actinomycetes predominately from the genus *Nocardia.* Whole genome sequencing of over 700 pathogenic *Nocardia* sp. at the Peter Doherty Institute, University of Melbourne indicates that this genus possesses a large array of secondary metabolite biosynthetic gene clusters that rivals that of *Streptomyces*. Moreover, such gene clusters vary greatly amongst different *Nocardia* sp. highlighting the potential of this biosynthetically talented group of bacteria to yield new leads for antimicrobial metabolites. With these factors in mind, this study was formulated with three distinct goals:

(i) To empirically screen 169 pathogenic actinomycete isolates on 19 alternate media types for the production of antimicrobial metabolites that exhibit activity against a panel of prominent multidrug resistant bacteria.

(ii) To utilise whole genome sequencing to aid in the selection of a transcriptionally silent, secondary metabolite biosynthetic gene cluster of unknown potential in *Nocardia* and attempt to activate this gene cluster via promoter refactoring and homologous recombination techniques.

(iii) To utilise LC-MS/MS and bioinformatic software to evaluate the whole secondary metabolome of ten distinct *Nocardia* species grown on multiple media types.

None of the *Nocardia* isolates that were empirically screened within this study inhibited the growth of any of the five multidrug resistant bacterial pathogens. However, 61 *Nocardia* isolates were observed to inhibit the growth of *S. aureus* Newman and 30 exhibited inhibitory activity against *E. coli* DH10B. Compound extraction, LC-MS and dereplication analysis of *S. aureus* Newman and *E. coli* DH10B inhibiting metabolites was beyond the scope of this project. However, it should be emphasised that these metabolites could represent new leads as antibiotic scaffolds for future development.

Due to the cryptic nature of many secondary metabolite biosynthetic gene clusters, media type was a major consideration in the empirical screening of all actinomycete isolates. A total of 19 media types were assessed to determine if media composition facilitated the induction of secondary metabolism. It is recommended that future empirical screening efforts focused on the discovery of antimicrobial metabolites in *Nocardia* sp. be carried out on Bennett's media. This media type best facilitated the induction of antimicrobial secondary metabolite biosynthetic genes active against *S. aureus* Newman and *E. coli* DH10B in a range of strains. In contrast, this study also highlights media types that performed poorly in regard to facilitating antibiotic production in *Nocardia* (ISP1 and A3M) and should be avoided in future antibiotic screening endeavours focused on *Nocardia* species.

Strains from a total of 13 distinct *Nocardia* species were empirically screened. Isolates that produced the largest number of zones of inhibition on a range of different growth media in response to overlay with *S. aureus* Newman and *E. coli* DH10B included *N. cyriacigeorgica*, *N. farcinica* and *N. brasiliensis*. It is recommended that these species are prioritised for further bioprospecting efforts over the remaining cohort of *Nocardia* assessed in this study. Interestingly, *Nocardia* isolates from each individual species did not uniformly exhibit the same antibiotic production

characteristics in response to *S. aureus* Newman and *E. coli* DH10B overlay. This suggests that the genes responsible for the biosynthesis of antimicrobial secondary metabolites in the *Nocardia* varies greatly, even amongst members of the same species. This observation is complemented by bioinformatic interrogation of the genomes of *Nocardia* indicating a high degree of variability of secondary metabolite biosynthetic gene clusters amongst individual species.

The traditional, agar-based empirical screening methodology employed in this study can be limiting, predominately due to that fact that antimicrobial producing isolates are prioritised based upon visually observable zones of inhibition. Weakly active antimicrobials or highly active antimicrobials expressed at low concentrations cannot be visualised and therefore can be overlooked when employing this methodology. Furthermore, the overlay assays can be time consuming, challenging to interpret and rely upon good growth characteristics from every isolate screened. Despite these limitations, this technique has proved fruitful with the identification of two *Streptomyces* isolates which inhibited the growth of multidrug resistant *A. baumannii* and *E. coli*. Additional LC-MS/MS analysis was carried out on media extracts from each *Streptomyces* isolate and subjected to GNPS DEREPLICATOR+ interrogation which yielded no corresponding hits to any previously known metabolites within the database. This is highly suggestive that both superbug inhibiting metabolites could be novel and warrant further investigation.

This study utilised whole genome sequencing to select a secondary metabolite biosynthetic gene cluster of unknown potential that appeared to be unique to a single *Nocardia* isolate. Attempts to activate this transcriptionally silent gene cluster via promoter refactoring were not successful. The constitutive promoter selected (*PermE*) has been successfully utilised in the activation of secondary metabolite biosynthetic gene clusters in *Streptomyces* sp. (Olano et al., 2014) (Dangel et al., 2008). However, in this instance, despite confirmation of insertion at the desired site by

whole genome sequencing, *PermE* did not induce secondary metabolite production, suggesting that this promoter may not function in *Nocardia*. The reasons behind this are not clear, it is suspected that tight transcriptional regulation of secondary metabolite biosynthetic machinery could be a consideration. To date, little is known regarding gene regulation in the *Nocardia* and further work remains to understand this phenomenon. Bioactivity screening of mutant and wild type strains on Bennett's media against multidrug resistant *A. baumannii* yielded a zone of inhibition in both strains. As this zone was observed in both mutant and wild type strains, it is clear that antibiotic production was not due to promoter refactoring efforts. DEREPLICATOR+ interrogation of media extracts yielded no hits to any previously known antibiotics. This indicates that this anti-*Acinetobacter* antibiotic could be novel and warrants further investigation and characterisation.

To gain further insight into the biosynthetic capabilities of *Nocardia*, a molecular network was generated focusing on ten distinct species and highlighted the vast unknown molecular diversity of the *Nocardia*. This work identified four distinct molecular families of bioactive metabolites that contained spectra matching previously known metabolites in the curated GNPS library. Placing these identified metabolites into their corresponding sub-networks aided in assessing the probable metabolites within four sub-networks of the molecular network generated. Evaluation of media composition indicated that the abundance of secondary metabolites present within the molecular network varied greatly depending upon the media type selected. This highlights the importance of growth media selection to optimise secondary metabolite recovery. Overall, rare 3 media facilitated the highest number of spectra within the generated molecular network. This emphasises the usefulness of molecular networking as a tool to facilitate optimal media selection prior to bioactivity screening to maximise secondary metabolite output.

The molecular network generated in this study aided by the DEREPLICATOR+ platform of GNPS has shown the vast untapped chemical diversity of the *Nocardia* allowing greater insight into the biosynthetic capabilities of this group of bacteria. This untargeted approach has additionally highlighted the potential of *Nocardia* to produce secondary metabolites that are not just antibacterial in nature, with known anti-malarial, antifungal and cytotoxic metabolites detected in media extracts. To date, whole genome sequences from three of the ten *Nocardia* species assessed in this chapter have been uploaded to GenBank. Accession numbers for each isolate is available in appendix (1). The generation of spectral data within this study contributes to a greater understanding of the biosynthetic capabilities of *Nocardia* which to date, is distinctly lacking within the literature and spectral databases.

References

ABDELMOHSEN, U. R., et al. (2015) Elicitation of secondary metabolism in actinomycetes. *Biotechnology Advances,* **33** (6, Part 1), 798-811.

ACKERMANN, B. L., et al. (1996) Rapid analysis of antibiotic-containing mixtures from fermentation broths by using liquid chromatography-electrospray ionization-mass spectrometry and matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry. *J Am Soc Mass Spectrom,* **7** (12), 1227-37.

ADLI, M. (2018) The CRISPR tool kit for genome editing and beyond. *Nature communications,* **9** (1), 1911-1911.

ANAND, S., et al. (2010) SBSPKS: structure based sequence analysis of polyketide synthases. *Nucleic Acids Research,* **38** (suppl_2), W487-W496.

AURA (2017) Second Australian report on antimicrobial use

and resistance in human health. 215.

BABA, T., et al. (2008) Genome Sequence of Staphylococcus aureus Strain Newman and Comparative Analysis of Staphylococcal Genomes: Polymorphism and Evolution of Two Major Pathogenicity Islands. *Journal of Bacteriology,* **190** (1), 300.

BALTZ, R. H. (2010) Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters. *Journal of Industrial Microbiology & Biotechnology,* **37** (8), 759-772.

BALTZ, R. H. (2018) Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. *Journal of Industrial Microbiology & Biotechnology*.

BANKEVICH, A., et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology,* **19** (5), 455-477.

BARGER, S. R., et al. (2012) Imaging secondary metabolism of Streptomyces sp. Mg1 during cellular lysis and colony degradation of competing Bacillus subtilis. *Antonie van Leeuwenhoek,* **102** (3), 435-445.

BARKA, E. A., et al. (2016) Taxonomy, Physiology, and Natural Products of <span class="named-content genus-species" id="named-content-1">Actinobacteria</span&gt. *Microbiology and Molecular Biology Reviews,* **80** (1), 1.

BARTLETT, J. G., et al. (2013) Seven Ways to Preserve the Miracle of Antibiotics. *Clinical Infectious Diseases,* **56** (10), 1445-1450.

BEHAL, V. (2000) Bioactive products from Streptomyces. *Adv Appl Microbiol,* **47**, 113-56.

BERDY, J. (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo),* **58** (1), 1-26.

BEUTLER, J. A., et al. (1990) Dereplication of Phorbol Bioactives: Lyngbya majuscula and Croton cuneatus. *Journal of Natural Products,* **53** (4), 867-874.

BODE, H. B., et al. (2002) Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem,* **3** (7), 619-627.

BOECK, L. D., et al. (1984) N-demethylvancomycin, a novel antibiotic produced by a strain of Nocardia orientalis. Taxonomy and fermentation. *J Antibiot (Tokyo),* **37** (5), 446-53.

BROWN-ELLIOTT, B. A., et al. (2006) Clinical and Laboratory Features of the Nocardia spp. Based on Current Molecular Taxonomy. *Clinical Microbiology Reviews,* **19** (2), 259.

BUULTJENS, A. H., et al. (2017) Evolutionary origins of the emergent ST796 clone of vancomycin resistant Enterococcus faecium. *PeerJ,* **5**, e2916.

CAMPBELL, D. R., et al. (2007) Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and metal sites of metal-responsive repressors are not congruent. *J Biol Chem,* **282** (44), 32298-310.

CASTELLI, J. B., et al. (2011) Infectious endocarditis caused by Nocardia sp.: histological morphology as a guide for the specific diagnosis. *Braz J Infect Dis,* **15** (4), 384-6.

CDC (2018) *Timeline of Antibiotic Resistance Compared to Antibiotic Development.* Centers of Disease Control and Infection, [cited: 13 July]. Available from https://www.cdc.gov/drugresistance/about.html.

CDC (2019) *Extensively Drug-Resistant Tuberculosis (XDR TB).* [cited: 9th July]. Available from https://www.cdc.gov/tb/publications/factsheets/drtb/xdrtb.htm.

CHAN, Y. A., et al. (2009) Biosynthesis of polyketide synthase extender units. *Natural product reports,* **26** (1), 90-114.

CHEN, G., et al. (2000) Enhanced production of microbial metabolites in the presence of dimethyl sulfoxide. *J Antibiot (Tokyo),* **53** (10), 1145-53.

COATES, A. R. M., et al. (2011) Novel classes of antibiotics or more of the same? *British Journal of Pharmacology,* **163** (1), 184-194.

CORTES, J., et al. (1990) An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature,* **348** (6297), 176-8.

CORTINA, N. S., et al. (2012) Myxoprincomide: A Natural Product from Myxococcus xanthus Discovered by Comprehensive Analysis of the Secondary Metabolome. *Angewandte Chemie International Edition,* **51** (3), 811-816.

CRONAN, J. E. and THOMAS, J. (2009) Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods in enzymology,* **459**, 395-433.

CRUSEMANN, M., et al. (2017) Prioritizing Natural Product Diversity in a Collection of 146 Bacterial Strains Based on Growth and Extraction Protocols. *J Nat Prod,* **80** (3), 588-597.

DANGEL, V., et al. (2008) novE and novG act as positive regulators of novobiocin biosynthesis. *Archives of Microbiology,* **190** (5), 509-519.

DAVIES, J., et al. (2006) The world of subinhibitory antibiotic concentrations. *Current Opinion in Microbiology,* **9** (5), 445-453.

DE LIMA PROCÓPIO, R. E., et al. (2012) Antibiotics produced by Streptomyces. *The Brazilian Journal of Infectious Diseases,* **16** (5), 466-471.

DOMALAON, R., et al. (2018) Antibiotic Hybrids: the Next Generation of Agents and Adjuvants against Gram-Negative Pathogens? *Clinical Microbiology Reviews,* **31** (2), e00077-17.

DOULL, J. L., et al. (1994) Conditions for the production of jadomycin B by Streptomyces venezuelae ISP5230: effects of heat shock, ethanol treatment and phage infection. *J Ind Microbiol,* **13** (2), 120-5.

DURFEE, T., et al. (2008) The Complete Genome Sequence of Escherichia coli DH10B: Insights into the Biology of a Laboratory Workhorse. *Journal of Bacteriology,* **190** (7), 2597.

DUTHIE, E. S. and LORENZ, L. L. (1952) Staphylococcal coagulase; mode of action and antigenicity. *J Gen Microbiol,* **6** (1-2), 95-107.

EBATA, E., et al. (1975) *Lysocellin, a new polyether antibiotic. I. Isolation, purification, physico chemical and biological properties.* Vol. 28.

ELLNER, J. J. and BENNETT, J. E. (1976) Chronic meningitis. *Medicine (United States),* **55** (5), 341-369.

ENGELHARDT, K., et al. (2010) Production of a new thiopeptide antibiotic, TP-1161, by a marine Nocardiovsis species. *Applied and Environmental Microbiology,* **76** (15), 4969-4976.

FAIR, R. J. and TOR, Y. (2014) Antibiotics and Bacterial Resistance in the 21st Century. *Perspectives in Medicinal Chemistry,* **6**, 25-64.

FLEISCHMANN, R. D., et al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science,* **269** (5223), 496.

FLEMING, A. (1929) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bulletin of the World Health Organization,* **79** (8), 780-790.

FOLEY, T. L., et al. (2009) Phosphopantetheinyl transferase inhibition and secondary metabolism. **276** (23), 7134-7145.

GALM, U. and SHEN, B. (2006) Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis. *Expert Opin Drug Discov,* **1** (5), 409-37.

GAO, C.-H., et al. (2012) Characterization of a novel ArsR-like regulator encoded by Rv2034 in Mycobacterium tuberculosis. *PloS one,* **7** (4), e36255-e36255.

GAUDELLI, N. M., et al. (2015) beta-Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. *Nature,* **520** (7547), 383-7.

GOERING, A. W., et al. (2016) Metabologenomics: Correlation of Microbial Gene Clusters with Metabolites Drives Discovery of a Nonribosomal Peptide with an Unusual Amino Acid Monomer. *ACS Central Science,* **2** (2), 99-108.

GONZALEZ-LERGIER, J., et al. (2005) Theoretical considerations and computational analysis of the complexity in polyketide synthesis pathways. *J Am Chem Soc,* **127** (27), 9930-8.

GUINAND, M. and MICHEL, G. (1966a) Structure d'un peptidolipide isole de Nocardia asteroides, la peptidolipine NA. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism,* **125** (1), 75-91.

GUINAND, M. and MICHEL, G. (1966b) [Structure of a peptidolipid isolated from Nocardia asteroides, peptidolipin NA]. *Biochim Biophys Acta,* **125** (1), 75-91.

GUO, F., et al. (2015) Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metabolic Engineering,* **28**, 134-142.

HAFERBURG, G., et al. (2008) Arousing sleeping genes: shifts in secondary metabolism of metal tolerant actinobacteria under conditions of heavy metal stress. **22** (2), 225.

HARTWIG, A. (2001) Role of magnesium in genomic stability. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* **475** (1), 113-121.

HIGASHIDE, E., et al. (1977) Ansamitocin, a group of novel maytansinoid antibiotics with antitumour properties from Nocardia. *Nature,* **270** (5639), 721-722.

HO, S., et al. (1996) The mechanism of action of cyclosporin A and FK506. *Clin Immunol Immunopathol,* **80** (3 Pt 2), S40-5.

HOSHINO, Y., et al. (2011) Identification of nocobactin NA biosynthetic gene clusters in Nocardia farcinica. *Journal of bacteriology,* **193** (2), 441-448.

HUBERT, J., et al. (2015) *Dereplication strategies in natural product research: How many tools and methodologies behind the same concept?* Vol. 16.

IGARASHI, M., et al. (2000) Tubelactomicin A, a novel 16-membered lactone antibiotic, from Nocardia sp. I. Taxonomy, production, isolation and biological properties. *Journal of Antibiotics,* **53** (10), 1096-1101.

IGARASHI, Y., et al. (2010) Alchivemycin A, a bioactive polycyclic polyketide with an unprecedented skeleton from Streptomyces sp. *Org Lett,* **12** (15), 3402-5.

IKEDA, Y., et al. (2005a) Nocardimicins G, H and I, siderophores with muscarinic M3 receptor binding inhibitory activity from Nocardia nova JCM 6044. *J Antibiot (Tokyo),* **58** (9), 566-72.

IKEDA, Y., et al. (2005b) Nocardimicins A, B, C, D, E, and F, Siderophores with Muscarinic M3 Receptor Inhibiting Activity from Nocardia sp. TP-A0674. *Journal of Natural Products,* **68** (7), 1061-1065.

ISHINO, Y., et al. (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol,* **169** (12), 5429-33.

JANSEN, R., et al. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol,* **43** (6), 1565-75.

KIM, J. H., et al. (2010) Cloning large natural product gene clusters from the environment: piecing environmental DNA gene clusters back together with TAR. *Biopolymers,* **93** (9), 833-44.

KLAUS, M. and GRININGER, M. (2018) Engineering strategies for rational polyketide synthase design. *Nat Prod Rep,* **35** (10), 1070-1081.

KO, Y. H., et al. (1999) Chemical mechanism of ATP synthase. Magnesium plays a pivotal role in formation of the transition state where ATP is synthesized from ADP and inorganic phosphate. *J Biol Chem,* **274** (41), 28853-6.

KOMAKI, H., et al. (2014) Genome based analysis of type-I polyketide synthase and nonribosomal peptide synthetase gene clusters in seven strains of five representative Nocardia species. *BMC Genomics,* **15** (1), 323.

KOPP, J., et al. (2017) Impact of Glycerol as Carbon Source onto Specific Sugar and Inducer Uptake Rates and Inclusion Body Productivity in E. coli BL21(DE3). *Bioengineering (Basel, Switzerland),* **5** (1), 1.

LERMINIAUX, N. A. and CAMERON, A. D. S. (2018) Horizontal transfer of antibiotic resistance genes in clinical environments. *Canadian Journal of Microbiology,* **65** (1), 34-44.

LI, J., et al. (2009a) [Evaluation of the activities of two promoters in Streptomycetes by reporter gene method]. *Wei Sheng Wu Xue Bao,* **49** (11), 1454-8.

LI, M. H., et al. (2009b) Automated genome mining for natural products. *BMC Bioinformatics,* **10** (1), 185.

LI, S., et al. (2015) Genome-wide identification and evaluation of constitutive promoters in streptomycetes. *Microbial cell factories,* **14**, 172-172.

LI, W., et al. (2003) Nocathiacins, new thiazolyl peptide antibiotics from Nocardia sp. I. Taxonomy, fermentation and biological activities. *J Antibiot (Tokyo),* **56** (3), 226-31.

LIM, Y. H., et al. (2018) Auroramycin: A Potent Antibiotic from Streptomyces roseosporus by CRISPR-Cas9 Activation. *ChemBioChem,* **19** (16), 1716-1719.
LING, L. L., et al. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature,* **517**, 455.
LUO, Y., et al. (2015) Systematic Identification of a Panel of Strong Constitutive Promoters from Streptomyces albus. *ACS Synth Biol,* **4** (9), 1001-10.
MAGET-DANA, R., et al. (1985) Bacterial lipopeptides induce ion-conducting pores in planar bilayers. *Biochemical and Biophysical Research Communications,* **129** (3), 965-971.
MARCET-HOUBEN, M., et al. (2008) Phylogenetic analysis of homologous fatty acid synthase and polyketide synthase involved in aflatoxin biosynthesis. *Bioinformation,* **3** (1), 33-40.
MARSHALL, D. D. and POWERS, R. (2017) Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Progress in nuclear magnetic resonance spectroscopy,* **100**, 1-16.
MARTÍNEZ-BUENO, M., et al. (1990) A transferable plasmid associated with AS-48 production in Enterococcus faecalis. *Journal of Bacteriology,* **172** (5), 2817-2818.
MARTINEZ, J. L. and BAQUERO, F. (2000) Mutation Frequencies and Antibiotic Resistance. *Antimicrobial Agents and Chemotherapy,* **44** (7), 1771-1777.
MCCLURE, R. A., et al. (2016) Elucidating the Rimosamide-Detoxin Natural Product Families and Their Biosynthesis Using Metabolite/Gene Cluster Correlations. *ACS Chemical Biology,* **11** (12), 3452-3460.
MCLEAN, C. T., et al. (2019) Dissolution of the Disparate: Co-ordinate Regulation in Antibiotic Biosynthesis. *Antibiotics,* **8** (2).
MEDEMA, M. H., et al. (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research,* **39** (Web Server issue), W339-W346.
MENÉNDEZ, R., et al. (1997) Pulmonary infection with Nocardia species: A report of 10 cases and review. *European Respiratory Journal,* **10** (7), 1542-1546.
MERINO, N., et al. (2019) Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context. *Frontiers in microbiology,* **10**, 780-780.
MOHIMANI, H., et al. (2018) Dereplication of microbial metabolites through database search of mass spectra. *Nature Communications,* **9** (1), 4035.
MOJICA, F. J., et al. (2000) Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol,* **36** (1), 244-6.
MONTIEL, D., et al. (2015) Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters. *Proceedings of the National Academy of Sciences,* **112** (29), 8953.
MOORE, J. M., et al. (2012) Chapter Eighteen - Use and Discovery of Chemical Elicitors That Stimulate Biosynthetic Gene Clusters in Streptomyces Bacteria. In: HOPWOOD, D. A. (Ed.) *Methods in Enzymology*. Vol. 517. Academic Press, pp. 367-385.
MORGENSTERN, A., et al. (2015) Divalent transition-metal-ion stress induces prodigiosin biosynthesis in Streptomyces coelicolor M145: formation of coeligiosins. *Chemistry,* **21** (16), 6027-32.
MUROI, M., et al. (1980) Macbecins I and II, new antitumor antibiotics. II. Isolation and characterization. *J Antibiot (Tokyo),* **33** (2), 205-12.
NAH, H.-J., et al. (2017) Cloning and Heterologous Expression of a Large-sized Natural Product Biosynthetic Gene Cluster in Streptomyces Species. *Frontiers in Microbiology,* **8** (394).

NICHOLS, D., et al. (2010) Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl Environ Microbiol,* **76** (8), 2445-50.

OCHI, K. and HOSAKA, T. (2013) New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. *Applied microbiology and biotechnology,* **97** (1), 87-98.

OCHMAN, H., et al. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature,* **405** (6784), 299-304.

OKADA, B. K. and SEYEDSAYAMDOST, M. R. (2017) Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS Microbiol Rev,* **41** (1), 19-33.

OLANO, C., et al. (2014) Activation and identification of five clusters for secondary metabolites in Streptomyces albus J1074. *Microbial biotechnology,* **7** (3), 242-256.

ONAKA, H., et al. (2011) Mycolic Acid-Containing Bacteria Induce Natural-Product Biosynthesis in Streptomyces Species. *Applied and Environmental Microbiology,* **77** (2), 400.

ONAKA, H., et al. (2001) Goadsporin, a chemical substance which promotes secondary metabolism and morphogenesis in streptomycetes. I. Purification and characterization. *J Antibiot (Tokyo),* **54** (12), 1036-44.

OUTHRED, A. C., et al. (2011) Nocardia infections of the face and neck. *Current Infectious Disease Reports,* **13** (2), 132-140.

PARTRIDGE, S. R., et al. (2018) Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clinical Microbiology Reviews,* **31** (4), e00088-17.

PEAK, N., et al. (2007) Abundance of six tetracycline resistance genes in wastewater lagoons at cattle feedlots with different antibiotic use strategies. *Environmental Microbiology,* **9** (1), 143-151.

PETTIT, R. K. (2011) Small-molecule elicitation of microbial secondary metabolites. *Microbial biotechnology,* **4** (4), 471-478.

PIDOT, S. J., et al. (2014) Antibiotics from neglected bacterial sources. *Int J Med Microbiol,* **304** (1), 14-22.

PIDOT, S. J., et al. Biosynthesis and ether bridge formation in nargenicin macrolides. *Angewandte Chemie,* **0** (ja).

PIMENTEL-ELARDO, S. M., et al. (2010) Anti-parasitic compounds from Streptomyces sp. strains isolated from Mediterranean sponges. *Marine Drugs,* **8** (2), 373-380.

PISHCHANY, G., et al. (2018) Amycomicin is a potent and specific antibiotic discovered with a targeted interaction screen. *Proceedings of the National Academy of Sciences,* **115** (40), 10124.

POURCEL, C., et al. (2005) CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology,* **151** (Pt 3), 653-63.

PROJAN, S. J. (2003) Why is big Pharma getting out of antibacterial drug discovery? *Current Opinion in Microbiology,* **6** (5), 427-430.

RAFIQ, M., et al. (2019) Geochemistry and Bacterial Recovery from Hindu Kush Range Glacier and Their Potential for Metal Resistance and Antibiotic Production. *Geomicrobiology Journal,* **36** (4), 326-338.

RATLEDGE, C. and SNOW, G. A. (1974) Isolation and structure of nocobactin NA, a lipid-soluble iron-binding compound from Nocardia asteroides. *Biochemical Journal,* **139** (2), 407-413.

RAVEH, A., et al. (2013) Discovery of potent broad spectrum antivirals derived from marine actinobacteria. *PLoS ONE,* **8** (12).

READ, A. F. and WOODS, R. J. (2014) Antibiotic resistance management. *Evolution, Medicine, and Public Health,* **2014** (1), 147-147.

RENWICK, M. and MOSSIALOS, E. (2018) What are the economic barriers of antibiotic R&D and how can we overcome them? *Expert Opinion on Drug Discovery,* **13** (10), 889-892.

RICE, L. B. (2008) Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. *The Journal of Infectious Diseases,* **197** (8), 1079-1081.

ROCHFORT, S. (2005) Metabolomics Reviewed: A New "Omics" Platform Technology for Systems Biology and Implications for Natural Products Research. *Journal of Natural Products,* **68** (12), 1813-1820.

ROESCH, L. F., et al. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *Isme j,* **1** (4), 283-90.

RUIZ, B., et al. (2010) Production of microbial secondary metabolites: Regulation by the carbon source. *Critical Reviews in Microbiology,* **36** (2), 146-167.

RUTLEDGE, P. J. and CHALLIS, G. L. (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology,* **13**, 509.

SANDMANN, A., et al. (2004) Identification and analysis of the core biosynthetic machinery of tubulysin, a potent cytotoxin with potential anticancer activity. *Chemistry and Biology,* **11** (8), 1071-1079.

SANTIN, O. and MONCALIAN, G. (2018) Loading of malonyl-CoA onto tandem acyl carrier protein domains of polyunsaturated fatty acid synthases. *J Biol Chem,* **293** (32), 12491-12501.

SEEMANN, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* **30** (14), 2068-9.

SENGES, C. H. R., et al. (2018) The secreted metabolome of Streptomyces chartreusis and implications for bacterial chemistry. *Proceedings of the National Academy of Sciences,* **115** (10), 2490-2495.

SEYEDSAYAMDOST, M. R. (2014) High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proceedings of the National Academy of Sciences of the United States of America,* **111** (20), 7266-7271.

SHIGEMORI, H., et al. (1998) Brasilicardin A. A novel tricyclic metabolite with potent immunosuppressive activity from actinomycete Nocardia brasiliensis. *Journal of Organic Chemistry,* **63** (20), 6900-6904.

SHIMA, J., et al. (1996) Induction of actinorhodin production by rpsL (encoding ribosomal protein S12) mutations that confer streptomycin resistance in Streptomyces lividans and Streptomyces coelicolor A3(2). *J Bacteriol,* **178** (24), 7276-84.

SHWAB, E. K., et al. (2007) Histone deacetylase activity regulates chemical diversity in Aspergillus. *Eukaryot Cell,* **6** (9), 1656-64.

SILVER, L. L. (2011) Challenges of Antibacterial Discovery. *Clinical Microbiology Reviews,* **24** (1), 71.

SIMPKIN, V. L., et al. (2017) Incentivising innovation in antibiotic drug discovery and development: progress, challenges and next steps. *The Journal of antibiotics,* **70** (12), 1087-1096.
SMITH, D. L., et al. (2002) Animal antibiotic use has an early but important impact on the emergence of antibiotic resistance in human commensal bacteria. *Proceedings of the National Academy of Sciences of the United States of America,* **99** (9), 6434-6439.
SOHNG, J. K., et al. (2008) Production, isolation and biological activity of nargenicin from Nocardia sp. CS682. *Archives of Pharmacal Research,* **31** (10), 1339-1345.
STARCEVIC, A., et al. (2008) ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res,* **36** (21), 6882-92.
STRIEKER, M., et al. (2010) Nonribosomal peptide synthetases: structures and dynamics. *Current Opinion in Structural Biology,* **20** (2), 234-240.
SUN, Y., et al. (2002) 'Streptomyces nanchangensis', a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters. *Microbiology,* **148** (2), 361-371.
SUZUKI, S., et al. (1965) A NEW ANTIBIOTIC, POLYOXIN A. *J Antibiot (Tokyo),* **18**, 131.
TANAKA, Y., et al. (2010) Rare earth elements activate the secondary metabolite–biosynthetic gene clusters in Streptomyces coelicolor A3(2). *The Journal Of Antibiotics,* **63**, 477.
TONG, Y., et al. (2018) CRISPR/Cas-based genome engineering in natural product discovery. *Nat Prod Rep*.
TORRES, O. H., et al. (2000) Infection Caused by Nocardia farcinica: Case Report and Review. *European Journal of Clinical Microbiology and Infectious Diseases*, **19** (3), 205-212.
TORSVIK, V., et al. (1990) High diversity in DNA of soil bacteria. *Appl Environ Microbiol,* **56** (3), 782-7.
TORTORELLA, E., et al. (2018) Antibiotics from Deep-Sea Microorganisms: Current Discoveries and Perspectives. *Marine drugs,* **16** (10), 355.
TSUDA, M., et al. (2005) Brasilibactin A, a Cytotoxic Compound from Actinomycete Nocardia brasiliensis. *Journal of Natural Products,* **68** (3), 462-464.
VAN DER HEUL, H. U., et al. (2018) Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Nat Prod Rep,* **35** (6), 575-604.
VENTER, J. C., et al. (2015) The Sequence of the Human Genome. *Clin Chem,* **61** (9), 1207-8.
WADE, W. (2002) Unculturable bacteria--the uncharacterized organisms that cause oral infections. *Journal of the Royal Society of Medicine,* **95** (2), 81-83.
WAKSMAN, S. A. and HENRICI, A. T. (1943) The Nomenclature and Classification of the Actinomycetes. *Journal of bacteriology,* **46** (4), 337-341.
WANG, G., et al. (2001) Spontaneous mutations that confer antibiotic resistance in Helicobacter pylori. *Antimicrobial agents and chemotherapy,* **45** (3), 727-733.
WANG, M., et al. (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology,* **34**, 828.
WHO (2017) GLOBAL PRIORITY LIST OF ANTIBIOTIC-RESISTANT BACTERIA

TO GUIDE RESEARCH, DISCOVERY, AND DEVELOPMENT OF

NEW ANTIBIOTICS. 7.
WHO (2018) *Global tuberculosis report* Available from https://www.who.int/tb/publications/global_report/en/.

WOLF, F., et al. (2018) Characterization of the Actinonin Biosynthetic Gene Cluster. **19** (11), 1189-1195.
WU, K.-M., et al. (2009) Genome Sequencing and Comparative Analysis of Klebsiella pneumoniae NTUH-K2044, a Strain Causing Liver Abscess and Meningitis. *Journal of Bacteriology,* **191** (14), 4492.
WYCHE, T. P., et al. (2012) Peptidolipins B-F, antibacterial lipopeptides from an ascidian-derived Nocardia sp. *Journal of natural products,* **75** (4), 735-740.
XU, J., et al. (2002) A rifampicin resistance mutation in the rpoB gene confers ppGpp-independent antibiotic production in Streptomyces coelicolor A3(2). *Molecular Genetics and Genomics,* **268** (2), 179-189.
YONEHARA, H., et al. (1968) The detoxin complex, selective antagonists of blasticidin S. *J Antibiot (Tokyo),* **21** (5), 369-70.
ZHANG, M. M., et al. (2017) CRISPR-Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. *Nature chemical biology,* 10.1038/nchembio.2341.
ZHOU, Q., et al. (2013) Structure and biosynthesis of xenoamicins from entomopathogenic Xenorhabdus. *Chemistry,* **19** (49), 16772-9.
ZUR WIESCH, P. A., et al. (2011) Population biological principles of drug-resistance evolution in infectious diseases. *The Lancet Infectious Diseases,* **11** (3), 236-247.

Appendix 1. Actinomycete isolates utilised within this study. 16S rRNA sequencing identification (SQID) and Peter Doherty Institute reference number.

Actinomycetes: Empirical Screening	Reference Number
16S rRNA Sequencing Identification (SQID)	
Nocardia beijingensis	SP0003
Streptomyces thermocarboxydus	AUSMDU00041313
Streptomyces thermoviolaceus* subsp *thermoviolaceus	AUSMDU00041314
Nocardia nova	AUSMDU00018955
Nocardia beijingensis	AUSMDU00018956
Nocardia nova	AUSMDU00041173
Nocardia cyriacigeorgica	AUSMDU00041174
Nocardia brasiliensis	AUSMDU00018957
Nocardia veterana/elegans/africana	AUSMDU00018958
Nocardia veterana/elegans/africana	AUSMDU00018959
Nocardia cyriacigeorgica	AUSMDU00041175
Nocardia asiatica/abscessus/arthritidis	AUSMDU00018960
Nocardia nova	AUSMDU00041176
Nocardia nova	AUSMDU00041177
Nocardia paucivorans	AUSMDU00018961
Nocardia paucivorans	AUSMDU00041178
Nocardia paucivorans	AUSMDU00041179
Nocardia nova	AUSMDU00018962
Nocardia paucivorans	AUSMDU00041180
Nocardia asiatica/abscessus/arthritidis	AUSMDU00018963
Nocardia cyriacigeorgica	AUSMDU00041181
Nocardia veterana	AUSMDU00018964
Nocardia veterana	AUSMDU00041182
Nocardia cyriacigeorgica	AUSMDU00041183
Nocardia abscessus	AUSMDU00041184
Nocardia cyriacigeorgica	AUSMDU00041185
Nocardia paucivorans	AUSMDU00041186
Nocardia cyriacigeorgica	AUSMDU00041187
Nocardia veterana/africana/elegans	AUSMDU00041188
Nocardia cyriacigeorgica	AUSMDU00041189
Nocardia nova	AUSMDU00041190
Nocardia veterana	AUSMDU00041191
Nocardia nova	AUSMDU00041192
Nocardia beijingensis	AUSMDU00018966
Nocardia cyriacigeorgica	AUSMDU00018967

Nocardia paucivorans	AUSMDU00018968
Nocardia nova	AUSMDU00041193
Nocardia paucivorans	AUSMDU00041194
Nocardia nova	AUSMDU00018969
Nocardia nova	AUSMDU00018970
Nocardia cyriacigeorgica	AUSMDU00018971
Nocardia cyriacigeorgica	AUSMDU00041198
Nocardia nova	AUSMDU00018996
Nocardia pseudobrasiliensis	AUSMDU00018992
Nocardia brasiliensis	AUSMDU00018998
Nocardia abscessus/arthritidis/asiatica	AUSMDU00018995
Nocardia veterana/africana/elegans	AUSMDU00019061
Nocardia nova	AUSMDU00041200
Nocardia nova	AUSMDU00041202
Nocardia veterana/africana/elegans	AUSMDU00019012
Dietzia cinnamea **or closely related**	AUSMDU00041306
Nocardia cyriacigeorgica	AUSMDU00019001
Nocardia cyriacigeorgica	AUSMDU00041203
Nocardia cyriacigeorgica	AUSMDU00041204
Nocardia veterana/elegans/africana	AUSMDU00019051
Nocardia cyriacigeorgica	AUSMDU00041205
Nocardia cyriacigeorgica	AUSMDU00041206
Nocardia cyriacigeorgica	AUSMDU00041208
Nocardia cyriacigeorgica	AUSMDU00019019
Nocardia nova	AUSMDU00041209
Nocardia beijingensis	AUSMDU00041210
Nocardia cyriacigeorgica	AUSMDU00041211
Nocardia cyriacigeorgica	AUSMDU00041212
Nocardia cyriacigeorgica	AUSMDU00041213
Nocardia nova	AUSMDU00041214
Nocardia veterana/elegans/africana	AUSMDU00041216
Nocardia cyriacigeorgica	AUSMDU00041217
Nocardia nova	AUSMDU00041218
Nocardia paucivorans	AUSMDU00041219
Nocardia nova	AUSMDU00041220
Nocardia nova	AUSMDU00041221
Nocardia beijingensis	AUSMDU00041222
Nocardia nova	AUSMDU00041223
Nocardia brasiliensis	AUSMDU00018987
Nocardia farcinica	AUSMDU00041224
Nocardia nova	AUSMDU00041225

Nocardia cyriacigeorgica	AUSMDU00041226
Nocardia cyriacigeorgica	AUSMDU00041227
Nocardia farcinica	AUSMDU00041228
Nocardia nova	AUSMDU00041229
Nocardia nova	AUSMDU00041230
Nocardia cyriacigeorgica	AUSMDU00041231
Nocardia nova	AUSMDU00041232
Nocardia nova	AUSMDU00041233
Nocardia cyriacigeorgica	AUSMDU00019004
Nocardia cyriacigeorgica	AUSMDU00019005
Nocardia veterana/elegans/africana	AUSMDU00019010
Nocardia cyriacigeorgica	AUSMDU00041234
Nocardia cyriacigeorgica	AUSMDU00019023
Nocardia beijingensis	AUSMDU00041235
Nocardia farcinica	AUSMDU00041236
***Dietzia* sp.**	AUSMDU00041308
Nocardia nova	AUSMDU00041237
Nocardia nova	AUSMDU00041238
Nocardia paucivorans	AUSMDU00041239
Nocardia paucivorans	AUSMDU00041240
Nocardia farcinica	AUSMDU00041242
Nocardia paucivorans	AUSMDU00041243
Nocardia otitidiscaviarum	AUSMDU00041244
Nocardia cyriacigeorgica	AUSMDU00041245
Nocardia veterana/elegans	AUSMDU00019002
Nocardia beijingensis	AUSMDU00041246
Nocardia paucivorans	AUSMDU00041247
Nocardia brasiliensis	AUSMDU00041248
Nocardia farcinica	AUSMDU00041249
Nocardia beijingensis	AUSMDU00041250
Nocardia cyriacigeorgica	AUSMDU00041252
Nocardia cyriacigeorgica	AUSMDU00041253
Nocardia veterana/elegans	AUSMDU00041254
Nocardia farcinica	AUSMDU00041255
Nocardia nova	AUSMDU00041256
Nocardia nova	AUSMDU00041257
Streptomyces thermocarboxydus	AUSMDU00041315
Nocardia paucivorans	AUSMDU00041258
Nocardia paucivorans	AUSMDU00041259
Nocardia cyriacigeorgica	AUSMDU00041260
Nocardia brasiliensis	AUSMDU00041261

Nocardia nova	AUSMDU00041262
Nocardia veterana/elegans/africana	AUSMDU00019007
Nocardia farcinica	AUSMDU00041263
Nocardia brasiliensis	AUSMDU00041264
Nocardia brasiliensis	AUSMDU00041265
Nocardia cyriacigeorgica	AUSMDU00041267
Nocardia puris	AUSMDU00041269
***Dietzia* sp.**	AUSMDU00041309
Nocardia veterana/elegans/africana	AUSMDU00041270
Nocardia paucivorans	AUSMDU00041271
Auritidibacter ignavus	AUSMDU00041307
Nocardia brasiliensis	AUSMDU00041272
Nocardia veterana/elegans/africana	AUSMDU00019009
Nocardia veterana/elegans/africana	AUSMDU00019013
Nocardia nova	AUSMDU00041273
Streptomyces albus/rangoonensis/gibsonii	AUSMDU00041317
Nocardia cyriacigeorgica	AUSMDU00041274
Nocardia paucivorans	AUSMDU00041275
Nocardia nova	AUSMDU00041276
Nocardia cyriacigeorgica	AUSMDU00041277
Nocardia brasiliensis	AUSMDU00041279
Nocardia nova	AUSMDU00041280
Streptomyces thermocarboxydus	AUSMDU00041319
Nocardia cyriacigeorgica	AUSMDU00041281
Nocardia farcinica	AUSMDU00041282
Nocardia nova	AUSMDU00041283
Nocardia nova	AUSMDU00041284
Nocardia cyriacigeorgica	AUSMDU00041285
Rhodococcus coprophilus	AUSMDU00041305
Nocardia farcinica	AUSMDU00041286
Nocardia species	AUSMDU00019003
Nocardia nova	AUSMDU00041287
Nocardia brasiliensis	AUSMDU00041288
Nocardia brasiliensis	AUSMDU00041289
Nocardia farcinica	AUSMDU00041290
***Nocardia* sp.**	AUSMDU00041291
Nocardia farcinica	AUSMDU00041293
Nocardia cyriacigeorgica	AUSMDU00041294
Nocardia farcinica	AUSMDU00041295
Nocardia veterana	AUSMDU00019011
Nocardia abscessus	AUSMDU00041296

Streptomyces thermoviolaceus	AUSMDU00041320
Nocardia nova	AUSMDU00041297
Nocardia cyriacigeorgica	AUSMDU00041298
Nocardia vaccinii	AUSMDU00018999
Nocardia nova	AUSMDU00041299
Nocardia nova	AUSMDU00041300
Nocardia species	AUSMDU00019016
Streptomyces cacaoi	AUSMDU00041311
Nocardia arthritidis	AUSMDU00018978
Nocardia farcinica	AUSMDU00041302
***Streptomyces* sp.**	AUSMDU00041318
Positive controls: Empirical screening	
***Streptomyces* sp.**	AUSMDU00041328
Nocardia arthritidis	AUSMDU00012717
Nocardia terpenica	AUSMDU00012715
Nocardia: Molecular network	
***Nocardia* sp.**	AUSMDU00041322
Nocardia abscessus	AUSMDU00041323
Nocardia pneumoniae	AUSMDU00041325
Nocardia beijingensis	AUSMDU00041326
Nocardia nova	AUSMDU00041327
Nocardia terpenica	AUSMDU00012715 GenBank Accession Number: CP046173.1
Nocardia arthritidis	AUSMDU00012717 GenBank Accession Number: CP046172.1
Nocardia brasiliensis	AUSMDU00024985 GenBank Accession Number: CP046171.1
Nocardia cerradoensis	AUSMDU00041324
Nocardia ignorata	AUSMDU00041321
Nocardia: Promoter refactoring	
***Nocardia* sp.**	AUSMDU00041268

Appendix 2. Media composition

2.1 Media composition: Empirical screening

Media	Composition (g/L)
A3M	Soluble starch 20.0g, Glucose 5.0g, Glycerol 20.0g, Pharmamedia (cottonseed flour) 15.0g, Yeast extract 3.0g, Diaion HP-20 10.0g, Agar 15.0g pH to 7.0
AMM	Starch 10.0g, Yeast extract 4.0g, Peptone 2.0g, Agar 18.0g
Asukamycin	Glucose 20.0g, Peptone 5.0g, K_2HPO_4 0.25g, $MgSO_4$ 0.25g, Agar 15.0g, Hopwood trace elements solution (20 ml)
ATCC-2	Yeast extract 5.0g, Beef extract 3.0g, Peptone 5.0g, Dextrose 1.0g, Potato starch 2.0g, $CaCO_3$ 1.0g, NZamine E 5.0g, Agar 15.0g
Bennett's	Yeast extract 1.0g, Meat extract 1.0g, NZ amine (casein digest from milk) 2.0g, Glucose 10.0g, Agar 15.0g pH to 7.2
Bennett's Modified	Yeast extract 1.0 g, Meat extract 0.8g Bacto-casitone 2.0g, Glycerol 10.0g, Agar 15.0g
GYEA	Glucose 10.0g, Yeast extract 10.0g, Agar 15.0g
ISP1	Pancreatic Digest of Casein 5.0g, Yeast Extract 3.0g, Agar 15.0g
ISP2	Yeast Extract 4.0g, Malt Extract 10.0g, Dextrose 4.0g, Agar 20.0g
ISP5	L-Asparagine 1.0g, Dipotassium phosphate 1.0g, *Trace salt solution (1.0ml), Agar 20.0g ***Trace salt solution (100 ml)** Ferrous sulphate heptahydrate 0.1g Manganese chloride tetrahydrate 0.1g Zinc sulphate heptahydrate 0.1g Final pH @ 25°C 7.4 ± 0.2
M1	Starch 1.0g, Yeast extract 0.4g, Peptone 0.2g, Agar 18.0g

M8	Starch 20.0g, Glucose 10.0g, Hydrolysed casein 4.0g, Yeast extract 2.0g, Meat extract 2.0g, $CaCO_3$ 3.0g, Agar 15.0g
Minimal	L-asparagine 0.5g, K2HPO4 0.5g, $MgSO_4.2H_2O$ 0.2g, $FeSO_4.7H_2O$ 0.01g, Glucose 10.0g, Agar 10.0g
Modified YEME	Yeast extract 3.0g, Difco bacto-peptone 5.0g, Malt extract 3.0g, Glucose 10.0g, Sucrose 340.0g, Agar 15.0g
MYM	Maltose 4.0g, Yeast extract 4.0g, Malt extract 4.0g, Bacto agar 18.0g
R5M	**Part A** Maltose 100.0g, $MgCL_2.6H_2O$ 10.12g, K_2SO_4 0.5g, Casamino acids 0.2g, Yeast extract 10.0g, TES 11.46g, Agar 20.0g **Part B** Trace element solution (4ml) K_2PO_4 (0.5% solution) (10ml) 5M $CaCl_2.2H_2O$ (4ml) 20% L-proline (15ml) 1M NaOH (7ml) **Trace element solution (1 L)** $ZnCl_2$: 40mg $FeCl_3.6H_2O$: 200mg $CuCl2.2H_2O$: 10mg $MnCl_2.4H_2O$: 10mg $Na_2B_4O_7.10H_2O$: 10mg $(NH_4)_6Mo_7O_{24}.4H_2O$: 10mg
RARE 3	Glucose 10.0g, Yeast extract 4.0g, Malt extract 10.0g, Bacto peptone 2.0g, $MgCl_2 \cdot H_2O$ 2.0g, Glycerol 5.0g, Agar 15.0g
SM3	Glucose 10.0g, Peptone 5.0g, Tryptone 3.0g, Agar 15.0g
V-22	Soluble starch 10.0g, NZ-case (casein digest from milk) 3.0g, Yeast extract 2.0g, Tryptone 5.0g, K_2HPO_4 1.0g, $MgSO_4.7H_2O$ 0.5g, $CaCO_3$ 3.0g, Glucose 3.0g, Agar 15.0g

Appendix 2.2 Media composition: Transformation of ultra-competent *Escherichia coli*

Media	**Composition (g/L)**
Super Optimal Agar (SOA)	Yeast extract 5.0g, Tryptone 20g, 0.58g NaCl, 0.19g KCL, 2.4g $MgSO_4$, Agar 15g
Super Optimal broth with Catabolite repression (SOC)	Yeast extract 5g, Tryptone 20g, 0.58g NaCl, 0.19g KCL, 2.4g $MgSO_4$, 3.6g Glucose

Appendix 2.3 Media composition: Molecular networking.

Media	**Composition (g/L)**
Bennett's	Yeast extract 1.0g, Meat extract 1.0g, NZ amine (casein digest from milk) 2.0g, Glucose 10.0g, Agar 15.0g
ISP2	Yeast Extract 4.0g, Malt Extract 10.0g, Dextrose 4.0g, Agar 20.0g
Minimal	L-asparagine 0.5g, K_2HPO_4 0.5g, $MgSO_4.2H_2O$ 0.2g, $FeSO_4$.7H2O 0.01g, Glucose 10.0g, Agar 10.0g
M1	Starch 1.0g, Yeast extract 0.4g, Peptone 0.2g, Agar 18.0g
Rare 3	Glucose 10.0g, Yeast extract 4.0g, Malt extract 10.0g, Bacto peptone 2.0g, $MgCl_2 \cdot H_2O$ 2.0g, Glycerol 5.0g, Agar 15.0g

Appendix 3. Multidrug resistant isolates utilised in this study. Highlighting strain designation and resistance profiles.

Isolate Name	**Strain Designation**	**Resistance Profile**
Escherichia coli	ST131	ampicillin, ciprofloxacin ceftriaxone ceftazidime, gentamycin and tobramycin
Klebsiella pneumoniae **(Wu et al., 2009)**	NTUH-K2044	fosfomycin, cephalosporin, monobactams, carbapenems, fluoroquinolones aminoglycosides and glycopeptides
Acinetobacter baumannii	BAA-1710	gentamycin and tetracycline
Enterococcus faecium **(Buultjens et al., 2017)**	ST796	vancomycin, trimethoprim, tetracycline, macrolides and aminoglycosides
Staphylococcus aureus	ST239	spectinomycin, streptomycin, erythromycin, fosfomycin, bleomycin, tigecycline, tetracycline, aminoglycosides