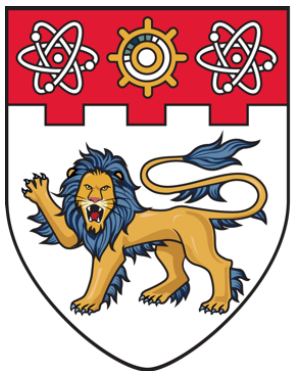


NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

A Genome-Guided Approach to the Study of Novel Secondary Metabolites in Quarry-Derived Actinomycetes

DARYN TAN FU ERN
G1802835K

SCHOOL OF BIOLOGICAL SCIENCES
2019

A Genome-Guided Approach to the Study of Novel Secondary Metabolites in Quarry-Derived Actinomycetes

DARYN TAN FU ERN
G1802835K

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the Masters of Science in Biological Sciences
2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

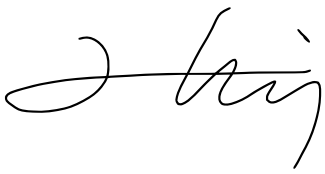
21 August 2019

. . . . . .
Date

. . . . . .
Daryn Tan Fu Ern

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.



21 August 2019

. . . . . .
Date

. . . . . .
Assoc Prof Liang Zhao Xun

Authorship Attribution Statement

Please select one of the following; *delete as appropriate:

*(A) This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

~~*(B) This thesis contains material from [x number] paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.~~

21 August 2019

. . . . . .

Date

. . . . . .

Daryn Tan Fu Ern

Acknowledgements

My journey pursuing the Masters of Science Degree has been an eventful and fun one. This could not have been possible without the constant support and encouragement from my family, friends, colleagues, supervisors and mentors. I would like to extend my sincere gratitude to the people who contributed to this thesis in one way or another.

Firstly, I would like to thank my supervisor Assoc Prof Liang Zhao Xun for giving me an opportunity to partake in such exciting work since my FYP and allowing me to continue my work for my M.Sc. He has always invoked me to think more critically as a scientist. My learning would not have been so fruitful without the freedom provided by him. His macro-management style inculcated an independent learning habit which drove me to develop problem-solving skills in the midst of exploring new ideas. I am grateful for all the wisdom and guidance from him that have helped me tremendously deal in this project.

I would also like to thank the following friends and colleagues for their wonderful company and assistance in the laboratory. Firstly, I would like to thank Hartono Candra, both PhDs to be, for his mentorship, critical insights and advice in this project; Dr. Howard Saw for his assistance, especially in the last stretch of experiments; Dr. Xin Lingyi, Dr. Ma Guang Lei, Dr. Chong Zhi Soon, Yeo Yee Phan, Ye Ruijuan, Daniel Wong, Melvin Lim, Rachel Andrea Chea, Hoa Tran for their support, guidance and assistance through this journey. Lastly, I would like to extend my gratitude to SBS office, for all administrative support and advice throughout my four years as an Integrated Programme student.

Table of Contents

ABSTRACT

The recent revival of microbial natural product discovery is motivated by the discovery of cryptic biosynthetic gene clusters (BGCs) with developments in genome sequencing technology. A genome-guided approach was employed in this study to discover natural products produced by the actinomycetes strain *Streptomyces* sp. P46 isolated from the Pulau Ubin Quarry. Genome sequencing and mining have revealed at least 33 BGCs, of which 18 were considered to be cryptic. Two methods were concurrently used to activate these cryptic BGCs – altering the fermentation media conditions (i.e., one strain many compounds or OSMAC approach) and overexpressing local regulator genes. A novel compound (named as Tasikamide A) with anticancer activity was isolated and its chemical structure was determined. Tasikamide A is putatively produced by BGC 20, a non-ribosomal peptide synthetase (NRPS)-polyketide synthase (PKS) hybrid cluster. Apart from Tasikamide A, several other potentially novel compounds have also been detected through the OSMAC approach and are in the process of being isolated and characterised.

In addition to the OSMAC approach, a genetic method was employed to target novel NRPS-PKS hybrid clusters BGCs 12, 14 and 28 to discover novel natural products. An overexpression plasmid was successfully used to express locally activating regulators to generate two overexpression mutant strains (P46 C12-ACT and P46 C28-ACT). The two mutant strains were sub-cultured in several culture media and the organic extract of the culture broth and biomass were subjected to metabolite profiling. Two compounds (**a** and **b**) were consistently produced by P46 C28-ACT and were identified to be of interest due to their unique retention time and UV-Vis spectra. While compound **a** is likely to be furaquinocin A, a known compound, the identity of compound **b** remains to be established. Unexpectedly, several compounds produced by the wild type P46 strain were not produced by P46 C12-ACT, suggesting that the overexpressed regulator could have been a repressor. Overall, my studies demonstrated that P46 is a biosynthetically talented *Streptomyces* strain in producing novel natural products and set the stage for the activation of cryptic BGCs using the genetic tools validated in this study.

CHAPTER 1: INTRODUCTION

1.1 Natural product discovery

Natural products from microorganisms have been historically proven as the most abundant source of compounds for drug discovery (Trivella & de Felicio, 2018). Natural products are usually derived from bacterial and fungal sources, and they contribute to more than 50% of all pharmaceutical drugs that are currently present in the market (Newman & Cragg, 2016). However, two decades ago, the pharmaceutical industry shifted to alternative means of drug discovery due to the problems posed by natural product discovery. The rediscovery of known compounds and low throughput methodologies have hindered drug discovery such that this approach was no longer economical for the industry (Baltz, 2017).

Hence, alternative techniques such as combinatorial chemistry and fragment-based drug design have been used to pursue new drugs from synthetic chemistry. Pharmaceutical companies utilise combinatorial chemistry to generate libraries comprising millions of synthetic compounds which will undergo high-throughput screening (HTS) in order to identify drug leads. An example of the method's success would be the synthesis of the antibiotic linezolid, patented as Zyvox by Pfizer. It was built off the foundation molecules oxazolidinones (Slee et al., 1987), developed by DuPont Pharmaceuticals in the 1980s. Oxazolidinones are antibiotic agents which possess a unique mechanism of inhibiting bacterial protein synthesis. The advantages of combinatorial chemistry allowed compounds to be designed in a systematic manner, continually improving from previous drug leads. HTS adds to this efficiency through which a larger number of compounds can be screened. Compared to the screening of natural products, pharmaceutical companies can avoid ambiguity in the protection of their intellectual property rights over their discovery when screening self-generated synthetic compounds (Gupta et al., 2017). However, Henkel et al. (1999) proved through statistical analysis that synthetic compounds had lacked structural complexity that was associated with bioactive natural products such as multiple chiral centres, heterocyclic substituents, and polycyclic structures.

Fragment-based drug design (FBDD) approaches drug discovery from a different angle. Fragments are defined as compounds possessing less than 20 non-hydrogen

atoms and FBDD focuses on these smaller compounds (Erlanson et al., 2016). Instead of screening a large number of libraries in search of drug leads, narrowing the focus to fragments significantly reduces the possible number of compounds which undergoes HTS. Moreover, applying the concept of 'Molecular Complexity' (Hann et al., 2001), smaller fragments which form fewer unfavourable interactions with neighbouring groups are hypothesised to bind to a greater number of sites on a larger number of proteins, resulting in more successful drug leads. Although theoretically, FBDD showed promising potential, more recent studies have highlighted that drug leads discovered were of poor quality and fragments did not possess suitable physicochemical properties such as solubility and aggregation.

The pharmaceutical industry is facing a reduction in the productivity and an increased economic barriers to entry of antibiotic discovery programs. Coupled with the imminent threat of multidrug resistance and resistant fungi, it has signalled a vital need to efficiently search for novel natural products (Nikaido, 2009) possessing therapeutic properties to serve as alternatives.

In this context, whole-genome sequencing has resulted in a major paradigm shift in approaching drug discovery today. Advanced genome sequencing technologies have allowed for successful genome mining of prokaryotes in search of novel bioactive natural products (Bachmann et al., 2014). Through a deeper understanding of the genetics and enzymology responsible for secondary metabolite biosynthesis, gene clusters encoding secondary metabolite pathways can be more effectively identified and analysed. A biosynthetic gene cluster can be identified as the core unit in charge of synthesising secondary metabolites, coding for biosynthetic enzymes, resistance determinants, transporters and regulatory elements (Nguyen et al., 2012). However, a vast majority of these biosynthetic gene clusters (BGCs) are poorly expressed in lab environments, and hence referred to as 'cryptic' or silent. Genome sequencing technologies and bioinformatics platforms have allowed the thorough coverage of all secondary metabolite genes, even those with cryptic expression. Through genome mining, cryptic BGCs can be targeted in search of novel natural products, significantly reducing the prospect of rediscovering known compounds.

Therefore, activating these silent and cryptic BGCs have been the focus of my research. Through closer analysis of the genomic sequences and understanding various mechanisms that regulate secondary metabolism in *Streptomyces*, I will be guided by a genome-based approach in elucidating possibly novel natural products from our quarry-derived *Streptomyces* strains.

1.2 Actinomycetes as a rich source of microbial natural products

Known for their ability to produce a wide diversity of natural products, actinomycetes have remained an important source for the discovery of novel antibiotics (Genilloud, 2017). Actinomycetes are defined as a metabolically active group of slow-growing, filamentous, gram-positive bacteria with high GC content of more than 70% (Tong et al., 2015). They grow through a combination of tip extension and branching of the hyphae and reproduce by sporulation.

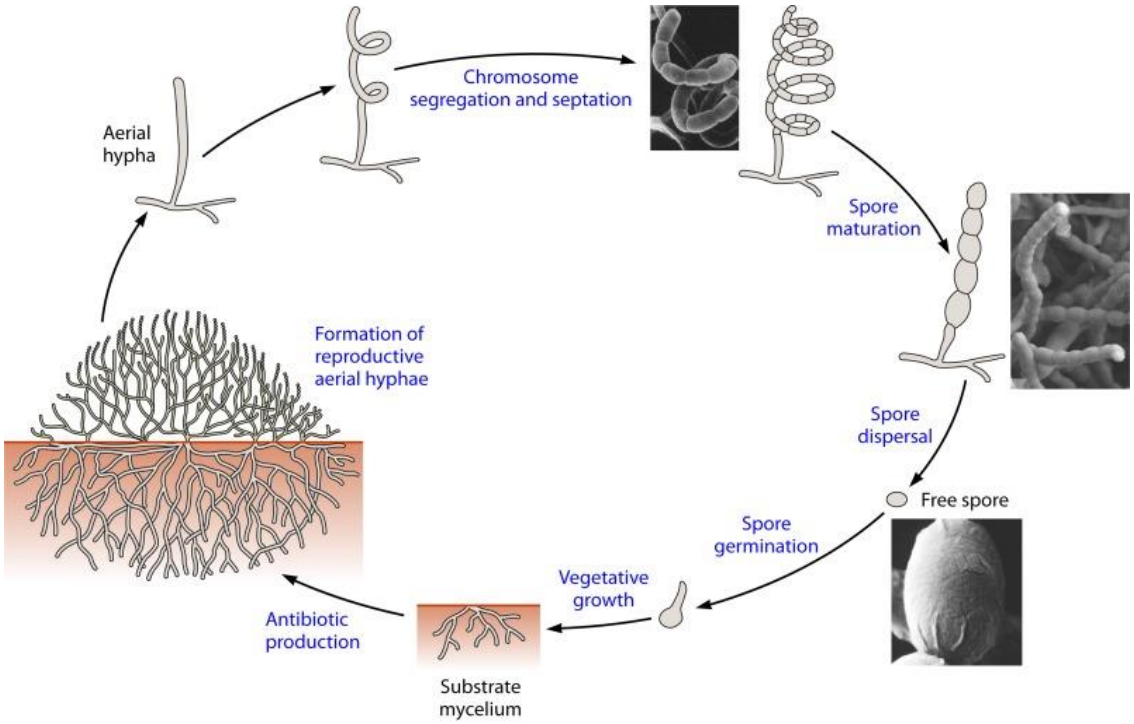


Figure 1. Schematic representation of the life cycle of sporulating actinomycetes adapted from (Barka et al., 2016). Actinomycetes reproduce through spores where spores can germinate and form a germ tube. The germ tube undergoes vegetative growth by first producing the primary or substrate mycelium which develops inside the media. Subsequently, it produces the secondary or aerial mycelium which grows above the substratum. When the vegetative hyphae maturates, a chain of spores are formed through septation. Spores are released when the vegetative hyphae swells and undergoes fragmentation.

Physiologically and ecologically, most actinomycetes are aerobic. Although they can be either heterotrophic or chemoautotrophic, most exists as chemoheterotrophic bacteria with the ability to harvest a wide variety of nutritional sources such as

various complex polysaccharides (Lechevalier & Lechevalier, 1965). Actinomycetes inhabit a diverse range of environments with species such as *Streptomyces*, *Micromonospora*, *Rhodococcus*, and *Salinispora* thriving in soil or aquatic environments, *Frankia* spp. a plant symbiont, *Corynebacterium*, *Mycobacterium*, or *Nocardia*species a plant or animal pathogen and *Bifidobacterium* spp. as gastrointestinal commensals.

A majority of the actinomycetes are saprophytic, soil-dwelling organisms (Mayfield et al., 1972), in particular, the *Streptomyces* species (which is used in this paper). *Streptomyces* can be found both on the soil surface as well as at depths of more than 2 m below ground (Goodfellow & Williams, 1983) and they account for over 95% (Williams et al., 1993) of actinomycetes strains isolated from soil. For optimum growth of *Streptomyces*, factors such as temperature, pH and soil moisture must be considered. Most *Streptomyces* are mesophilic, growing optimally at temperatures between 25 to 30°C while thermophilic species can grow at higher temperatures between 50 to 60°C. Low humidity and neutral pH also favours the vegetative growth of *Streptomyces* (Edwards, 1993). Both marine and terrestrial actinomycetes have been intensively explored and have contributed to the continual discovery of bioactive compounds (Monciardini et al., 2014).

Two-thirds of all antibiotics were of biological origin or were the semi-synthetic derivatives of natural products. Actinomycetes, specifically *Streptomyces*, are one of the most promising sources of natural products (Jakubiec-Krzesniak et al., 2004). Secondary metabolites are a category of natural products that have been of particular interest (Hanson, 2003). These secondary metabolites are not essential for growth and survival of the organism in nutrient-rich conditions, unlike the products of primary metabolism. However, they are regarded to be important during stress and nutrient-deprived conditions and considered to provide organisms with a competitive edge over others in the natural environment. It has been reported that higher production of secondary metabolites is associated with growth and development of microorganisms (Calvo et al., 2002). For example, in *Streptomyces*, secondary metabolite synthesis correlates with aerial hyphae formation (Bibb, 2005). With a diverse secondary metabolism, these bacteria acquire a huge amount of compounds which can serve as leads for the development of antibiotics, antivirals,

immunosuppressants, antifungals, insecticides, and antitumorals (Paulus et al., 2017).

Developments in genome sequencing technology has made tapping into *Streptomyces* as a promising source of natural products viable. Novel bioactive compounds discovered through genome mining approaches of actinomycete strains include: stambomycin, a large macrolactone from *Streptomyces ambofaciens* with cytotoxic activity (Laureti et al., 2011); the siderophore coelichelin from *Streptomyces coelicolor* (Lautru et al., 2005); orfamide, a cyclic lipopeptide with antifungal activity produced by *Pseudomonas fluorescens* (Gross et al., 2007); and the morphogenic peptide catenulipeptin from *Catenulispora acidiphila* (Huan Wang & van der Donk, 2012).

Recognising the potential of *Streptomyces* strains in producing bioactive compounds, we have isolated and sequenced *Streptomyces* strains obtained from soil samples in the Pulau Ubin Quarry to carry out a genome-based approach in search of novel natural compounds.

1.3 Classes of secondary metabolites

Secondary metabolites are classified into several major categories based on their structure, function and biosynthetic origins – polyketides and fatty acid-derived substances, ribosomal or non-ribosomal peptides, alkaloids and terpenoids (Table S3).

Polyketides are naturally produced by bacteria, fungi and plants, and are synthesised from repeated condensation reactions linking small carbon precursors (usually acyl CoA and malonyl CoA) by polyketide synthases (PKS). Due to the different methods of synthesis of the fatty acids, and post-PKS modifications, polyketides are extremely structurally diverse (Pfeifer & Khosla, 2001). Polyketides are known to possess specific therapeutic characteristics such as being anticancer, antifungal, and anticholesteremic agents; antibiotics; parasiticides and immunomodulators. Some well-known pharmaceutical polyketides include the antibiotics tetracycline and erythromycin (Klopries et al., 2013), the anticancer drug doxorubicin (Malla et al., 2010), the antioxidants epigallocatechin gallate (EGCG) and resveratrol (Lussier et al., 2012), and the cholesterol-lowering lovastatin (Guo &

Wang, 2014). Polyketides hold a considerable market share in pharmaceuticals, having annual sales of more than $20 billion (Hopwood, 2009).

Non-ribosomal peptides are mostly produced by soil-inhabiting and marine microorganism, and eukaryotic filamentous fungi. They are synthesised by non-ribosomal peptide synthetases (NRPS) using either proteogenic or non-proteogenic amino acid monomers. Proteogenic amino acids are part of the genetic code, while non-proteogenic amino acids are those not naturally encoded or found in the genetic code. Non-ribosomal peptides typically have macrocyclic or branched macrocyclic structures and possess a wide range of properties from being toxins, siderophores, pigments, antibiotics, cytostatics, immunosuppressants or anticancer agents (Wang et al., 2014). Some examples are the antibiotics surfactin (Mor, 2000) and gramicidin (Lipmann et al., 1971), siderophore vibriobactin (Griffiths et al., 1984) synthesised by *Vibrio cholerae*, and the immunosuppressant cyclosporin A (Finking & Marahiel, 2004).

Alkaloids are organic bases usually synthesised from tryptophan and dimethylallyl pyrophosphate (DMAPP) building blocks. Although they are mostly produced by plants, they are also produced by bacteria, fungi and animals (Kabera et al., 2014). They have been of interest due to their diverse range of bioactivities such as having anticancer, diuretic, antiinflammatory, antimicrobial, and antihypertensive effects (Connolly et al., 2005), and in particular their potent inhibition of epidermal growth factor receptor (EGFR) (Cruz-López et al., 2011). An example is amoebicide pyreudiones synthesised by *Pseudomonas fluorescens*.

The majority of terpenoids are synthesised by plants, including taxol (diterpene) of *Taxus buccata* and artemisinin (sesquiterpene lactone) from *Artemisia annua*, fungi are also known to synthesise important terpenes such as aristolochenes, carotenoids, gibberellins, indole-diterpenes and trichothecenes (Keller et al., 2005). Examples of terpenoids synthesised by *Micromonosporaceae* include platencin, platensimycin (Smanski et al., 2011) and carotenoids (Richter et al., 2015). These lipophilic compounds are composed of isoprenyl pyrophosphate (IPP) and DMAPP subunits and are assembled using terpene synthases or cyclases. These two subunits are generated from basic C_5 isoprenoid units and are synthesised through

two main metabolic pathways – mevalonate (MVA) pathway and non-mevalonate (MEP) pathway (Dairi et al., 2000). Terpenoids have a wide variety of structures being linear or cyclic, and saturated or unsaturated. Additionally, these compounds can be further modified.

Hybrid products produced using a combination of different biosynthetic pathways add to the variety of structures possessed by secondary metabolites. Well-known hybrid metabolites have NRPS and PKS origins and have been of interest due to their high potential as a bioactive metabolite. Examples of such hybrid compounds from bacterial sources includes bleomycin (Boettger & Hertweck, 2013), epothilone (Tang et al., 2000), yersiniabactin (Pelludat et al., 1998), and rapamycin (Khaw et al., 1998) while fusarin C (Khaw et al., 1998), cyclopiazonic acid (Tokuoka et al., 2008), and flavipucine (Gressler et al., 2011) are of fungal origins.

This project will focus on secondary metabolites (or natural products) that are synthesised by multi-modular, multi-domain proteins NRPSs and PKS. Well-known metabolites produced by *Streptomyces* species are of NRP, PK or NRP-PK hybrid origins. Examples such as antibiotics cephalosporin (NRP) produced by *Streptomyces clavuligerus*, tetracycline (PK) produced by *Streptomyces aureofaciens*; antitumorals bleomycin (hybrid NRP-PK) produced by *Streptomyces verticillus*; and immunosuppressants rapamycin (PK) produced by *Streptomyces hygroscopicus* (Siezen & Khayatt, 2008). Both NRPS and PKS systems are molecular assembly lines that facilitate the oligomerisation of multiple-amino/hydroxyacids or acyl-CoA precursors respectively, into complex polymers (Donadio et al., 2007). These complex polymers are then often further modified into other unique structures. The biosynthesis facilitated by both systems involves three fundamental steps – initiation, elongation and termination, performed by specialised modules of the synthases as depicted in Figure 2 below. These modules are usually encoded in large gene clusters (Khosla et al., 1999).

Non-ribosomal peptide synthetases modules usually comprise four protein domains (Figure 2A): an adenylation domain (A) that selects, activates and loads monomers such as proteinogenic or non-proteinogenic amino acids or carboxylic acids, a thiolation domain (T) also referred to as a peptidyl carrier protein (PCP) that

covalently fixes the amino acid onto the synthetase, a condensation domain that catalyses the peptide bond formation and lastly, a thioesterase domain (TE) that releases the assembled peptide from the synthetase (Sieber & Marahiel, 2005). Distinct specificity of the adenylation domain and subsequent modifications by cluster-embedded or standalone additional domains (e.g. methylation, epimerisation, cyclisation (Walsh et al., 2001), etc.) can give rise to diverse products differing in their composition and structure. Peptide products can be linear, branched, partially cyclic, cyclic or bicyclic. An example of a linear product is the pentadecapeptide, gramicidin (Kessler et al., 2004), branched product like vibriobactin (Keating et al., 2000), cyclic product like gramicidin S (Erlanger & Goode, 1960) and bicyclic product like actinomycin (Pfennig et al., 1999).

On the other hand, polyketide synthases usually also contain four core domains (Figure 2B): an acyltransferase domain (AT) that selects and activates the acyl-CoA monomers (e.g. acetyl-CoA), an acyl carrier domain (ACP), a keto-acyl synthase condensation domain (KS) and a releasing thioesterase domain (Te). Apart from the core domains, other modification domains such as ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) are also present. Similar to fatty acid synthesis, PKS creates enzyme-bound ketoacyl intermediates through step-wise decarboxylative condensations between the incoming monomers and the growing polyketide chain (Siezen & Khayatt, 2008).

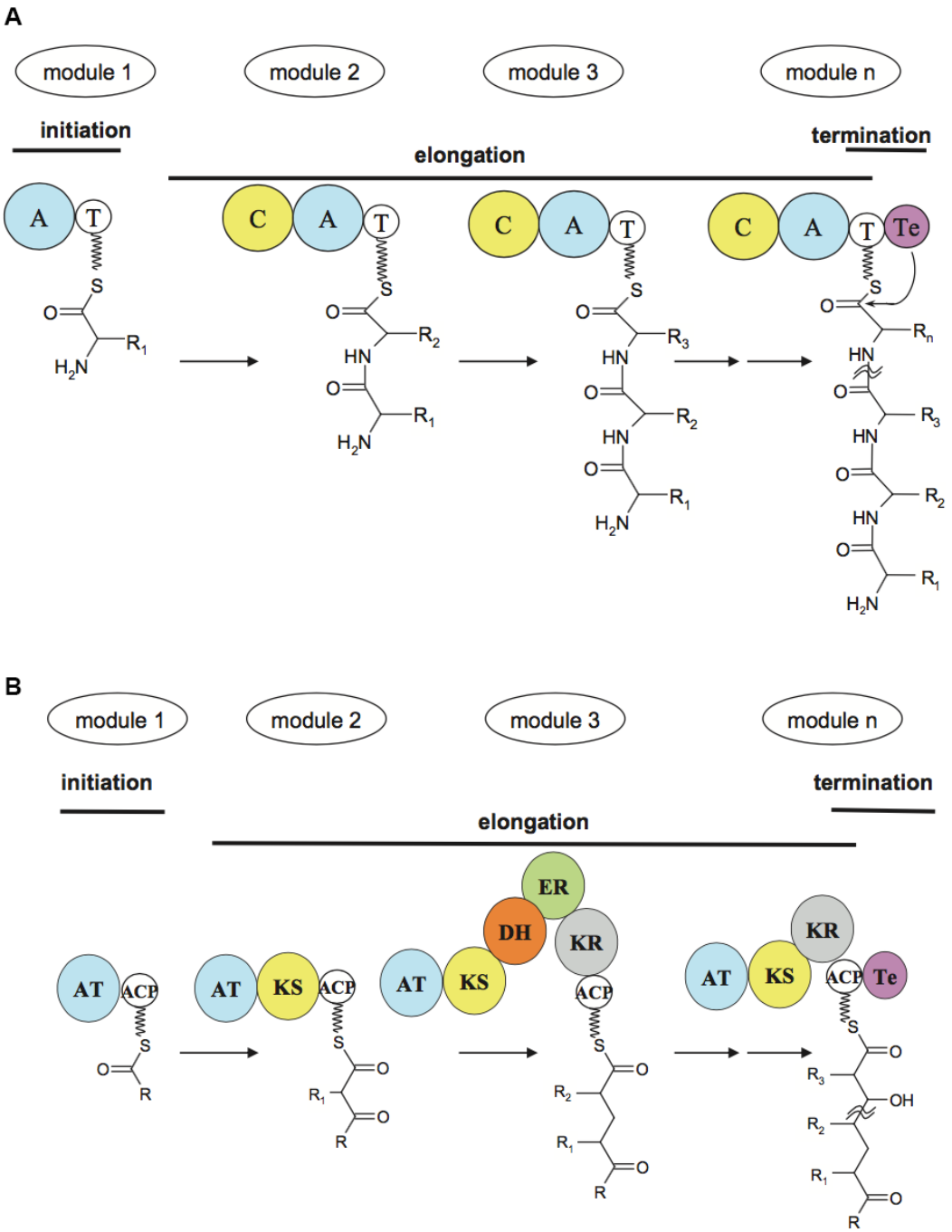


Figure 2. Fundamental steps in **(A)** non-ribosomal peptide and **(B)** polyketide synthesis adapted from (Donadio et al., 2007). The three basic steps involved are the initiation, elongation and termination. These steps are performed by specialised domains present in the synthases.

There are three general types of PKS, classified according to the structure of their functional domains – type I PKSs, type II PKSs and type III PKSs. Type I PKSs can be further categorised into type I iterative PKSs, type I modular PKSs, *trans*-AT modular PKSs and enediyne PKS (PKSE). Type I iterative PKSs consist of only a single PKS module, responsible in conducting multiple rounds of chain extension and β-keto processing. Varying degrees of reduction of β-keto in each round of chain

extension results in the diversity of products. Both non-reducing and partially reducing type I iterative PKSs are commonly found in actinomycete systems while fully reducing type I iterative PKSs in fungal systems (Cox & Simpson, 2009).

On the other hand, type I modular PKSs are large multi-domain proteins that function using a linear, processive mechanism to produce its final polyketide product. Each module is usually responsible for a single round of chain elongation, although exceptions may be present where modules may be iterative in a programmed manner (He & Hertweck, 2003) or due to certain irregularities. *Trans*-AT modular PKSs produces its final polyketide using similar mechanisms as type I modular PKSs; however, they lack AT domains (Helfrich & Piel, 2016). These *trans*-acting AT domains structurally resemble type II AT domains (Wong et al., 2011) and perform a role similar to *cis*-acting AT domains in type I modular PKSs. *Trans*-acting AT domains serve as proof-reading checkpoints when extender units are loaded onto the ACPs during polyketide chain elongation (Jensen et al., 2012). Enediyne PKSs are fully reducing iterative type I PKSs that are encoded in BGCs responsible for the production of natural products with 9 or 10-membered enediyne core moiety (Chen et al., 2014). They catalyse the production of a heptaene metabolite (Horsman et al., 2010) which is subsequently modified by accessory enzymes encoded by the BGC to produce the eventual enediyne natural product.

Type II PKSs function as a family of multienzyme systems that catalyse the formation of aromatic polyketide natural products such as actinorhodin and tetracenomycin (Dreier & Khosla, 2000). The absence of the KR and DH domains in type II PKSs results in the unstable polyketide products undergoing spontaneous cyclisation into aromatic moieties. Another core difference between the type I and II PKSs is that type II PKSs exists as an α/β-heterodimer complex where the α-subunit catalyses decarboxylative Claisen condensation during chain elongation while the β-subunit is responsible for the priming, elongation and termination of the polyketide product, controlling its chain length (Castaldo et al., 2008).

Type III PKSs exist as self-contained enzymes that form homodimers. Type III PKSs were previously believed to be plant-specific; however, recent research has revealed that type III PKSs are also widely present in microbes such as actinomycetes, and

fungi (Yu et al., 2012). Examples of bioactive compounds produced by type III PKSs in actinomycetes includes germicidin (Song et al., 2006) and violapyrones (Huang et al., 2016).

The collinearity rule is often used to explain the order and the structure of monomers present in the secondary polypeptide product formed through the NRPS and PKS system. The order of biochemical steps in the formation of the polypeptide product will follow the chromosomal order of the underlying modular genes present in the biosynthetic gene clusters (Reimer & Bode, 2014). The rule has provided a way of predicting the structures of possible compounds through analysis of the genomic sequences, especially in the more conserved PKS system. However, the collinearity rule does not always apply as some systems skip a module or utilise a module twice because of protein-protein interaction (Gokhale et al., 1999). Additionally, the modular nature of the NRPS and PKS system holds a huge potential for structural diversity and complexity of PKs and NRPs products. Variations in the selection and combination of monomers in the elongation steps, the degree of reduction in the condensation reactions and post-synthetic processing of the products (Menzella et al., 2005) can give rise to a wide range of products that cannot be accurately predicted based on genomic sequences alone. These variations are hence being extensively studied in the search for novel natural products of NRPS, PKS or NRPS-PKS hybrid origins.

With our focus being on specialised metabolites of NRPS, PKS or NRPS-PKS hybrid origins, we utilised next-generation sequencing and bioinformatics tools such as antiSMASH to identify potential gene clusters of such origins in our *Streptomyces* strains.

1.4 Activation of silent BGCs for the discovery of microbial natural products

There have been several successful approaches in the field of microbial natural products discovery, with three main traditional approaches being the modification of fermentation conditions, the use of chemical elicitors, and a co-cultivation approach to active cryptic BGCs in actinomycetes.

The former approach, generally named the ‘One Strain Many Compounds’ (OSMAC) method, was first conceptualised by Bode and colleagues (Bode et al., 2002). It was discovered that although a strain has the potential to produce numerous compounds, specific groups of compounds are only produced under specific culture conditions. Fermentation parameters that stimulate secondary metabolite productions include nutrient composition, trace elements and physical parameters (pH, temperature, etc.). An advantage of such an approach would be that a priori genomic information of the type of BGCs and regulatory processes guiding their expressions is not required. The OSMAC approach can hence also be extended to strains which are less responsive to genetic manipulation (Rutledge & Challis, 2015). Statistical experimental methods such as response surface methodology (RSM) and Plackett-Burman design (PBD) has also proved to be effective as an additional means to optimise secondary metabolite production by optimising selected, highly influential factors in the fermentation media (Jose et al., 2013).

Secondly, culture media can also be supplemented with specific chemical elicitors like sub-lethal concentrations of antibiotics or signalling molecules (Romano et al., 2018) to trigger the activation of cryptic BGCs in actinomycetes. The purpose of these chemicals is to mimic molecules present in the natural environment, usually signalling molecules that modulate quorum sensing communication processes. These chemicals affect proteins that are involved in certain global regulatory networks, trigger stress in the bacteria, or are precursors for the production of secondary metabolites (Seyedsayamdost, 2014). However, the effects of most chemical elicitors may still be unpredictable and random due to the complex nature of the entire metabolome and the cross-influence of associated pathways. For example, chemical elicitors like DMSO (Chen et al., 2000), ethanol and heavy metals are some compounds which precise mechanism of inducing secondary metabolite synthesis have yet been discovered (Pettit, 2011). A well-studied chemical would be *N*-acetyl-d-glucosamine (GlcNAc), a signalling molecule, which activates pathways for secondary metabolite biosynthesis in many fungi and bacteria. GlcNAc is a monomer of peptidoglycan in bacterial cell walls and chitin in fungal cell walls and is released during cell growth (Dashti et al., 2017). It has been studied that GlcNAc controls the activity of regulator proteins in actinomycetes, which binds to operator sequences of genes known to be responsible for growth and secondary metabolism

(Sánchez et al., 2010). The activation of secondary metabolite production by GlcNAc is triggered by a feast/famine signal where rich media conditions repress production while poor media condition activates production. For example, GlcNAc was studied to induce the production of undecylprodgiosin and actinorhodin antibiotics in *S. coelicolour* in poor nutritional conditions (Rigali et al., 2008). There are also studies reporting the effectiveness of sub-inhibitory concentrations of antibiotics to activate secondary metabolite productions in actinomycetes. Chemicals can also aid in the increase in yield of secondary metabolite production by increasing the supply of precursors in the culture media. For example, supplementing media with sodium acetate or sodium propionate in millimolar quantities have been reported to increase the yield of polyketide secondary metabolites by multiple times (Koju et al., 2012).

Co-cultivation is another approach to mimic the growth conditions in the natural environment and such a method has been studied to activate the genes synthesising cryptic secondary metabolites successfully. Target strains are cultivated together with either a collaborator or a competitor strain. Collaborator strains can activate secondary metabolite production through the production of quorum sensing molecules which act as communication signals while competitor strains usually do so by producing defence molecules as triggers (Pérez et al., 2011). For example, in *Streptomyces* strains, secondary metabolite production was induced when *Streptomyces lividans* was cultivated with mycolic acid (Onaka et al., 2011) producing bacteria (*Corynebacteriaceae* in particular). A co-cultivation system can be established on both plate and liquid cultures. Both plate and liquid cultures serve as a means to study direct cell-cell interactions between target strains. Plate cultures are effective in offering a visual indication of the production of potential antimicrobial compounds, for example, growth inhibition of bacterial strains (Dalmas et al., 2013). Otherwise, liquid cultures are typically preferred as they would allow for the possibility of compounds to be isolated in higher yield (Hoshino et al., 2015). Although such an approach may seem easy and useful, it requires rigorous optimisation of the culture conditions between bacterial strains to maintain optimum growth. Hence, binary combinations are often favoured over co-cultivation of multiple strains.

In the post-genomic era, easy access to a plethora of genomic data is available, largely favouring genome mining efforts. Such advancement has allowed for activation of silent BGCs through a genome-guided approach which we shall attempt. Current bioinformatics based approaches have provided insights into the identification of BGCs and prediction of the structure of the secondary metabolite potentially produced (Kim et al., 2017). One particular bioinformatics tool that is regularly used for genome mining is the antibiotics & Secondary Metabolite Analysis Shell (antiSMASH) (Weber et al., 2015). AntiSMASH works as an integrated platform combining prediction algorithms from other standalone prediction tools to provide a comprehensive identification of all putative BGCs present in a given DNA sequence. The main features of this platform include gene annotation, cluster identification, cluster comparison and substrate prediction amongst many others (Figure 3). AntiSMASH also allows comparison of BGCs with a known database from the 'Minimum Information about a Biosynthetic Gene cluster' (MIBiG) project, facilitating identification of orphan BGCs that could potentially be of interest. The MIBiG is effective in itself or can be used in tandem with antiSMASH to access a comprehensive database of experimentally validated BGCs (Medema et al., 2015). Apart from this, the Atlas of Biosynthetic gene Clusters within the Integrated Microbial Genomes system (IMG-ABC), the largest gene cluster database to date, also facilitates discovery and analysis of experimentally validated and computationally predicted biosynthetic gene clusters present in genomes and metagenomes (Hadjithomas et al., 2017). Similarly, annotation and analysis can be integrated with antiSMASH or used standalone.

The Dictionary of Natural Products (http://dnp.chemnetbase.com), Natural Product Alert (NAPRALERT) (Loub et al., 1985), and Super Natural II (Banerjee et al., 2015) provides information on basic structural and physicochemical properties, mechanism of action (MoA) and pathway information of natural compounds experimentally validated or already commercialised. *Streptomycetes* strains have proven to be such an important source of natural products that StreptomeDB, a freely accessible specialised database focusing on compounds produced by *Streptomycetes* is curated. It similarly holds extensive information such as host organisms, predicted physicochemical properties, synthesis pathways and bioactivities (Klementz et al., 2016).

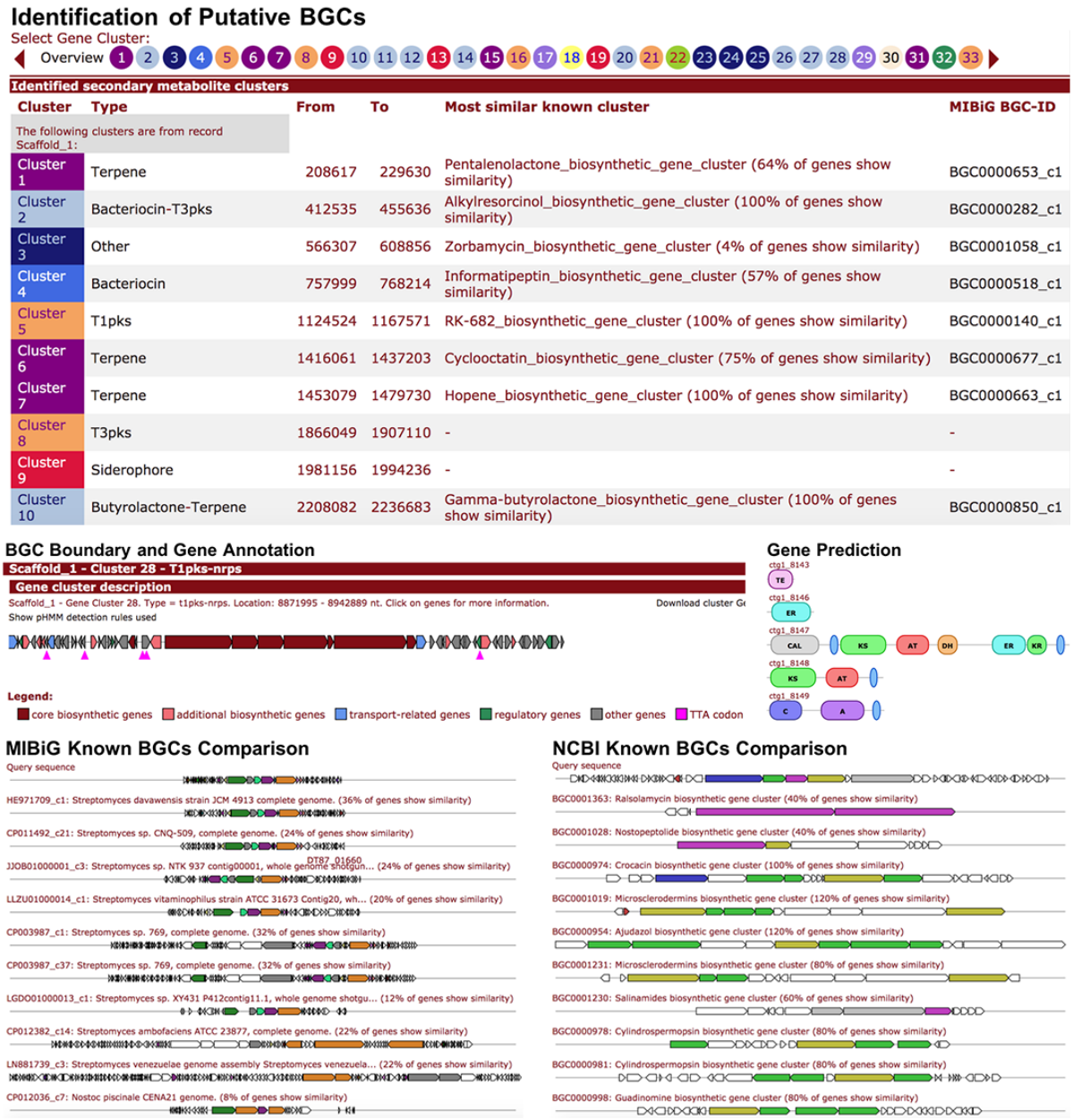


Figure 3. A schematic diagram illustrating the functions of antiSMASH. The analysis features putative BGCs identified, predicted BGC boundary and gene annotation within each cluster and gene predictions. Homologous comparison with known BGCs and sequences from MIBiG and NCBI databases respectively aids rigorous analysis.

1.5 Synthetic biological tools in natural products discovery

The regulation of BGCs involved in secondary metabolism can occur during both transcription and translation. Such regulation can be global or almost entirely specific for its respective pathway (Brakhage & Schroeckh, 2011). Pathway-specific methods involves transcription factors that are encoded by and regulate genes belonging to that particular cluster (Brakhage, 2013). These methods usually involve the over-expression of pathway-specific activator genes or deletion of pathway-specific repressor genes. Certain approaches are synergistic to each other, involving the

tandem knock out of a repressor and overexpression of a positive regulator (Baltz, 2016).

Pathway-specific regulatory proteins can play the role of an activator, a repressor or both, of which many have been identified and characterised. Examples of regulators that preferentially act as transcriptional activators come from the SARP (*Streptomyces* antibiotic regulatory proteins) family (Bruheim et al., 2002) or AraC/XylS (Gallegos et al., 1997) families while regulators from families such as the GntR (Haydon & Guest, 1991) and TetR (Ramos et al., 2005) families frequently function as transcriptional repressors. However, there are exceptions where some regulator families comprise both positive and negative regulators such as the LysR (Maddocks & Oyston, 2008) and MarR (Oh et al., 2007) families. A thorough study of the functions of pathway-specific regulators can allow us to engineer them in our plan to enhance metabolites production successfully.

Shuttle plasmids have been the most common method utilised to facilitate overexpression of genes. *E. coli* are a great host for these shuttle plasmids to replicate steadily for subsequent cloning. Conjugation remains the most favoured method for transforming actinomycetes with plasmids; hence, selected shuttle plasmids must contain an OriT required for conjugal transfer of plasmid from *E. coli* to actinomycetes. Such conjugal transfer can only be possible through the formation of a pilus between the donor *E. coli* and actinomycete strain, and an RP4 mediated transfer of the OriT containing plasmid in the form of single-stranded DNA. The required machinery is encoded by genes present in a broad host range RK2 plasmid which must also be present in the donor *E. coli* (Motallebi-Veshareh et al., 1992). Overexpression plasmids are a class of shuttle plasmids that will be essential in heterologous expression of targeted gene clusters and can be categorised as replicative or integrative. Replicative plasmids are better suited for stably maintaining only small-sized inserts but not larger ones (>40 kb) (Wohlleben & Pühler, 1987). However, heterologous expression of desired inserts would require the use of integrative plasmids which has the ability to integrate the inserts into specific sites in the chromosome of actinomycetes (Kuhstoss et al., 1991). Integrative plasmids utilise the serine integrases from phages to integrate inserts into specific attachment sites within the actinomycetal chromosome (Fogg et al., 2014). Examples of such

attachment sites that have been identified so far are *attB*$_{\Phi BT1}$ (Gregory et al., 2003), *attB*$_{\Phi C31}$ (Keravala & Calos, 2008) and *attB*$_{VWB}$ (Van Mellaert et al.,1998). These sites can integrate up to three inserts into different regions of an actinomycetal genome, facilitating its advantage as a biosynthetic tool for heterologous expression.

Another synthetic method in generating novel natural products would be through modifying enzymes responsible for the biosynthesis of natural products. Multiple enzymes catalyse the chemical transformation of intermediate substrates to form the final natural product. Through an in-depth understanding of the roles these enzymes play in the biosynthetic pathway, it provides the potential for engineering the biosynthesis of novel analogues. Such engineering strategies mainly target tailoring enzymes responsible for the latter stages of the biosynthetic pathway, reducing the risk of modified intermediates being incompatible for downstream processes. Engineered biosynthesis can increase the natural products repertoire and it has also been successful in improving pharmacological properties of bioactive natural products identified as potential drug leads (Pascolutti & Quinn, 2014). For example, polynik A and polyoxin N are two compounds produced through engineered combinatorial biosynthesis of the nikkomycin X nucleoside and the polyoxin dipeptidyl moiety, enhancing their antifungal properties compared to their parent antibiotics nikkomycin X and polyoxin B (Li et al., 2011). Engineered biosynthesis have even been further complemented by chemical synthesis in GenoChemetics where chemically synthesised molecules can be incorporated as precursors for an originally biosynthetic pathway, generating novel analogues with moieties not derived naturally (Sharma et al., 2017).

Most synthetic biology tools have been developed and optimised for *Streptomyces*, the largest and well-studied genus of *Actinomycetales*. With actinomycetes holding the potential as a rich source of natural product and secondary metabolite diversity is correlated to taxonomic diversity (Zotchev, 2014), it is worthwhile to explore cryptic BGCs in the taxonomically diverse *Streptomyces* strains.

However, optimisation of protocols for the genetic manipulation of individual actinomycete strains may be time-consuming, a limitation that suggests that heterologous expression of BGCs may be a more feasible approach. Heterologous

expression of targeted BGCs requires firstly, a well-studied actinomycete host with an easily controllable transcriptional machinery (Reen et al., 2015) and secondly, a robust cloning system. Characteristics of an ideal host include fast growth, small genome size, facile genetic tractability, etc. Moreover, to further aid their function as a host for secondary metabolite synthesis, redundant genes may be deleted from the genome while extra copies of useful genes may be added (Figure 4A). An example of such a host is *Streptomyces albus* J1047 which can be cultivated quickly, possesses a small genome size of about 6.8 Mb and is easily-manipulated (Zaburannyi et al., 2014). It has been successfully applied for heterologous production of compounds like napyradiomycin (Winter et al., 2007), cyclooctation (Kim et al., 2009), iso-migrastatin (Feng et al., 2009) and thiocoraline (Lombó et al., 2006). Additional mutations to regulate metabolic flux in *S. albus* J1047 have recently also shown effectiveness in increasing secondary metabolite yield (Kallifidas et al., 2018).

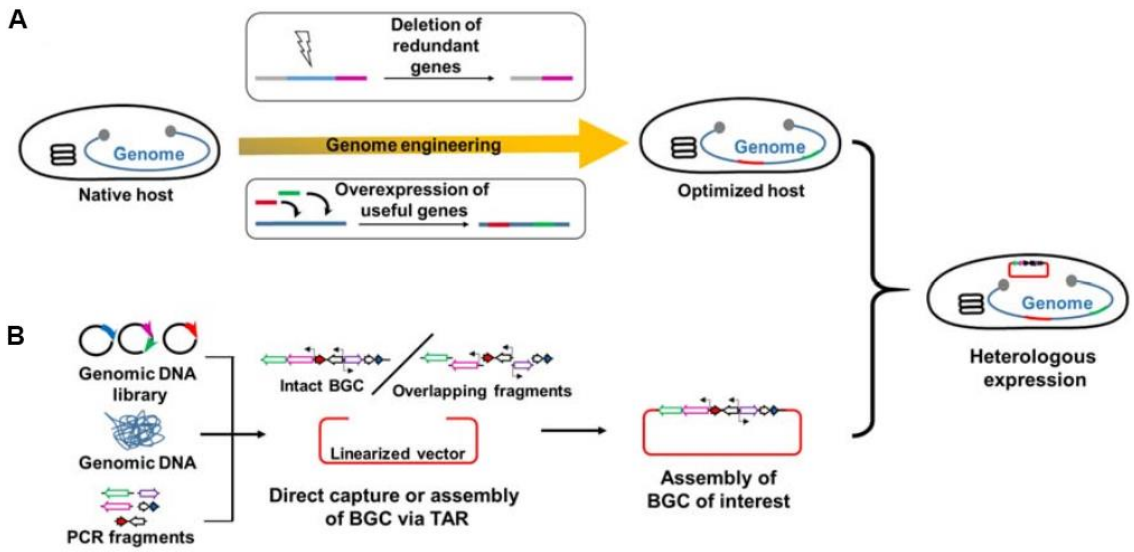


Figure 4. Strategy for heterologous expression of BGCs. **(A)** Optimisation of actinomycete host through genome engineering for secondary metabolism. **(B)** Constructed gDNA library, digested gDNA and chemically synthesised PCR fragments are cloned into a linearised vector for subsequent expression in a surrogate host. Adapted from Low, unpublished.

Targeted BGCs can be cloned through several strategies depending on its size. Conventional methods include the construction of a genomic DNA (gDNA) library of an actinomycete using cosmids (Denis & Brzezinski, 1992) or bacterial artificial chromosome (BAC) vectors (Alduina & Gallo, 2012). Cosmid and BAC vectors can accommodate insert sizes of about 40 kb and 150 kb respectively. Hence, BAC vectors are usually favoured to facilitate studies of actinomycete derived BGCs that

typically range from about 20 kb to 200 kb. The use of gDNA libraries is exceptionally useful if multiple BGCs need to be targeted. Otherwise, a targeted cloning approach like transformation associated recombination (TAR) cloning or Gibson assembly would be better suited. TAR cloning exploits the high level of homologous recombination in *Saccharomyces cerevisiae* to isolate and manipulate large gDNA inserts (up to 300 kb) into a yeast artificial clone (YAC) (Kouprina & Larionov, 2016). The effectiveness of TAR cloning has also been successfully extended to gDNA with high GC content, which makes it feasible for cloning majority of the actinomycete genes (Noskov et al., 2012). TAR cloning has recently been optimised as a tool to clone BGCs for biosynthetic studies using heterologous expression, making the requirement of a gDNA library redundant (Kim et al., 2010).

Conversely, Gibson assembly cloning operates *in vitro* to assemble DNA fragments with overlapping sequences into a single fragment for subsequent heterologous expression (Gibson et al., 2009). The usefulness of Gibson assembly has even been enhanced with the development of the technique Cas9-assisted targeting of chromosome (CATCH) (Jiang et al., 2015). The adaptability of the CRISPR/Cas9 system to cleave gDNA at any desired region is utilised by CATCH to produce homologous ends which can then be cloned into a vector via Gibson assembly.

Unlike the aforementioned methods, reporter-based screening techniques do not manipulate target BGCs for activation but instead facilitate the identification of activated BGCs by conventional, pleiotropic approaches. A reporter gene, typically a fluorescence protein or an enzyme that synthesises a coloured pigment is engineered under the control of a native promoter from an essential operon or one within the target BGC (Lünse & Mayer, 2017). The result of such engineering correlates the transcriptional activation of the targeted BGC with that of the reporter gene (Figure 5). This technique aims to complement genetic manipulation methods by easing the detection of an activated BGC, allowing screening to be scalable to high throughput capacity. The process of isolation and identification of natural products can be made less cumbersome.

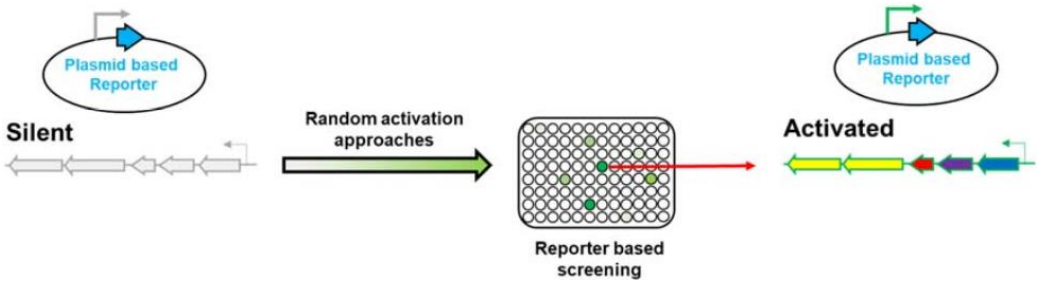


Figure 5. Reporter Based Screening. Conditions that are suitable for the activation of target BGCs are identified. Adapted from Low, unpublished.

The advantages and disadvantages of all these synthetic tools studied have guided this project in the selection and optimisation of methods most suitable for genome engineering of our *Streptomyces* strains.

1.6 The role of the *bldA* gene in the expression of biosynthetic genes

Beyond transcription, regulation of secondary metabolite production can occur during translational. The *bldA* gene has been identified to play a regulatory role on the translation of proteins in the *Streptomyces* strains, hence affecting their morphological differentiation and secondary metabolite production (Hackl & Bechthold, 2015). The *bldA* gene codes for the only tRNA in the entire genome that has the ability to translate a rare UUA codon into the amino acid leucine. Genes which contain a TTA codon can only be successfully translated into a protein when a functional *bldA* gene is present by subsequently generating a leu-tRNA$_{UUA}$ (Chater & Chandra, 2006). *BldA* mutants of *Streptomyces coelicolour*, *Streptomyces griseus* and *Streptomyces calvus* are unable to produce aerial hyphae and spores, and hence are labelled as having a 'bald' phenotype (Lawlor et al., 1987). A direct dependence of secondary metabolite producing genes on the *bldA* gene have been shown by the biosynthesis of actinorhodin, undecylprodigiosin, and methylenomycin in *S. coelicolor* (Chater, 2006); puromycin in *Streptomyces alboniger* (Tercero et al., 1998); a bacteriocin in *Streptomyces ipomoeae* (Wang et al., 2009); and expression of a landomycin regulatory gene in *Streptomyces globisporus* (Rebets et al., 2006). When a functional copy of the *bldA* gene is introduced through complementation into the mutants, aerial mycelium formation and sporulation was restored (Hackl & Bechthold, 2015).

It was also discovered that the main gene that *bldA* targets is *adpA*. The *adpA* gene similarly contains a TTA codon and mutants are also unable to produce aerial

mycelium and spores. When the TTA codon in *adpA* mutants are replaced with one of the other five available leucine codons, it only partially restored aerial mycelium formation and sporulation. Furthermore, when a set of other TTA-containing codons were disrupted, there was no visible difference in phenotypes compared to the wildtype. Hence, this illustrates that *adpA* is the main target of the *bldA* gene, resulting in the morphological differentiation displayed by the mutant strains compared to the wildtype (Hackl & Bechthold, 2015). Moreover, all research studies of cases where *bldA* dependent expression of natural product gene clusters have been noted, the disruption of the *bldA* gene in the host chromosome has resulted in the loss of production of secondary metabolites. A study of *S. calvus* has revealed that the reverse approach of expressing the correct copy of a *bldA* gene triggered the production of secondary metabolites that were encoded by cryptic genes (Kalan et al., 2013). This finding is of significance due to the presence of multiple TTA codons identified in the cryptic BGCs of our *Streptomyces* strains, highlighting the necessity to address such translational regulation when aiming to activate a cryptic BGC.

1.7 Aims and objectives

The main aim of my research project is to discover novel natural products from the actinomycetes culture collection from our lab. I have chosen to focus on *Streptomyces sp.* P46, a biosynthetically talented strain isolated from the sediment of a quarry. Two main approaches – an OSMAC approach and a genetic-based approach, were employed to activate cryptic BGCs in P46 for novel compound discovery. To achieve this aim, bioinformatic analysis will be performed for genome mining and identifying novel cryptic BGCs. Genetic vectors and protocols will be validated and used for *E. coli* – *Streptomyces* conjugation and gene overexpression. Additionally, I will rely on HPLC and LC-MS for metabolite profiling of the crude extracts for the detection and identification of novel secondary metabolites produced by the cryptic BGCs in both the wildtype and overexpression mutant strains.

CHAPTER 2: MATERIALS AND METHODS

2.1 Isolation of *Streptomyces* sp. P46 and its genomic DNA

A 100 metre-depth sediment was previously obtained from the Pulau Ubin Quarry, Singapore (Latitude:1.40590; Longitude: 103.95632) by members of the lab. The actinomycetes (consisting of my strain of interest, *Streptomyces* sp. P46) were subsequently isolated using various pre-treatments and culturing conditions as follows. Strains were cultured in 20 ml of CRM medium (10 g/L glucose, 103 g/L sucrose, 10.12 g/L $MgCl_2.6\ H_2O$, 15 g/L tryptic soy broth and 5 g/L yeast extract) in a 100 ml flask and incubated at 30°C for 2 – 3 d (early stationary phase). Mycelium of strains was prepared as described previously (Gomez-Escribano et al., 2015). The culture was divided into 5 mL aliquots in 15 mL centrifuge tubes (Greiner, Singapore) and the mycelium was harvested by centrifugation at 135000 × *g* for 5 min. The supernatant was discarded, and mycelium was washed with 5 mL of 20% glycerol. Aliquots that were not processed immediately were stored in 20% glycerol at -20°C. Washed mycelium were harvested by centrifugation at 135000 × *g* for 5 min. High quality of genomic DNA was extracted using the salting-out method as described in Practical Streptomyces Genetics (Kieser et al., 2000). Mycelium was resuspended in 5 mL of SET buffer (75 mM NaCl, 25 mM EDTA pH 8.0, 20 mM Tris-HCl pH 7.5) and homogenised using a Dounce homogeniser (Bellco, Sigma Aldrich, Singapore). Lysozyme of 140 µM was added and incubated at 37°C for 1.5 hours. Proteinase K solution of 17 M and 1% SDS were added and incubated at 55°C for another 2 h. NaCl of 1.25 M was added to the mixture and cooled to 37°C before chloroform extraction. A volume of 5 mL of chloroform was added and mixed gently by inversion until the mixture was homogenous. The mixture was centrifuged at 200000 × *g* for 5 min. The aqueous layer (top) was transferred to a new tube and chloroform extraction step was repeated until a clear layer was obtained. RNase A of 7 M was added to the aqueous layer and incubated at 37°C for 1.5 h. A proportion of 0.6 volume isopropanol was added and mixed by inversion to precipitate the genomic DNA. Genomic DNA was spooled out with a Pasteur pipette and washed with 70% ethanol. Genomic DNA was dried overnight in 2 mL centrifuge tube at room temperature. Tris buffer (pH 8.0) of 1M was added and shaken in cold room for a few days for dissolving genomic DNA. Quality of genomic DNA was analysed by

nanodrop readings, qubit readings and visualisation of agarose gel. The 260/280 ratio should be within 1.8 – 2.0 and the 260/230 ratio should be within 2.0 – 2.2 from nanodrop readings. The concentration of genomic DNA obtained from nanodrop and qubit readings should be within the range of ±50 µg/µL. The integrity of genomic DNA (at least 300 ng) has to be determined on a low percentage (0.6%) agarose gel at 100 V for 2.5 h with 1 kb DNA extension ladder. Genomic DNA should not show signs of degradation (smearing DNA bands).

2.2 Genome sequencing and genome annotation

The genome sequence of P46 was determined by former lab colleague Pang (unpublished). The genomic DNA was first sequenced using Illumina Miseq technology to obtain a draft genome which contains large gaps. Subsequent complete genome sequence of P46 was obtained using Single Molecule Real Time (SMRT) sequencing technology (Pacific Biosciences, California, USA) offered by PacBio RS II platform technology. The PacBio long sequencing reads successfully yielded 73 contigs with a total size of 9,824,301 bp by using the Hierarchical Genome Assembly Process 2 (HGAP2) protocol from SMRT Analysis version 2.0. Coding sequences (CDS) annotations of P46 was created by Rapid Annotation using Subsystem Technology (RAST) server, while secondary metabolite gene clusters were identified by ClusterFinder through antibiotics & Secondary Metabolite Analysis Shell (antiSMASH). AntiSMASH also aligned known biosynthetic gene clusters to its closest relative from databases using ClusterBlast and KnownClusterBlast. The BLAST Ring Image Generator (BRIG) software (Alikhan et al., 2011) was utilised to construct a circular genome map with the information above.

2.3 Overexpression plasmid construction (generation of mutants)

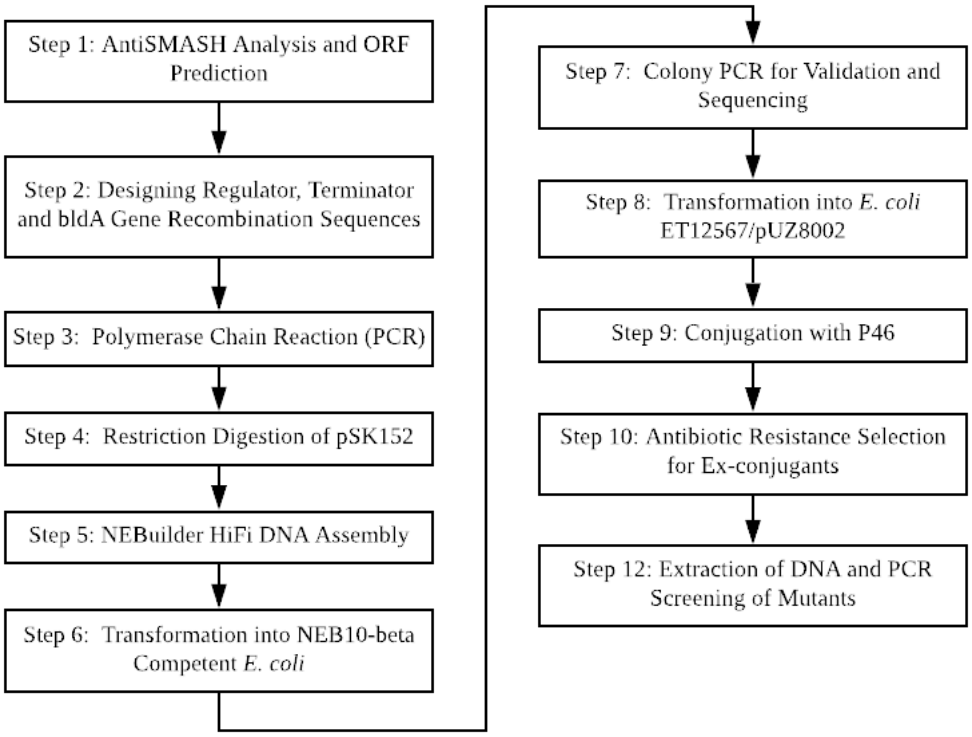


Figure 6. Workflow for the construction of overexpression plasmids.

The full genome sequence of P46 was analysed using the antiSMASH (version 5.0) online platform (Weber et al., 2015) and a BlastP search against the NCBI genome database (Altschul et al., 1990) validated the results.

The plasmid of choice is the pSK152 plasmid, an integrative and conjugative plasmid optimised for overexpression in actinomycetes (Low, unpublished). The design utilised a promoter and an RBS sequence that was separated by a ribozyme, RiboJ, which eliminates the interference of both elements through autocatalytic cleavage of sequence upstream of the ribozyme (Lou et al., 2012). The kasOp* promoter was selected to form the kasOp*–riboJ–RBS (from ΦC31 phage tail protein gene) cassette (Bai et al., 2015) and the cassette was subsequently inserted into a promoterless integrative and conjugative plasmid pSET152 to create pSK152.

The regulator, terminator (Deng et al., 1987) and *bldA* gene (Kalan et al., 2013) were constructed by PCR (Q5/Phusion DNA polymerase) using primers listed in Appendices, Table S1. The three PCR fragments were assembled and inserted into the *Not*I pre-digested pSK152 plasmid, using NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs, USA) in the order as illustrated by Figure 7 below. The

plasmid was transformed into NEB 10-beta competent *E. coli* cells and was screened with PCR (Taq DNA polymerase) using primers Terminator-F and *bldA*-R, and validated by Sanger sequencing (1st Base, Singapore).

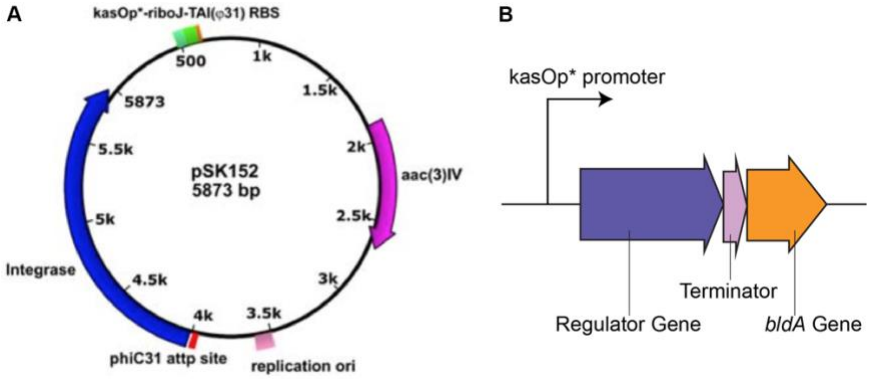


Figure 7. (A) psk152 plasmid utilised for overexpression in which **(B)** the three fragments (regulator, terminator and *bldA* gene) are inserted into following the order. The modified sequences are cloned into the plasmid downstream of the kasOp* promoter.

The validated plasmids were transformed into *E. coli* ET12567/pUZ8002 (Paranthaman & Dharmalingam, 2003) followed by conjugation with P46 according to standard protocol (Tong et al., 2015) to generate mutants P46-C12ACT, P46-C14ACT and P46-C28ACT. The ex-conjugants were selected and streaked on mannitol soy agar supplemented with nalidixic acid (50 µg/mL) and apramycin (50 µg/mL). Next, the colonies were transferred to GYM broth supplemented with apramycin (50 µg/mL) and incubated for 14 d at 37°C. The crude DNA of colonies were extracted and screened by PCR (Q5 polymerase) using primers PSK152-F and PSK152-R. Successfully transformed colonies are then subsequently sent for sequencing and concurrently screened by PCR (Q5 polymerase) using primers ΦC31-F and ΦC31-R for plasmid integration into the P46 genome.

2.4 Fermentation of *Streptomyces* sp. P46

P46 starter culture was prepared through the inoculation into 10 mL of GYM media and cultured for 5 d at 30°C with shaking. After which, 2% of the P46 starter culture was inoculated into 100 mL flasks containing 50 mL of media and cultured for 10 d at 30°C with shaking. Eight different media were selected (Table 1). The fermentation culture was then centrifuged for 15 min at 285000 × *g* and the culture broth (supernatant) and biomass were collected. The culture broth was extracted with an equal volume of ethyl acetate while the biomass was extracted with 50 mL of

acetone. The solvents, acetone and ethyl acetate, are combined and were evaporated using Genevac EZ-2 elite personal evaporator (Genevac Ltd, United Kingdom) at low boiling point mode, before switching to aqueous mode to completely remove traces of water. The dried extracts were dissolved in methanol (Optima™ LC/MS Grade, Fisher Chemical, USA) and centrifuged at 435000 × *g* for 2 min to remove any precipitate. Next, 30 µL of the extract was transferred to a microvolume glass vial insert for analytical HPLC.

Table 1. Recipe for the medium used for fermentation of *Streptomyces* sp. P46.

Media	Recipe in 1 L	Source
Complete Regeneration Media (CRM)	Glucose 10 g, Sucrose 103 g, $MgCl_2 \bullet 6 H_2O$, 10.12 g, Tryptic Soy Broth 15 g, Yeast extract 5 g	(Pigac & Schrempf, 1995)
Glucose, Yeast and Malt extract (GYM)	Glucose 4 g, Yeast extract 4 g, Malt extract 10 g (For agar, supplement with $CaCO_3$ 2 g, Agar 12 g)	(Flinspach et al., 2014)
Gauze	Soluble starch 20 g, KNO_3 1 g, NaCl 0.5 g, $MgSO_4 \bullet 7 H_2O$ 0.5 g, K_2HPO_4 0.5 g, $FeSO_4 \bullet 7 H_2O$ 0.01 g	(Dai et al., 2010)
Humic Acid Vitamin (HAV)	Humic acid 1 g, KCl 1.7 g, Na_2HPO_4 0.5 g, $MgSO_4 \bullet 7 H_2O$ 0.5 g, $CaCO_3$ 0.02 g, $FeSO_4 \bullet 7 H_2O$ 0.01 g, Vitamin B (0.5 g each of Thiamine-HCl, Riboflavin, Niacin, Pyridoxine-HCl, Calcium-Panthothenate, Inositol, p-Aminobenzoic acid and 0.25 mg of Biotin), Cycloheximide 0.05 g, Nalidixic acid 0.02 g, Agar 15 g	(Tseng et al., 2011)
L-Medium	Glucose 4 g, Yeast extract 4 g, Malt extract 10 g, $MnCl \bullet 4 H_2O$ 0.00015 g, $FeSO_4 \bullet 7 H_2O$ 0.00015 g, $ZnSO_4 \bullet 7 H_2O$ 0.0015 g, NaCl 0.0075 g, Boric 0.0015 g, Cu 0.0015 g	(Abbas & Edwards, 1990)
Oatmeal Agar	Oatmeal 20 g, Trace Salts solution ($FeSO_4 \bullet 7 H_2O$ 0.1 g, $MnCl_2 \bullet 4 H_2O$ 0.1 g, $ZnSO_4 \bullet 7 H_2O$ 0.1 g), Agar 18 g	(Xi et al., 2011)
Pharmamedia	Glucose 5 g, Cornsteep powder 5 g, Oatmeal 10 g, Cottonseed flour 10 g, K_2HPO_4 5 g, $MgSO_4 \bullet 7 H_2O$ 5 g, Trace metal solution 1 mL	(Lancini et al., 1967)
Starch Casein Nitrate Agar (SCN)	Starch 10 g, Sodium caseinate (Difco) 0.3 g, KNO_3 2 g, Agar 15 g	(Supong et al., 2013)
SG2	20 g Glucose, 5 g Yeast extract, 10 g Soytone and 2 g $CaCO_3$.	(Koshla et al., 2017)

Yeast extract Peptone Dextrose (YPD)	Peptone 20 g, Yeast extract 10 g, Glucose 20 g	('YPD media,' 2010)

2.5 Metabolite profiling of crude extracts using analytical HPLC

The following procedure was performed for P46 crude extracts with Agilent Eclipse XDB-C18, 100Å, (4.6 mm x 250 mm, 5 µm) using Agilent 1200 HPLC system (Agilent Technologies, Germany) equipped with DAD for UV detection. Mobile phase used: Buffer A consists of water + 0.1% formic acid and Buffer B consists of acetonitrile + 0.1% formic acid. HPLC gradient elution program was set with 10% B at 0 min, 20% B at 5 min, 70% B at 35 min, 90% B at 50 min, 100% B at 60 min. A volume of 20 µL of sample was injected with a flow rate at 1 mL/min. Analytes were monitored at λ = 220 nm, 260 nm, 284 nm, 314 nm, 330 nm, 360 nm, 400 nm and 420 nm.

2.6 Fractionation of organic extract with preparative HPLC

P46 was cultured to high cell density for 12 d. The resultant supernatant was extracted using ethyl acetate. A rotary evaporator was used to dry the extract, which would be subsequently dissolved in methanol. P46 EA extract was fractionated with preparative HPLC, Shimadzu Preparative Liquid Chromatography (Shimadzu Corporation, Japan) using a C18 column (9.4 x 250 mm; Agilent, Germany). The following conditions were used: Buffer A (acetonitrile) and Buffer B (water) were used for gradient elution from 30% Buffer A + 70% Buffer B to 100% Buffer A + 0% Buffer B; flow rate at 4.7 mL/min; injection volume of 1 mL of EA extract dissolved in methanol; analytes were monitored at λ = 220 nm, 254 nm, 280 nm, and 333 nm. The collection method used was time-based fractions, where each vial collected 3.2 mL.

2.7 Metabolite profiling by LC-MS

Liquid chromatography-mass spectrometry (LC-MS) was performed using the UltiMate 3000 UPLC-focused LC system coupled with an LTQ FT Ultra mass spectrometer equipped with a Zorbax® Eclipse plus C18 (2.1 mm x 50 mm, 5 µm). Water with 0.1% formic acid was set as mobile phase A while acetonitrile with 0.1% formic acid was set as mobile phase B. A flow rate of 0.3 mL/min was used and the

gradient elution program was similar to that employed for HPLC-DAD analysis. Mass spectra were measured using positive and negative ion modes of ESI and conditions of MS were set accordingly (capillary voltage 7 kV; capillary temperature 275°C; sheath gas flow rate 10 arb; auxiliary gas flow rate 0; spray voltage 5 kV) with a mass range of 100 – 2000 Da.

2.8 Genomic DNA extraction for gene cloning

The crude DNA of P46 strains for PCR reactions were extracted using Hexadecyl trimethyl-ammonium bromide (CTAB) buffer method. The liquid culture was centrifuged at 435000 × *g* for 2 min and the liquid medium was aspirated. The cell pellet was resuspended with 300 µL of CTAB (A stock solution of CTAB 2 g, 1 M Tris (pH 8) 10 mL, 0.5 M ethylenediamine tetraacetic acid disodium salt (pH 8) 4 mL, 5 M NaCl 28 mL, polyvinyl pyrrolidone (Mw 40,000) 1 g and water 58 mL) and incubated at 55°C for 15 min, centrifuged at 435000 × *g* for 2 min. The supernatant was transferred into a new 1.5 mL Eppendorf tube and mixed well with an equal volume of chloroform. The mixture was centrifuged at 435000 × *g* for 5 min and the top aqueous phase was transferred into a new 1.5 mL Eppendorf tube. Isopropanol was added based on 60% volume of aqueous phase solution, mixed well and centrifuged at 435000 × *g* for 5 min. Isopropanol was decanted and the DNA pellet was washed twice with 70% ethanol. The tube was dried in the oven at 55°C for 5 min and the DNA pellet was resuspended in 30 µL nuclease-free water.

2.9 Total RNA extraction for RT-qPCR

The RNA of P46WT and mutants were extracted using the following protocol. A volume of 1 mL of 10 d fermentation culture of P46 in GYM media was extracted and centrifuged at 435000 × *g* for 2 min. The pelleted cells were then resuspended thoroughly in 480 µL of 50 mM EDTA. As P46 is a gram-positive bacterial strain, lysozyme was sufficient for cell lysis. A volume of 120 µL of 700 µM lysozyme was added to the resuspended cell pellet and mixed gently. The samples were incubated at 37°C for 1 h before being centrifuged at 435000 × *g* for 2 min. The supernatant was removed.

The pelleted cells were further lysed using the freeze-thaw lysis method. Following which, the lysate was homogenised using a tissue homogeniser and TRIzol®

(Invitrogen, USA) reagent was added to allow complete dissociation of the nucleoprotein complexes. Cells were further lysed by repetitive pipetting using a 23G 1.25 TW (0.6 mm x 32 mm) PrecisionGlide™ needle (BD, United States of America) and were incubated at room temperature for 5 min. Add 0.2 mL of chloroform per 1 mL TRIzol® used, shake the tube vigorously by hand for 15 s and incubate at room temperature for 2-3 min. Centrifuge the sample at 400000 × *g* for 15 min at 4°C. The mixture will separate into a lower, red phenol-chloroform phase, an interphase, and a colourless upper aqueous phase containing the RNA. A volume of 450 µL of the colourless upper phase was transferred to a fresh RNase-free tube and an equal volume of 70% ethanol was added, vortexed to mix well. The tubes were gently inverted to disperse any visible precipitate that may form after adding ethanol.

A volume of 700 µL of the sample was transferred to a spin cartridge with a collection tube and centrifuged at 400000 × *g* for 15 s. The flow-through was discarded and the Spin cartridge reinserted into the same collection tube. This process was repeated until the entire sample was processed. Subsequently, 350 µL of Wash Buffer I was added from the PureLink® RNA Mini Kit (Invitrogen, USA) to the spin cartridge containing the bound RNA and centrifuged at 400000 × *g* for 15 s. The flow-through was discarded and the spin cartridge reinserted into the same collection tube. A volume of 80 µL of the PureLink® DNase was prepared by combining 8 µL of 10 x DNase 1 reaction buffer, 10 µL of resuspended DNA and 62 µL of RNase free water. This 80 µL of PureLink® DNase was directly added onto the surface of the spin cartridge membrane and incubated at room temperature for 15 min. Subsequently, 350 µL of Wash Buffer I was added into the same spin cartridge and centrifuged at 400000 × *g* for 15 s. The flow-through was discarded and the spin cartridge reinserted into a new collection tube. A volume of 500 µL of Wash Buffer II with ethanol was added to the spin cartridge and centrifuged at 400000 × *g* for 15 s. The flow-through was discarded and the spin cartridge reinserted into the same collection tube. The washing step was repeated once more. The spin cartridge was centrifuged at 400000 × *g* for an additional minute to dry the membrane containing the bound RNA. The collection tube was discarded and the spin cartridge was inserted into a recovery tube. A volume of 40 µL of RNase-free water was added to the centre of the spin cartridge and incubated at room temperature for 1 min. The

spin cartridge and recovery tube was centrifuged at 400000 × *g* for 1 min to elute the RNA. The DNA-free RNA samples were stored at -20°C.

2.10 RT-qPCR

The total RNA was subjected to pre-RT-PCR DNase treatment using RQ1 RNase-free DNase (Promega, USA). A mass of 700 ng of RNA was added to the DNase digestion mix and the reaction was incubated at 37°C for 30 min after which 1 µL of RQ1 DNase Stop Solution was added to terminate the reaction. Further incubate at 65°C for 10 min to inactivate the DNase.

Subsequently, the DNA-free RNA samples were reversely-transcribed using the RevertAid™ H Minus Reverse Transcriptase (RT) following the First Strand cDNA Synthesis protocol (ThermoScientific, USA). 1 µL DNA-free RNA, 1 µL of random hexamer primers was added to 10.5 µL of DEPC-treated water. The mix was then incubated at 65°C for 5 min. Subsequently, the following components were added to the mixture in the order – 4 µL 5x Reaction Buffer, 0.5 µL of ThermoScientific RiboLock RNase Inhibitor, 2 µL of 10mM dNTP mix and lastly 1 µL of RevertAid H Minus Reverse Transcriptase. The reaction mix was incubated at 25°C for 10 min, followed by 42°C for 60 min. The reaction was then terminated by heating at 70°C for 10 min. The reverse transcription products were either directly used in PCR or stored at -20°C.

CHAPTER 3: RESULTS

3.1 General features of the P46 genome

The assembly of *Streptomyces* sp. P46 genome using PacBio sequencing yielded 3 contigs - a linear chromosome with a total length of 9,249,452 bp with two additional plasmids of length 307,041 and 267,708 bp. General features of the chromosome sequence are shown in Figure 8 and Table 2. The circular genome map constructed by BRIG illustrates *Streptomyces* sp. P46's GC content, GC skew and 33 BGCs (Figure 8). The positions and type of the 33 BGCs in the genome are demarcated by their individual coloured bands. As annotated by RAST, the *Streptomyces* sp. P46 genome contains 449 functional metabolic pathways, 9116 coding sequences and 92 RNAs with a considerably high GC content at 70.3%.

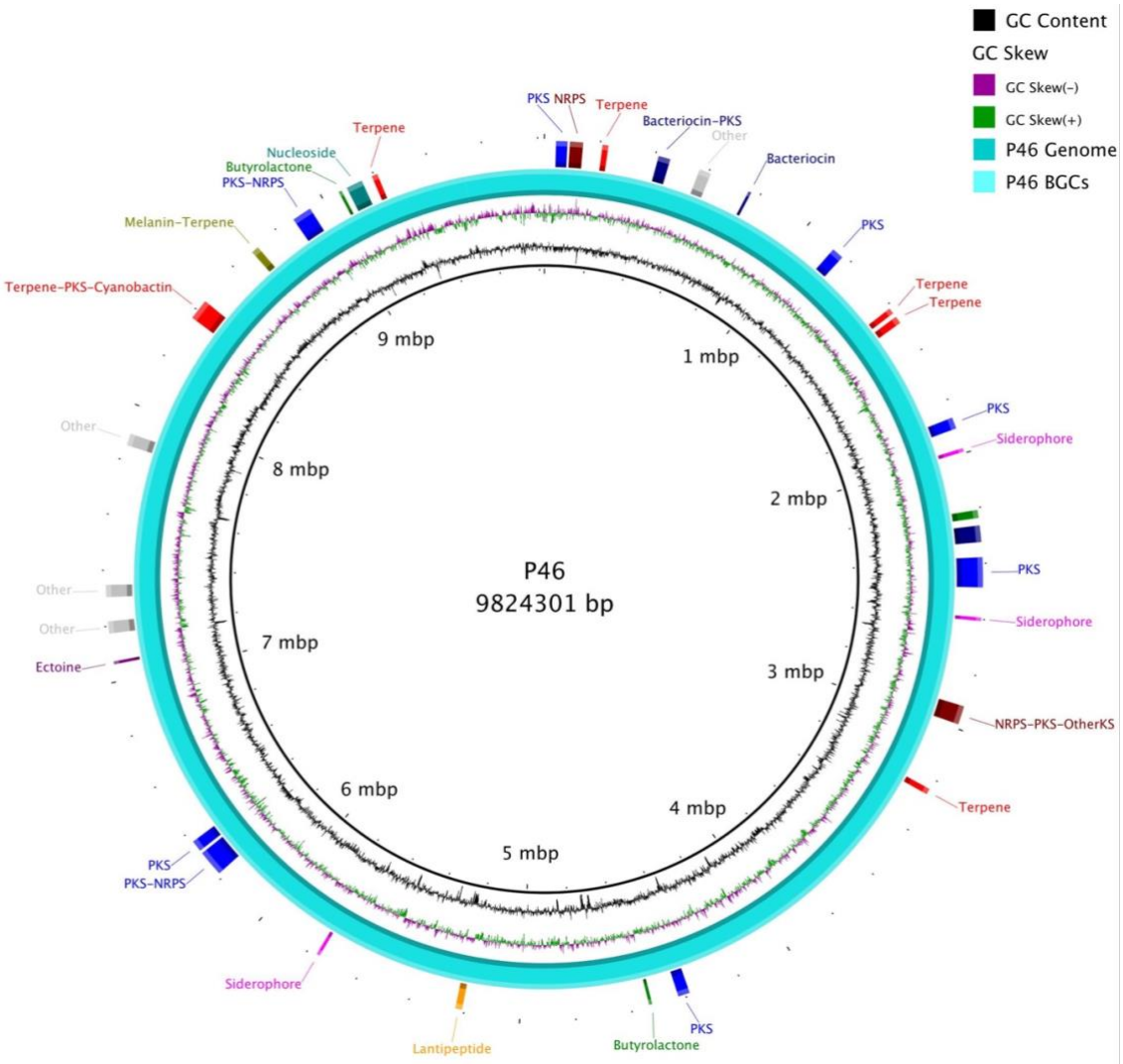


Figure 8. Circular representation of the genomic map of *Streptomyces* sp. P46. The Innermost ring depicts the GC content followed by the normalised GC skew with positive values in green and negative values in magenta. Secondary metabolite BGCs are annotated as bars in the outer ring with colours representing different classes of BGCs.

AntiSMASH analysis revealed 33 gene clusters among which PKS BGCs (11) were most abundant, followed by terpene BGCs (7). Information with regards to these gene clusters and the compounds which are predicted to be produced by the gene products is summarised in Table 2 below.

Table 2. Putative BGCs of *Streptomyces* sp. P46 identified by antiSMASH 5.0.

Cluster No.	Locus	Type	Predicted Compound Produced	Function
1	208617 – 229630	Terpene	Pentalenolactone	Antibacterial, antifungal, and antiviral
2	412535 – 455636	Bacteriocin-T3pks	Alkylresorcinol	Potential phytoanticipins and allelochemicals
3	566307 – 608856	Other		
4	757999 – 768214	Bacteriocin		
5	1124524 – 1167571	T1pks	RK-682	PTPase inhibitor
6	1416061 – 1437203	Terpene	Cyclooctatin	Lysophospholipase inhibitor
7	1453079 – 1479730	Terpene	Hopene	
8	1866049 – 1907110	T3pks		
9	1981156 – 1994236	Siderophore		
10	2208082 – 2236683	Butyrolactone-Terpene	Gamma-Butyrolactone	Antibacterial and antifungal
11	2263456 – 2323470	Bacteriocin-Oligosaccharide		
12	2377540 – 2485425	T2pks-T1pks	Pyrrolomicin	Antibacterial
13	2593519 – 2605837	Siderophore		
14	2916645 – 2983608	Nrps-T1pks-Otherks		
15	3236438 – 3257523	Terpene	Albaflavenone	Antibacterial
16	4384114 – 4426629	T2pks	Spore Pigment	Spore Colour
17	4523864 – 4534904	Butyrolactone		
18	5207537 – 5232046	Lantipeptide		
19	5753549 – 5765321	Siderophore	Desferrioxamine B	Iron Chelator
20	6217864 – 6312374	T1pks-Nrps	SCO-2138	-

21	6323478 – 6364563	T3pks		
22	7062297 – 7072701	Ectoine	Ectoine	Prevents osmotic stresses
23	7173564 – 7217319	Other		
24	7303966 – 7347817	Other		
25	7853953 – 7897222	Other		
26	8366841 – 8442028	Terpene-T3pks-Cyanobactin	Furaquinocin A	Antitumour
27	8684208 – 8710892	Melanin-Terpene	Melanin	
28	8871995 –– 8942889	T1pks-Nrps		
29	9061033 – 9071800	Butyrolactone		
30	9093038 – 9152249	Nucleoside	Nikkomycin	Competitive inhibitor of chitin synthetase. Inhibits growth of fungi, insects and acarids
31	9191547 – 9212470	Terpene		
32	41753 – 82829	T3pks		
33	91146 – 138903	Nrps		

Through antiSMASH analysis, out of the total 33 BGCs identified, 15 BGCs are already being studied. This suggests there are at least 18 potentially novel compounds from P46 WT that can be discovered (BGCs 3,4,8,9,11,13,14,17,18,21,23-25,28,29,31-33). These potentially novel compounds are of interest and there are ongoing efforts to activate their corresponding BGCs.

3.2 Metabolite profiling of P46 cultured in liquid media

Compounds with unique retention times and corresponding UV-Vis spectrums, through cross referencing with a natural compound library, were of interest (Lech & Jarosz, 2011). The HPLC chromatograms of the crude extracts of P46 cultured in SG2 media are shown in Figure 9. We have identified two known compounds after comparing the chromatograms of P46 WT (Figure 9B) with the control (Figure 9A) – 1-deoxypentalenic acid and alkylresorcinol. 1-deoxypentalenic acid has a UV_{max} at 220 nm (Figure 10A) and a retention time of 42.114 min. The LC-MS analysis of 1-deoxypentalenic acid observed an m/z = 234.162 $[M+H]_+$. Alkylresorcinol has a UV_{max} at 216 nm and 276 nm (Figure 10B), and a retention time from 61 – 65 min

(61.248, 62.394, 64.601 and 65.528 min). These four compounds are grouped together and since they possess the same UV-Vis spectrum, they are postulated to be analogues of alkylresorcinol. The LC-MS analysis of alkylresorcinol observed an m/z = 321.4167 $[M+H]_+$.

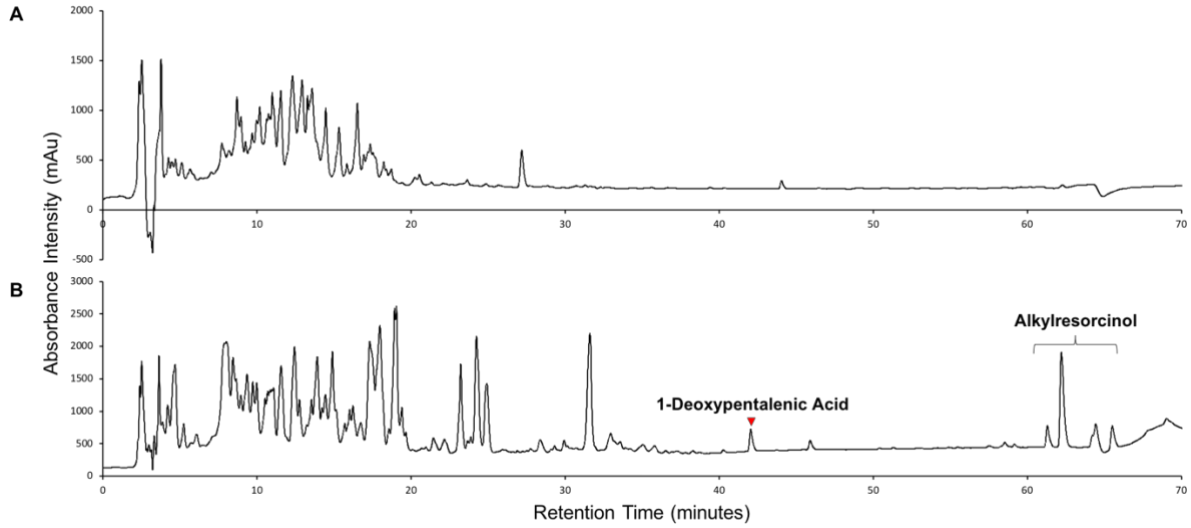


Figure 9. HPLC chromatogram at 220 nm of **(A)** SG2 media blank control and **(B)** P46 WT in SG2 media.

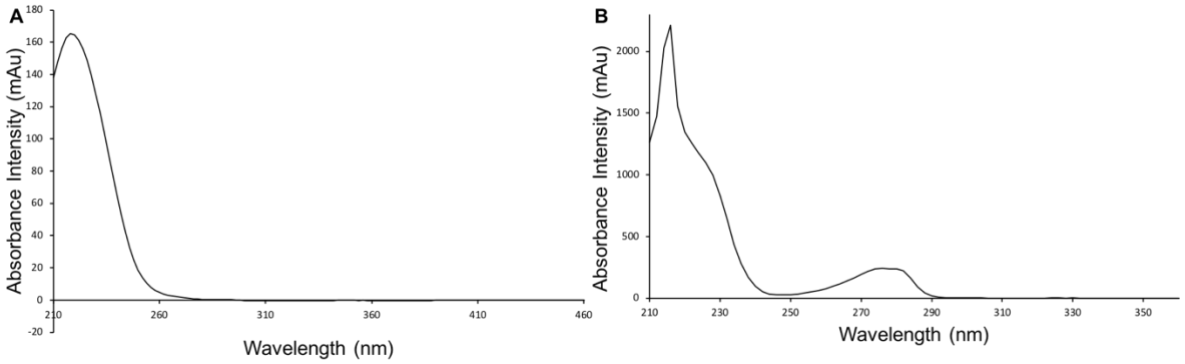


Figure 10. UV-Vis spectra of **(A)** 1-deoxypentalenic acid and **(B)** alkylresorcinol isolated from P46 WT in SG2 media.

3.3 Tasikamide A and its analogues

A group of potentially new compounds was identified after comparing the chromatograms of P46 WT (Figure 11B) and the control (liquid medium only) (Figure 11A). We identified and grouped them because although they possess retention times from 26 – 30 min, they share the same UV spectrum as depicted in Figure 12. The characteristics of these compounds do not match any of the known compounds recorded in the natural product library, hence it is identified and named as Tasikamide A by us. These compounds are in the process of being further purified.

The group of peaks were then further isolated and separated using 55% isocratic ACN with a highload C18 semiprep column. We have identified at least 12 analogues of Tasikamide A, identified as **T1** – **T12** (Figure 13). Their corresponding UV_{max} and m/z are indicated in Table 3. Analogue **T5** was further isolated and separated to obtain five additional analogues, identified as **T5.1** – **T5.5** (Figure 14). Their corresponding UV_{max} and m/z are similarly indicated in Table 3. Tasikamide A has a calculated m/z of 854.39 $[M+H]_+$ and its chemical structure has been confirmed through NMR. The NMR was carried out by Dr. Ma from my lab and the structure of Tasikamide A has also revealed its amino acid composition to be: hydroxyphenylalanine, isoleucine, N-methylphenylalanine, O-methyltryptophan, baba and a 4 carbon linker (Figure 15).

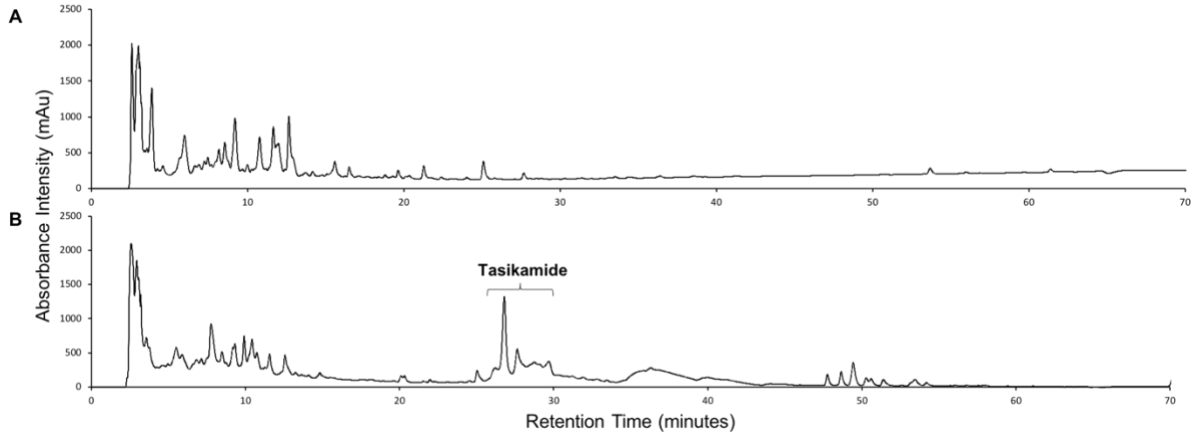


Figure 11. HPLC chromatogram at 220 nm of **(A)** GYM media blank control and **(B)** P46 Wildtype in GYM media.

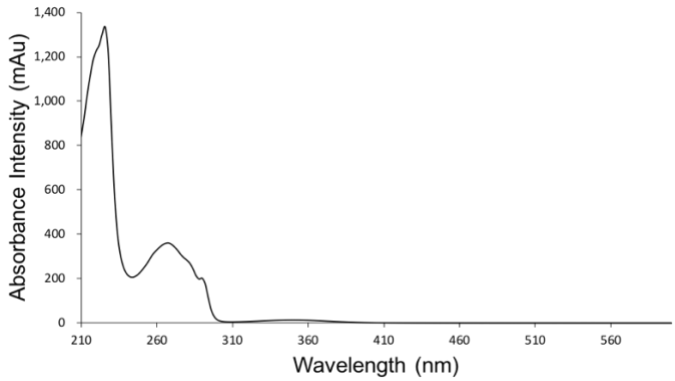


Figure 12. UV-Vis spectrum of Tasikamide isolated from P46 WT in GYM media.

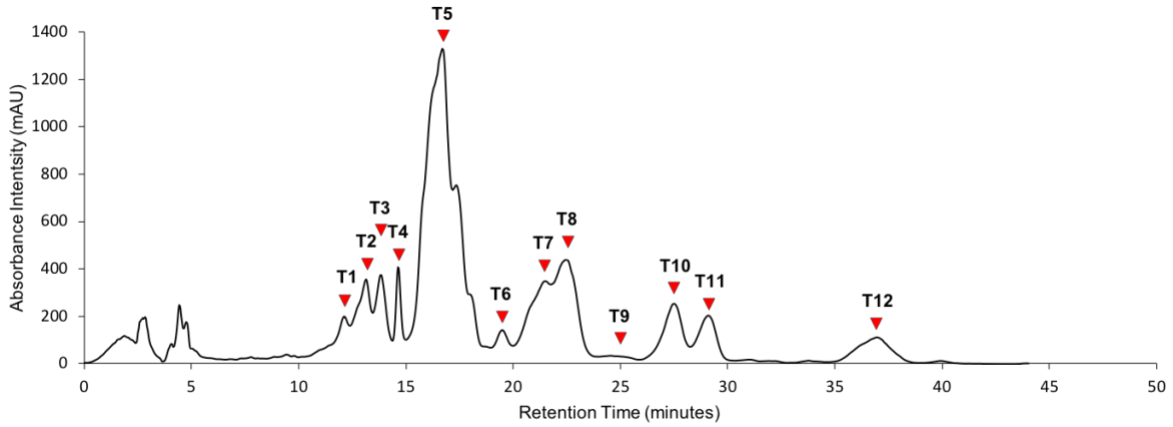


Figure 13. HPLC chromatogram at 284 nm of Tasikamide analogues (**T1** – **T12**).

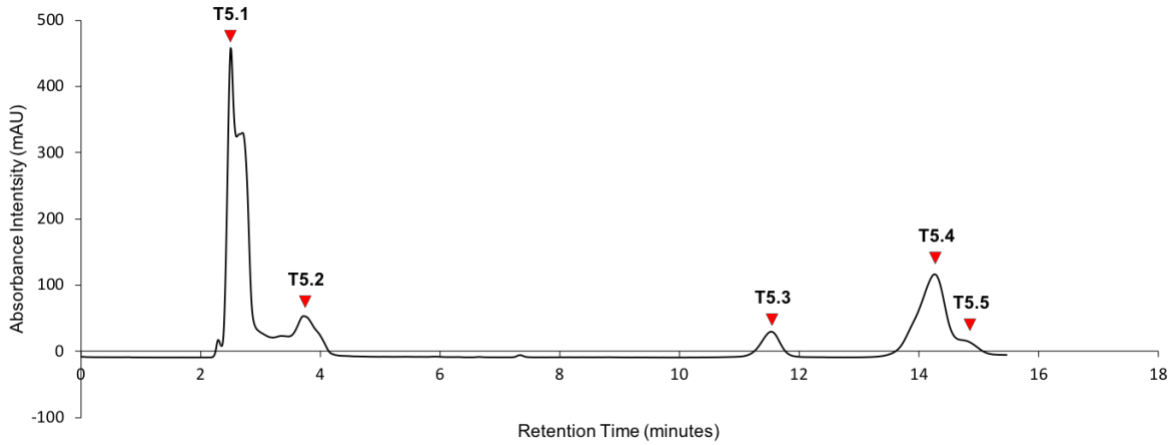


Figure 14. HPLC chromatogram at 260 nm of compound **T5** analogues (**T5.1** – **T5.5**).

Table 3. UV-Vis spectra and mass spectrometry data for Tasikamide A and its analogues.

Tasikamide Analogues	UV_{max} (nm)	m/z ($[M+H]_+$)
T1	194, 217, 269, 290	827.3949
T2	197, 267, 289	857.4065
T3	193, 217, 267	871.3840
T4	217, 264	811.4009
T5	229, 268	841.4112
T5.1	197, 268; 192, 218, 268	855.3890; 855.3890
T5.2	218, 267	855.3887
T5.3	219, 268	887.4162; 887.4175; 855.3891; 855.3897
T5.4	192, 218, 268	841.4110
T5.5	219, 267	841.4095
T6	217, 290	839.4305
T7	222, 268	825.4117
T8	220, 268	839.3948
T9	217, 267	811.3990

T10	218, 268	869.4061
T11	195, 218, 268	869.4057
T12	196, 218, 268	853.4122

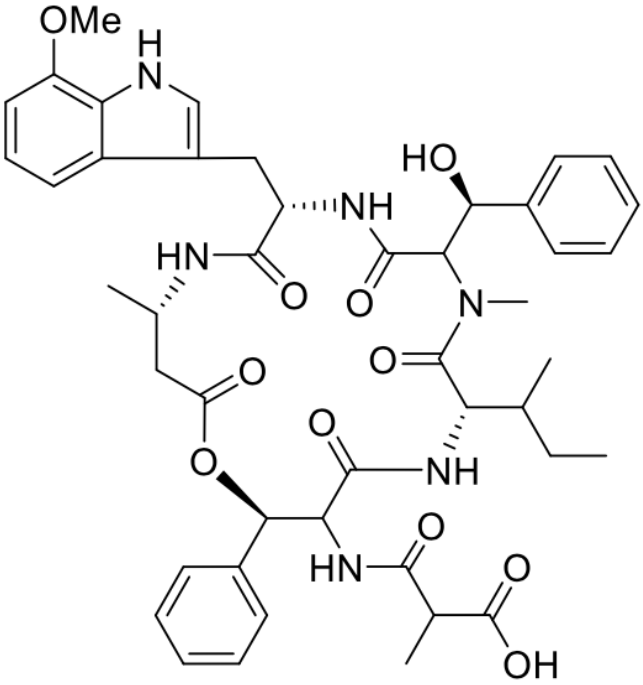


Figure 15. Chemical structure of Tasikamide A confirmed through NMR.

3.4 Potential novel compounds detected by the OSMAC approach

Other than compounds mentioned in **3.2**, there were more potential compounds that were detected through the OSMAC approach, contributing to our efforts in isolating novel compounds from *Streptomyces* sp. P46.

HPLC analysis has revealed another two potentially novel compounds isolated from the ethyl acetate extraction of the P46 WT supernatant after culture in GYM media – compounds **c** and **d** (Figure 16) at retention times of 8 and 14 min respectively. The UV-Vis absorption spectrum revealed UV_{max} at 225 and 310 nm for compound **c** (Figure 17A) and LC-MS analysis observed an m/z = 213.0903 $[M+H]_+$. The UV-Vis absorption spectrum revealed UV_{max} at 217, 244 and 332 nm for compound **d** (Figure 17B) and LC-MS analysis observed an m/z = 179.0810 $[M+H]_+$.

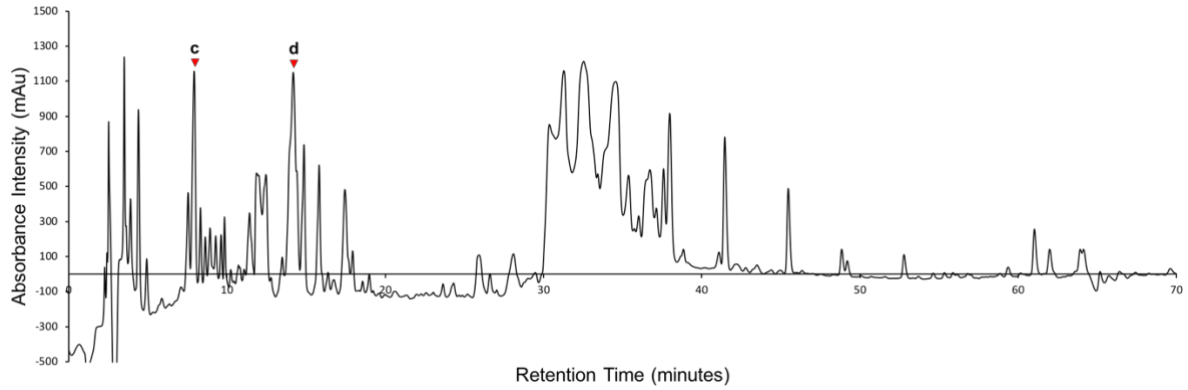


Figure 16. HPLC chromatogram at 220 nm of compound **c** and **d** isolated from ethyl acetate extraction of P46 WT supernatant in GYM.

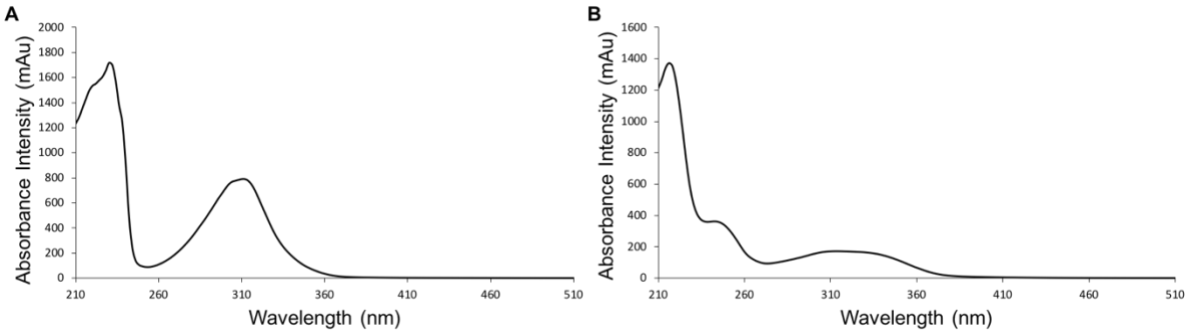


Figure 17. UV-Vis spectra of **(A)** compound **c** and **(B)** compound **d** isolated from ethyl acetate extraction of P46 WT supernatant in GYM media.

Additionally, another potentially novel compound was isolated from the acetone extraction of the P46 WT biomass after culture in GYM media - compound **e**, with retention times from 45 – 55 min. Five analogues (Figure 18) were identified and grouped according to their similar UV-Vis spectrum (Figure 19). Their UV_{max} and respective m/z are indicated in Table 4.

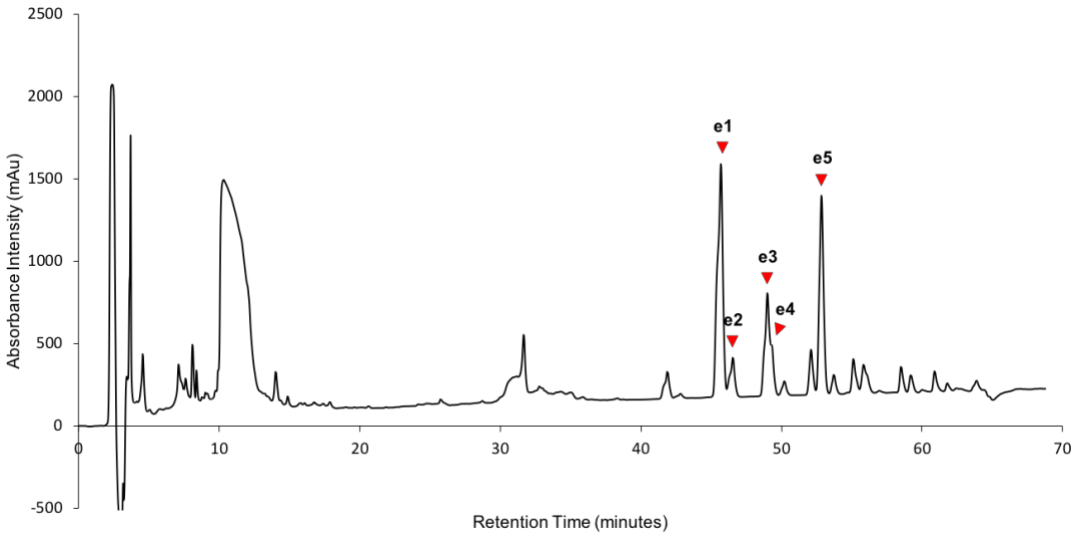


Figure 18. HPLC chromatogram at 220 nm of analogues of compound **e** isolated from acetone extraction of P46 WT biomass in GYM.

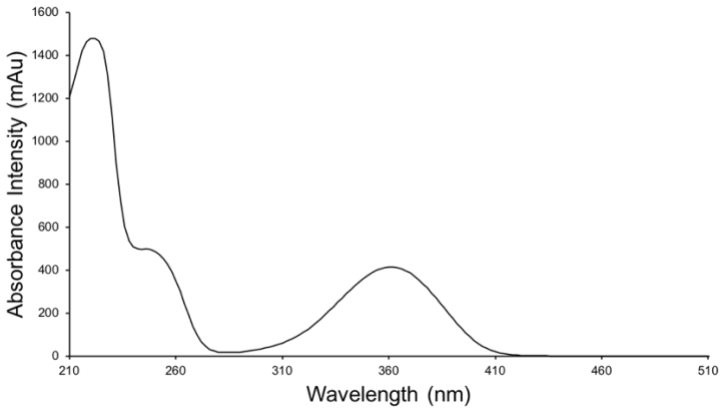


Figure 19. UV-Vis spectrum of compound **e** isolated from P46 WT in GYM media.

Table 4. UV-Vis spectra and mass spectrometry data of compounds **e1**, **e2**, **e3**, **e4**, and **e5**.

Potential Compounds	UV_{max} (nm)	m/z ($[M+H]_+$)
e1	220, 360	294.2053
e2	221, 360	294.2053
e3	220, 361	308.2208
e4	220, 361	308.2208
e5	220, 361	322.2364

Another group of potentially novel compounds was identified after comparing the chromatograms of P46 WT in SG2 with the profile of the media control. The compounds were identified as **f**, **g**, **h** and **j** (Figure 20) with retention times of 8, 19, 23 and 24 min respectively. Their UV_{max} and m/z indicated in Table 5.

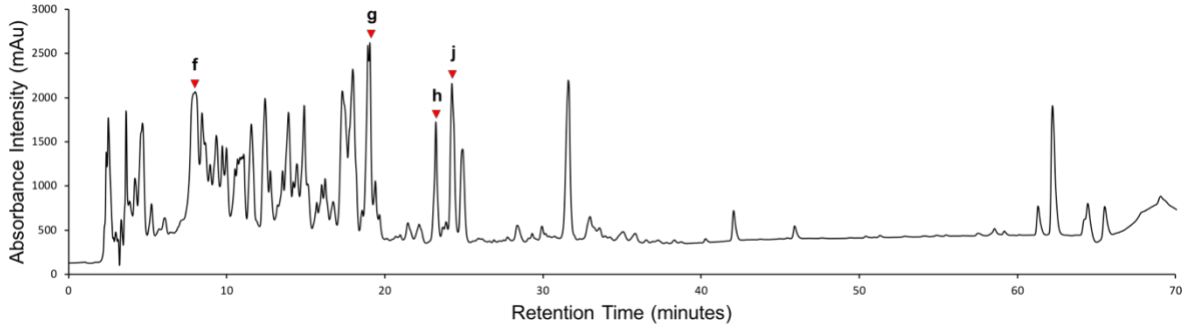


Figure 20. HPLC chromatogram at 220 nm of compounds **f**, **g**, **h** and **j** from P46 WT in SG2.

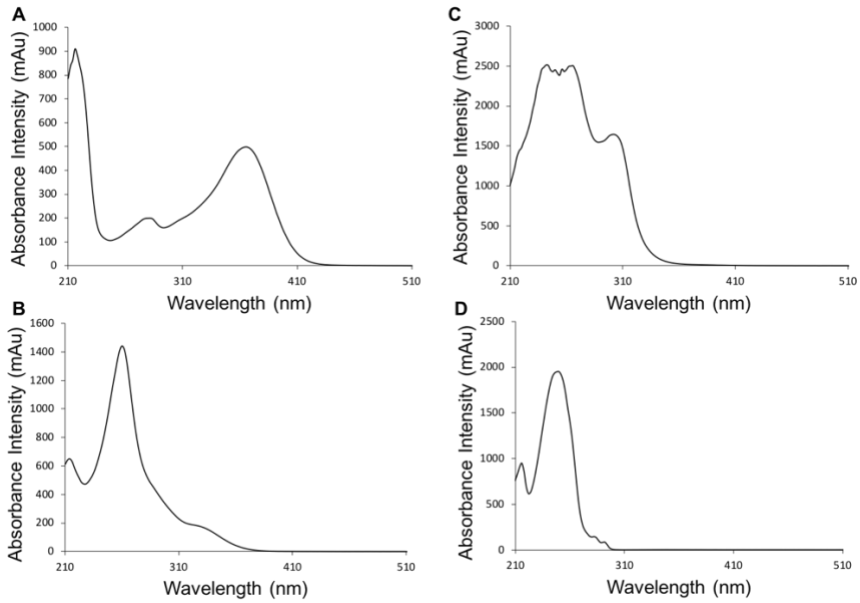


Figure 21. UV-Vis spectra of compounds **(A) f**, **(B) g**, **(C) h**, and **(D) j** isolated from P46 WT in SG2 media.

Table 5. UV-Vis spectra and mass spectrometry data of compounds **f**, **g**, **h** and **j**. The m/z of these compounds have yet been confirmed.

Potential Compounds	UV_{max} (nm)	m/z ($[M+H]_+$)
f	219, 322	-
g	195, 248, 299	-
h	220, 269	-
j	216, 242	-

These seven compounds (**c**, **d**, **e**, **f**, **g**, **h**, **j**) and their analogues are undergoing characterisation. Identification of their mass and chemical structure by LC-MS and NMR respectively will determine if they are novel compounds produced by *Streptomyces* sp. P46.

3.5 Gene annotation of BGCs 12, 14 and 28

BGCs of NRPS, PKS or hybrids thereof origin were identified and specifically focused on. Prominent classes of bioactive compounds isolated from microbial origin have been shown to be synthesised by such BGCs (Müller et al., 2015). Focusing on BGCs of these origins, BGCs 12, 14 and 28 piqued our interest (Figure 22). AntiSMASH analysis has not identified known compounds that were predicted to be produced by BGCs 12, 14 and 28, suggesting their potential to produce novel compounds. AntiSMASH analysis has provided us with individual gene annotations and predicted operon boundaries (identified by A, B, C, D, P, Q, R, S, T, X, Y, Z) (Figure 22) while BlastP analysis (Tables 6 – 8) has provided us with the predicted protein product and function of an individual gene within the three BGCs of interest. Correlating the operon boundaries identified with BlastP results (Tables 6 – 8), we were able to predict possible products of the genes present in the operons. Furthermore, the BlastP analysis identified the presence of pathway-specific regulators in these BGCs. The regulator genes targeted are ORF2243 in C12, ORF2694 in C14 and ORF8145 in C28. ORF2243 and ORF2694 are both predicted to code for AfsR/SARP family transcriptional regulators (identity: 71% and 56% respectively) while ORF8145 is predicted to code for transcriptional regulator NovG (identity: 51%). The *Streptomyces* antibiotic regulatory protein (SARP) family and NovG transcriptional regulators have been predicted to be positive regulators of secondary metabolites in *Streptomyces* and the overexpression of such pathway-specific regulators have proven successful in activating cryptic BGCs. Hence, we adopted a targeted genomic approach was adopted to activate these regulator genes. Moreover, the *bldA* gene was also deliberately overexpressed to target the presence of TTA codons in these regulator genes (ORF2243, ORF2694 and ORF8145), similarly identified by antiSMASH analysis. This approach aims to target the regulation of the BGCs during both the transcription and translation.

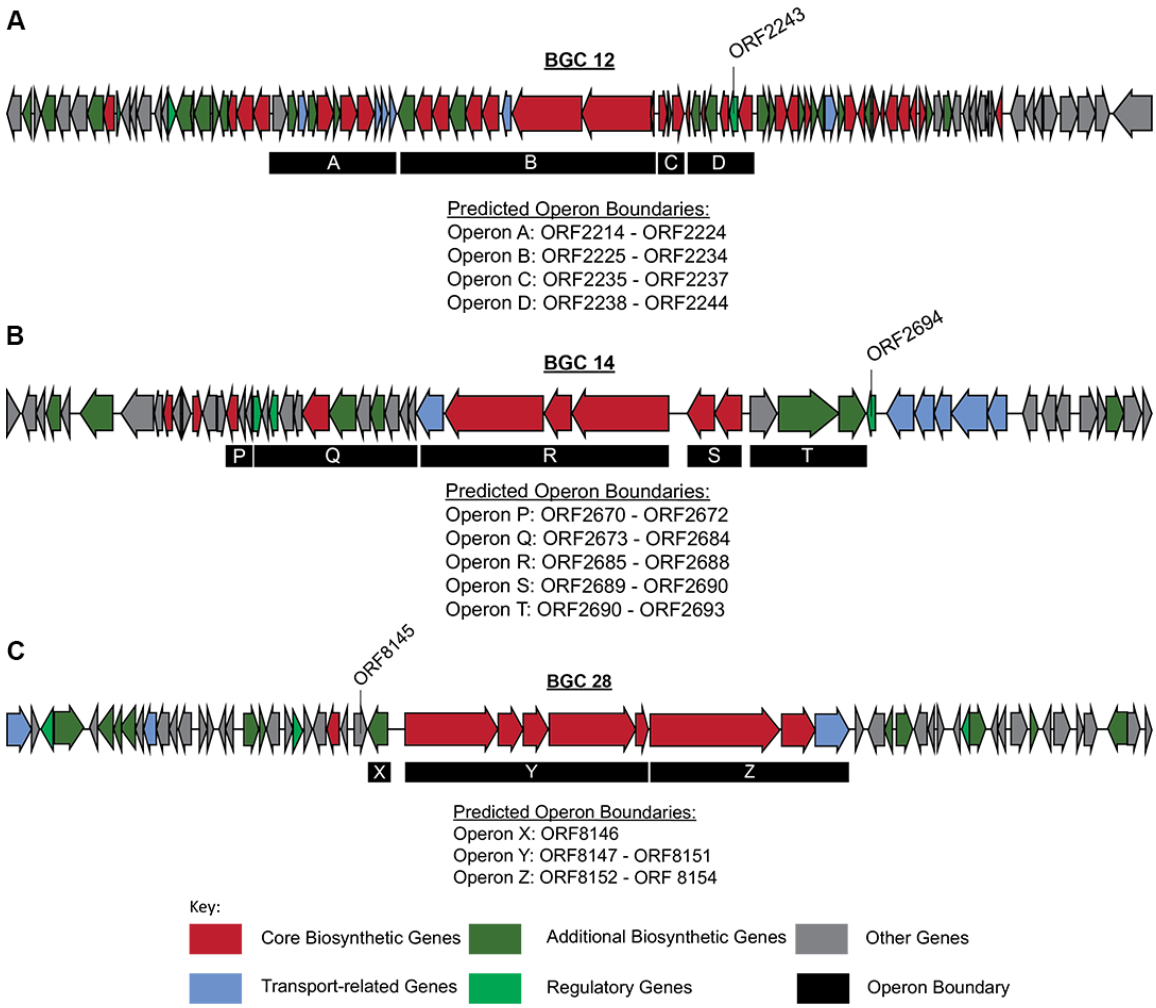


Figure 22. Genetic organisation of P46 **(A)** BGC12, **(B)** BGC14 and **(C)** BGC28. The major operon boundaries highlighted and annotated.

Table 6. Predicted function of genes from *Streptomyces* sp. P46 BGC 12. ORF2243, the predicted regulator we targeted for overexpression is highlighted in grey.

CDS	Protein and/or function	BLASTP best hit
2214	2-oxoglutarate oxidoreductase, alpha subunit (EC 1.2.7.3)	ref|WP_093719770.1|; 2-oxoacid:acceptor oxidoreductase subunit alpha; *Streptomyces* sp. strain DI166; identity, 95%
2215	2-oxoglutarate oxidoreductase, beta subunit (EC 1.2.7.3)	ref|WP_015657246.1| 2-oxoacid:ferredoxin oxidoreductase subunit beta; *Streptomyces davaonensis*; identity, 93%
2216	Na+/H+ antiporter	ref|WP_093719768.1|; cation/H(+) antiporter; *Streptomyces* sp. strain DI166; identity, 93%
2217	Butyryl-CoA dehydrogenase (EC 1.3.99.2)	ref|WP_015657248.1|; acyl-CoA dehydrogenase family protein; *Streptomyces davaonensis*; identity, 95%
2218	Peptide synthetase	ref|WP_015657249.1|; amino acid adenylation domain-containing protein; *Streptomyces davaonensis*; identity, 95%

2219	Nitrilotriacetate monooxygenase component B (EC 1.14.13.-)	ref|WP_015657250.1|; flavin reductase; *Streptomyces davaonensis*; identity, 97%
2220	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|AFP87527.1|; FADH2-dependent halogenase; *Streptomyces* sp. strain CNQ-418; identity, 70%
2221	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|AFP87528.1|; FADH2-dependent halogenase; *Streptomyces* sp. strain CNQ-418; identity, 79%
2222	ABC transporter permease protein	ref|AFP87529.1|; ABC transporter; *Streptomyces* sp. strain CNQ-418; identity, 51%
2223	ABC drug efflux pump, inner membrane subunit, DrrB family	ref|AFP87530.1|; ABC transporter; *Streptomyces* sp. strain CNQ-418; identity, 43%
2224	PROBABLE CONSERVED ATP-BINDING PROTEIN ABC TRANSPORTER	ref|WP_083978474.1|; ABC transporter ATP-binding protein; *Micromonospora rosaria*; identity 67%
2225	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|WP_093719765.1|; hypothetical protein; *Streptomyces* sp. strain DI166; identity, 92%
2226	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|WP_093719783.1|; FAD-dependent oxidoreductase; *Streptomyces* sp. strain DI166; identity, 95%
2227	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|WP_093719764.1|; NAD(P)/FAD-dependent oxidoreductase; *Streptomyces* sp. strain DI166; identity, 93%
2228	monooxygenase, FAD-binding	ref|WP_015657254.1|; 3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase Short=3-HPP hydroxylase; *Streptomyces davaonensis*; identity, 79%
2229	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|CCK26858.1|; halogenase B; *Streptomyces davaonensis* JCM 4913; identity, 96%
2230	Monodechloroaminopyrrolnitrin-3-halogenase PrnC	ref|WP_093719761.1|; FAD-dependent oxidoreductase; Streptomyces sp. strain DI166; identity, 94%
2231	transmembrane efflux protein	ref|WP_015657258.1|; DHA2 family efflux MFS transporter permease subunit; *Streptomyces davaonensis*; identity, 87%
2232	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_015657259.1|; type I polyketide synthase; *Streptomyces davaonensis*; identity, 93%
2233	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_015657260.1|; acyltransferase domain-containing protein; *Streptomyces davaonensis*; identity, 87%
2234	putative polyketide synthase	ref|CCK26867.1|; Putative polyketide synthase pksL Short=PKS; *Streptomyces davaonensis* JCM 4913; identity, 89%
2235	putative thioesterase	ref|WP_106435759.1|; thioesterase; *Streptomyces*

		davaonensis; identity, 93%
2236	hypothetical protein	ref|WP_041821273.1|; acyl carrier protein; Streptomyces davaonensis; identity, 93%
2237	FIG022199: FAD-binding protein	ref|WP_015657264.1|; NAD(P)/FAD-dependent oxidoreductase; *Streptomyces davaonensis*; identity, 99%
2238	Acyl carrier protein	ref|WP_015657265.1|; acyl carrier protein; *Streptomyces davaonensis*; identity, 90%
2239	4'-phosphopantetheinyl transferase (EC:2.7.8.-)	ref|WP_015657266.1|; 4'-phosphopantetheinyl transferase superfamily protein; *Streptomyces davaonensis*; identity, 95%
2240	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|APD72020.1|; type I polyketide synthase 1; *Streptomyces lunaelactis*; identity, 60%
2241	Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.8.1)	ref|CCK26874.1|; Isovaleryl-CoA dehydrogenase 2; *Streptomyces davaonensis* JCM 4913; identity, 93%
2242	Polyketide synthase type I	ref|WP_093719749.1|; acyltransferase domain-containing protein; *Streptomyces* sp. strain DI166; identity, 90%
2243	Regulatory protein dnrI	ref|WP_016468329.1|; MULTISPECIES: AfsR/SARP family transcriptional regulator; *Streptomyces*; identity, 71%
2244	Non-ribosomal peptide synthase	ref|WP_078598784.1|; fatty acyl-AMP ligase; *Streptomyces davaonensis*; identity, 93%

Table 7. Predicted function of genes from *Streptomyces* sp. P46 BGC 14. ORF2694, the predicted regulator we targeted for overexpression is highlighted in grey.

CDS	Protein and/or function	BLASTP best hit
2670	Cysteine desulfurase (EC 2.8.1.7)	ref|WP_015657679.1|; cysteine desulfurase; *Streptomyces davaonensis*; identity, 96%
2671	FIG01122820: hypothetical protein	ref|WP_037677709.1|;hypothetical protein; *Streptomyces griseus*; identity, 86%
2672	FIG01127629: hypothetical protein	ref|SBT93913.1|; hypothetical protein GA0115233_107426; *Streptomyces* sp. DI166; identity, 86%
2673	Transcriptional regulator, TetR family	ref|WP_015657682.1|; TetR family transcriptional regulator; *Streptomyces davaonensis*; identity, 89%
2674	hypothetical protein	ref|WP_048819859.1|; hypothetical protein; *Streptomyces ipomoeae*; identity, 84%
2675	Transcriptional regulator, DeoR family	ref|WP_015657683.1|; YafY family transcriptional regulator; *Streptomyces davaonensis*; identity, 89%

2676	TesB-like acyl-CoA thioesterase 3	ref|WP_093723330.1|; thioesterase family protein; *Streptomyces* sp. strain DI166; identity, 94%
2677	putative membrane protein	ref|CCK27304.1|; UPF0126 membrane protein; *Streptomyces davaonensis* JCM 4913; identity, 95%
2678	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.41)	ref|WP_007461640.1|; 3-oxoacyl-ACP synthase III; *Micromonospora lupini*; identity, 61%
2679	hypothetical protein	ref|SCF71788.1|; hypothetical protein GA0115260_101676; *Streptomyces* sp. strain MnatMP-M27; identity, 46%
2680	Dihydrolipoamide acetyltransferase component (E2) of acetoin dehydrogenase complex (EC 2.3.1.-)	ref|WP_003977562.1|; MULTISPECIES: acyltransferase; *Streptomyces*; identity, 59%
2681	Acetoin dehydrogenase E1 component beta-subunit (EC 1.2.4.-)	ref|WP_125633525.1|; alpha-ketoacid dehydrogenase subunit beta; *Streptomyces* sp. strain KPB2; identity, 74%
2682	Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)	ref|WP_076973300.1|; thiamine pyrophosphate-dependent dehydrogenase E1 component subunit alpha; *Streptomyces* sp. strain M1013; identity, 79%
2683	hypothetical protein	ref|WP_093603730.1|; MbtH family NRPS accessory protein; *Lentzea waywayandensis*; identity, 52%
2684	MbtH-like protein	ref|WP_005464209.1|; MbtH family NRPS accessory protein; *Saccharomonospora glauca*; identity 75%
2685	Macrolide-efflux protein	ref|WP_071803109.1|; MFS transporter; *Couchioplanes caeruleus*; identity 53%
2686	Non-ribosomal peptide synthetase	ref|WP_046563512.1|; hypothetical protein; *Micromonospora* sp. strain HK10; identity 49%
2687	Thioesterase in siderophore biosynthesis gene cluster	ref|WP_084262887.1|; thioesterase; *Actinomadura formosensis*; identity, 48%
2688	putative non-ribosomal peptide synthetase	ref|AKA59480.1|; non-ribosomal peptide synthetase; *uncultured bacterium* AZ_40; identity, 52%
2689	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_071803111.1|; type I polyketide synthase; *Couchioplanes caeruleus*; identity, 56%
2690	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_099933868.1|; amino acid adenylation domain-containing protein; *Streptomyces* sp. strain 70; identity, 51%
2691	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	ref|WP_065489939.1|; 4-hydroxyphenylpyruvate dioxygenase; *Streptomyces* sp. strain PTY087I2; identity, 60%

2692	(S)-2-hydroxy-acid oxidase (EC 1.1.3.15)	ref|WP_125938378.1|; aminotransferase class I/II-fold pyridoxal phosphate-dependent enzyme; *Streptomyces* sp. strain WAC 06738; identity, 70%
2693	Arogenate dehydrogenase (EC 1.3.1.43)	ref|WP_014061549.1|; prephenate dehydrogenase; *Streptomyces* violaceusniger; identity, 62%
2694	transcriptional regulator, SARP family	ref|WP_123558715.1|; AfsR/SARP family transcriptional regulator; *Micromonospora* sp. HM5-17; identity, 56%

Table 8. Predicted function of genes from *Streptomyces* sp. P46 BGC 28. ORF8145, the predicted regulator we targeted for overexpression is highlighted in grey.

CDS	Protein and/or function	BLASTP best hit
8145	StaQ	ref|OSY46292.1|; Transcriptional regulator NovG; *Streptomyces platensis*; identity, 51%
8146	Crotonyl-CoA carboxylase/reductase, ethylmalonyl-CoA producing	ref|WP_015655465.1|; crotonyl-CoA carboxylase/reductase; *Streptomyces davaonensis*; identity, 92%
8147	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_015655464.1|; type I polyketide synthase; *Streptomyces davaonensis*; identity, 84%
8148	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_015655463.1|; type I polyketide synthase; *Streptomyces davaonensis*; identity, 85%
8149	Non-ribosomal peptide synthetase	ref|WP_015655462.1|; non-ribosomal peptide synthetase; *Streptomyces davaonensis*; identity, 88%
8150	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	ref|WP_106435855.1|; type I polyketide synthase; *Streptomyces davaonensis*; identity, 83%
8151	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	ref|WP_015655460.1|; beta-ketoacyl-ACP reductase; *Streptomyces davaonensis*; identity, 89%
8152	Siderophore biosynthesis non-ribosomal peptide synthetase modules	ref|WP_015655459.1|; non-ribosomal peptide synthetase; *Streptomyces davaonensis*; identity, 86%
8153	putative cytochrome P450 hydroxylase	ref|WP_015655458.1|; cytochrome P450; *Streptomyces davaonensis*; identity, 92%
8154	major facilitator superfamily MFS_1	ref|WP_015655457.1|; MFS transporter; *Streptomyces davaonensis*; identity, 89%

3.6 Construction of the overexpression plasmid

Through the use of the conjugative and integrative plasmid pSK152 optimised for overexpression, we attempted to activate cryptic BGCs in P46 that may potentially lead to the production and discovery of novel bioactive compounds. To construct the overexpression plasmid, the purified regulator (ORF2243 – 831 bp, ORF2694 – 792

bp, ORF8145 – 1035 bp), terminator (56 bp) and *bldA* (502 bp) template (Figure 23A) were assembled using the primers 2243/2694/8145-F & *bldA*-R (Appendix, Table S1). Only combined gene templates 2243-terminator-*bldA* (1567 bp) and 8145-terminator-*bldA* (1771 bp) were cloned successfully. The purified amplicons (Figure 23B) were then cloned into *Not*I pre-digested pSK152 plasmid to produced pSK152-2243-terminator-*bldA* and pSK152-8145-terminator-*bldA* plasmids. Only pSK152-2243-terminator-*bldA* and pSK152-8145-terminator-*bldA* mutant plasmids were cloned successfully, pSK152-2694-terminator-*bldA* was not.

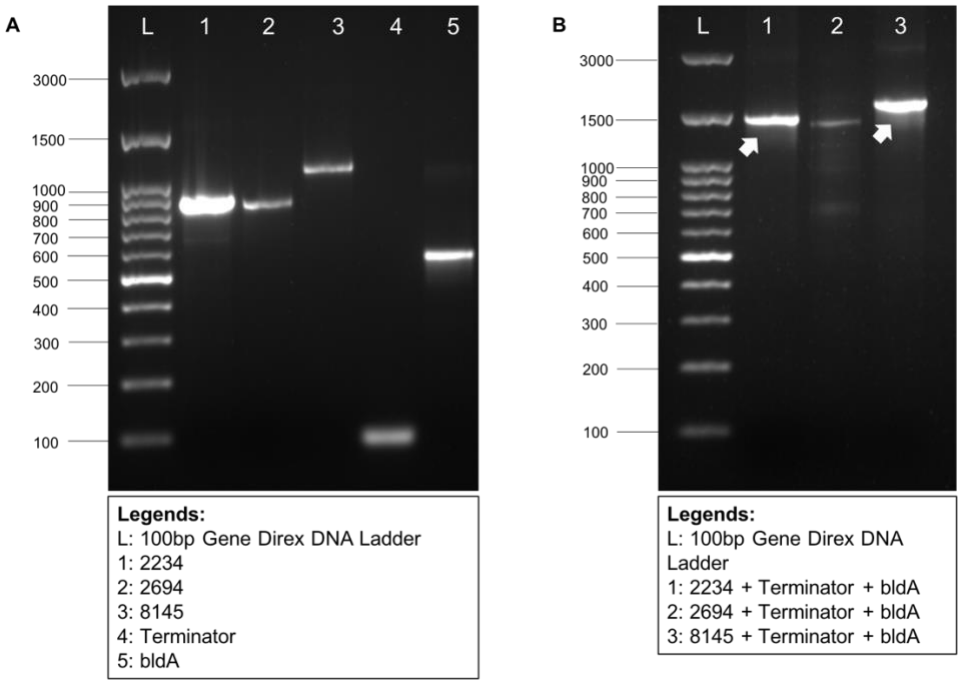


Figure 23. PCR amplicons of **(A)** the purified templates and **(B)** the combined regulator-terminator-*bldA* gene construct. The white arrows indicate successfully cloned amplicon.

The constructed plasmids were first transformed into NEB 10-beta *E. coli* and transformed colonies were screened with the primer pairs Terminator-F & *bldA*-R (Figure 24). The expected amplicon size was 558 bp and three colonies from mutants P46-C12ACT and three colonies from mutants P46-C28ACT were successfully transformed. This successful transformation was then confirmed through sequencing.

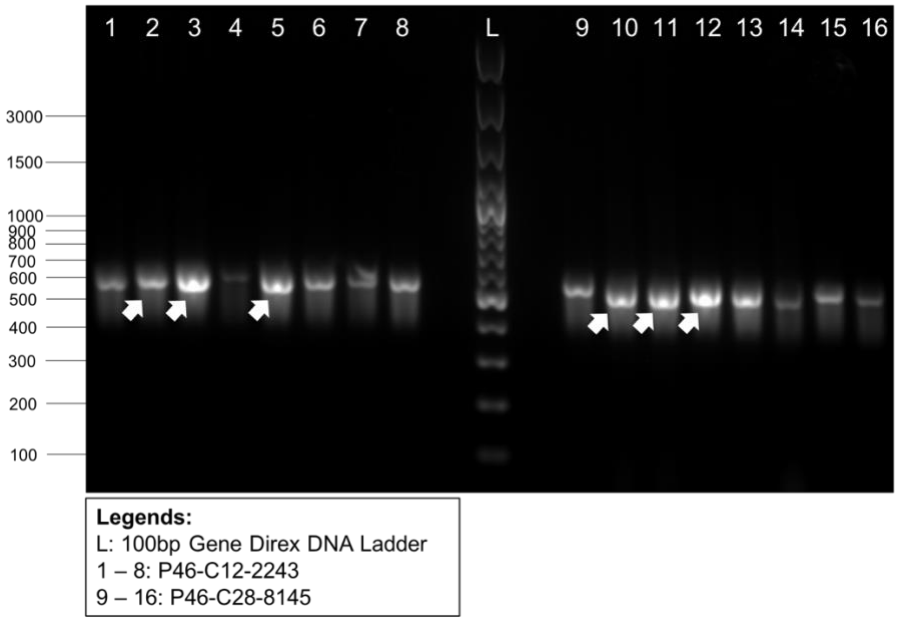
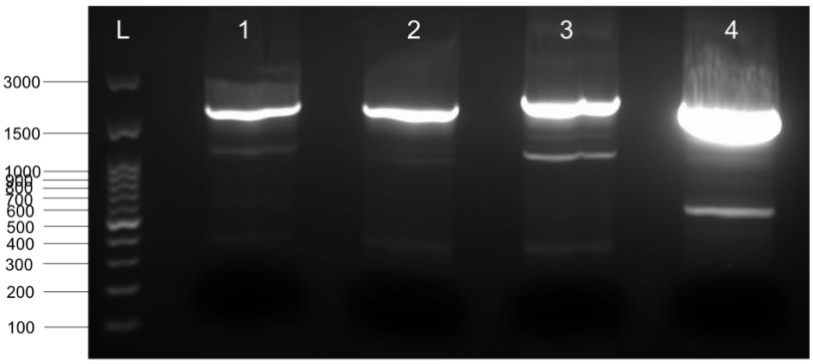


Figure 24. Colony PCR after transformation of constructed plasmid into NEB 10-beta *E.coli* using the primers Terminator-F & *bldA*-R. The white arrows indicate successfully cloned amplicon.

The validated plasmids were transformed into *E. coli* ET12567/pUZ8002 (Paranthaman & Dharmalingam, 2003) before being transferred into the P46-WT strain via intergenic conjugation. The P46 strains that harbour the overexpression plasmid (i.e. P46-C12ACT and P46-C28ACT) were screened by PCR using the primers pairs 2243/8145-F and *bldA*-R respectively. The expected amplicon size were 1567 bp and 1771 bp for P46-C12ACT and P46-C28ACT respectively (Figure 25). Two mutants for each P46 strain had successfully undergone conjugation (P46-C12ACT-1 and -2, P46-C18ACT-1 and -2). This successful conjugation was subsequently validated by sequencing. Concurrently, the successfully conjugated P46 mutants were screened for the integration of the plasmid into the P46 genome using the primers pairs ΦC31-F and -R. The expected amplicon size was 515 bp (Figure 26) and the plasmid had successfully integrated in all mutant strains, producing two P46-C12ACT and P46-C28ACT mutants each.

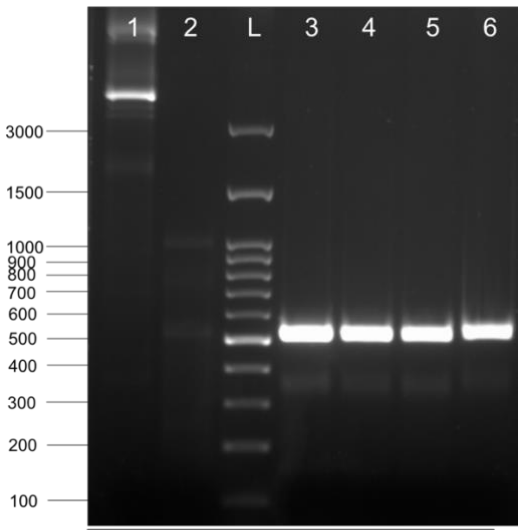
Cloning of the pSK152-2694-terminator-*bldA* plasmids was optimised by restriction digestion followed by T4 ligation instead of the Gibson Assembly method. The 2694-terminator-*bldA* fragment of 1462 bp was successfully constructed (Figure 27) and subsequently digested using restriction enzymes *NotI* and *EcoRV*. The pSK152 plasmid similarly was digested using the same restriction enzymes and was treated with alkaline diphosphatase to prevent the re-circularisation of the empty plasmid.

After the T4 ligation, the pSK152-2694-terminator-*bldA* plasmids were transformed into NEB 10-beta *E. coli* and transformed colonies were screened with the primer pairs pSK152-F & -R (Figure 28). However, although the expected amplicon size was 1728 bp and two colonies from mutants P46-C14ACT were successfully transformed, sequencing results indicated that the mutant plasmid was not cloned successfully.



Legends:
L: 100bp Gene Direx DNA Ladder
1: P46-C12-2243-1
2: P46-C12-2243-2
3: P46-C28-8145-1
4: P46-C28-8145-2

Figure 25. Validation of successful conjugation of *E.coli* ET12567/pUZ8002 with P46 using the primers 2243/8145-F and *bldA*-R.



Legends:
L: 100bp Gene Direx DNA Ladder
1: pSK152 plasmid
2: p46 Genome
3: P46-C12-2243-1
4: P46-C12-2243-2
5: P46-C28-8145-1
6: P46-C28-8145-2

Figure 26. Validation of plasmid integration into the P46 chromosome using the primers ΦC31-F and -R.

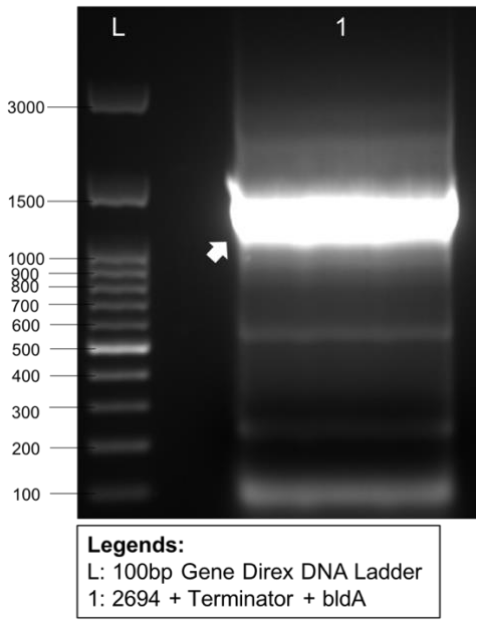


Figure 27. PCR amplicons of combined 2694-terminator-*bldA* gene construct using the primers 2694-F and *bldA*-R.

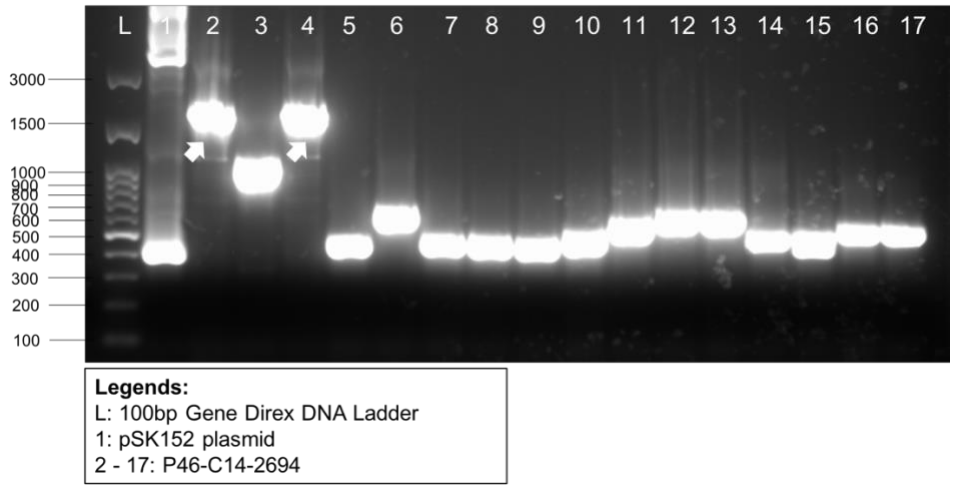


Figure 28. Colony PCR of constructed plasmid into NEB 10-beta *E.coli* using the primers 2694-F and *bldA*-R. The white arrows indicate successfully cloned amplicon.

Only two mutants of P46 were generated through successful intergenic conjugation and plasmid integration into the P46 genome. P46-C12-ACT and P46-C28-ACT are the mutants for which BGCs 12 and 28 are activated in P46 respectively. Metabolite profiling of these mutants will be subsequently studied. Other efforts are currently in place to successfully clone the last mutant P46-C14ACT.

3.7 Metabolite profiling of the overexpression mutants

HPLC was used to identify compounds present in the combined ethyl acetate and acetone extract of P46 fermented in various media. This time, we compare the profiles of P46 WT, P46-C12ACT, P46-C28ACT and the control (liquid medium only). Compounds with unique retention times and corresponding UV-Vis spectra that appeared in the mutant strains but not the wild type were identified.

The HPLC chromatograms of the crude extracts of P46 cultured in GYM media are shown in Figure 29. The P46-C12ACT mutant did not seem to produce any new compounds that were already discovered from the P46 WT culture. Surprisingly, the HPLC chromatogram of P46-C12ACT had fewer peaks present compared to the wild type. Several compounds that were previously produced in the P46 WT were not produced in the P46-C12ACT mutant. On the other hand, the P46-C28ACT mutant produced several compounds (including compounds **a** and **b**) that were not present in the culture of P46 WT in GYM. UV-Vis spectrum of compound **a** displayed a characteristic five absorption peaks with a UV_{max} at 222, 266, 304, 367 and 416 nm (Figure 33A); however, its mass has not been determined at the moment. Compound **b** displayed three absorption peaks with a UV_{max} at 218, 340, 363 nm (Figure 33B) and it has been discovered to have at least 12 analogues that have been isolated and further separated (Figure 30). The m/z of analogues of compound **b** are deduced by LC-MS and displayed in Table 9, although some have not been detected yet.

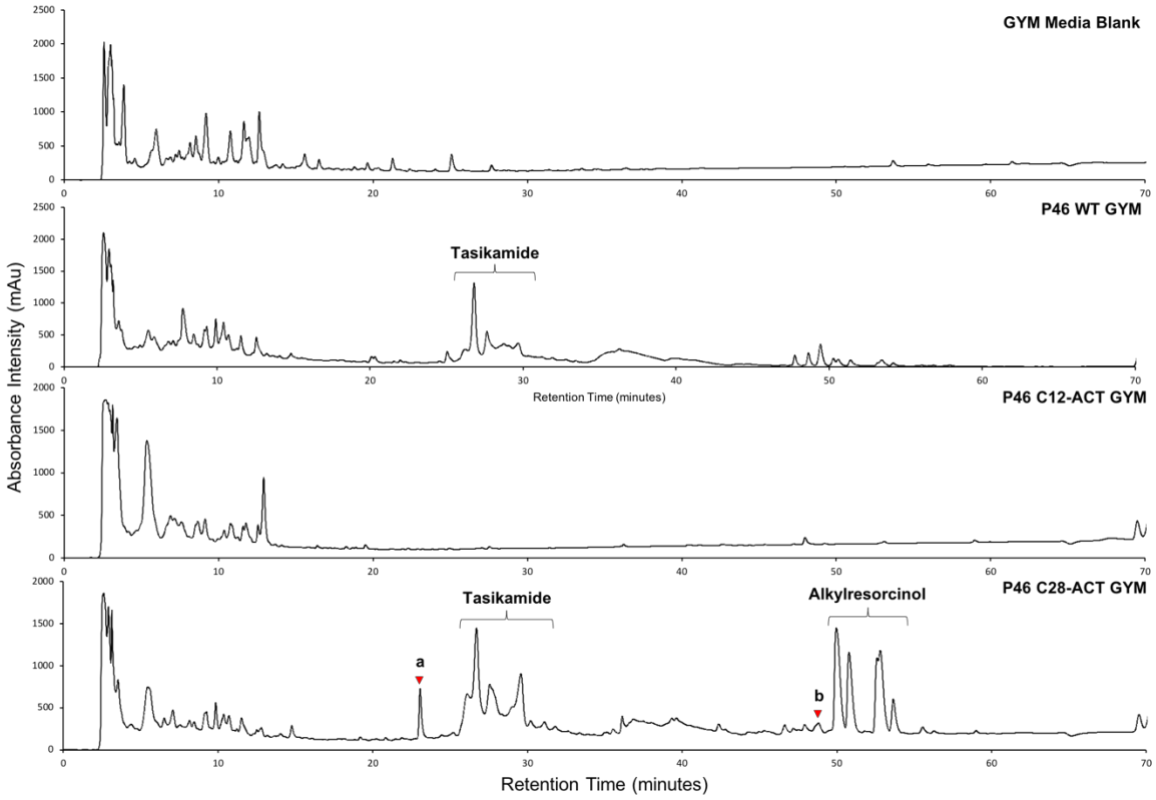


Figure 29. Comparison of HPLC chromatograms of P46 WT, P46 C12-ACT, P46 C28-ACT with GYM. Compounds with unique retention times and UV-Vis spectrum are highlighted.

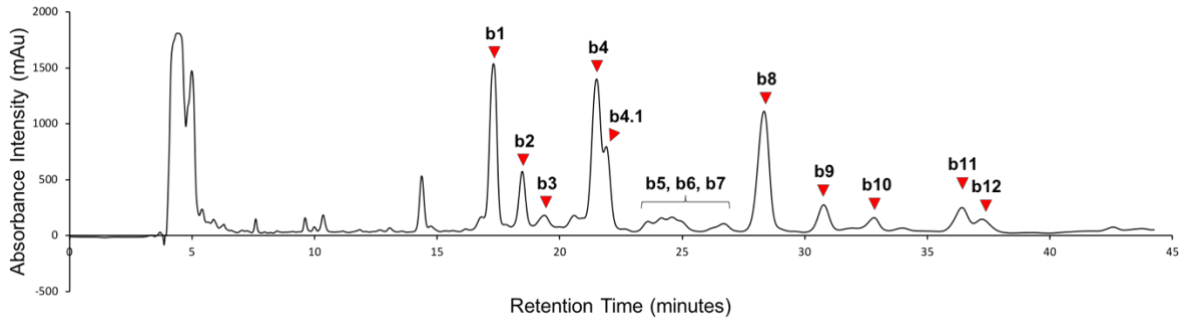


Figure 30. HPLC Chromatogram at 220 nm of compound **b** analogues (**b1** – **b12**).

Table 9. Corresponding UV_{max} and m/z data of compounds **b1**, **b2**, **b4, b9** and **b10**. The rest of the analogues have not been completely detected.

Potential Compounds	UV_{max} (nm)	m/z ($[M+H]^+$)
b1	218, 340, 363	1115.4167
b2	217, 339, 363	1115.4167
b4	218, 340, 363	1129.4167
b9	216, 338, 362	1143.4167
b10	216, 337, 361	1127.5000

The HPLC chromatograms of the crude extracts of P46 cultured in pharmamedia are shown in Figure 31. Similar to the profiles of GYM, the HPLC chromatogram of P46-C12ACT had fewer peaks present compared to the wild type. The P46-C12ACT mutant did not produce any compounds that were already previously discovered when P46 WT was cultivated in GYM. Instead, some compounds that were previously produced in the P46 WT were not produced in P46-C12ACT mutant. Significant amounts of compound **a** was produced by P46-C28ACT mutant that was not present in the culture of P46 WT. However, compound **b** was not produced by the P46-C28ACT mutant in this media.

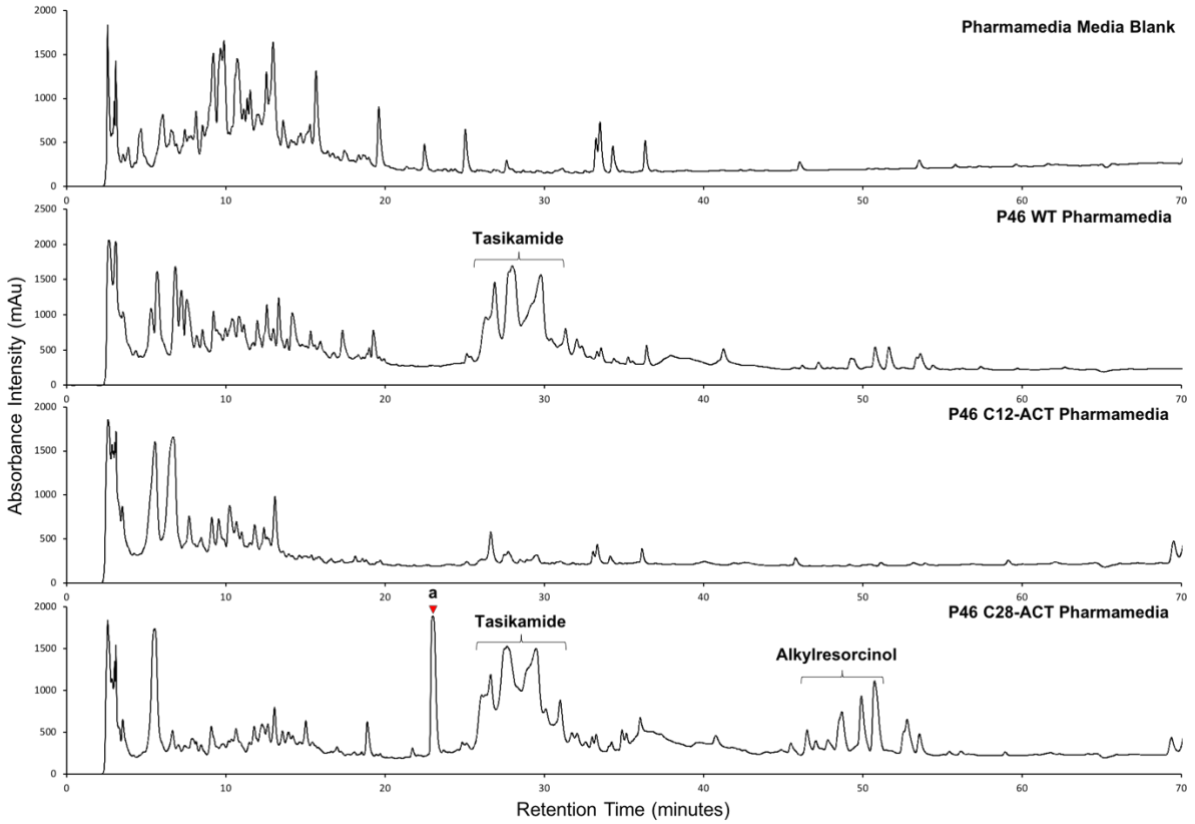


Figure 31. Comparison of HPLC chromatograms of P46 WT, P46 C12-ACT, P46 C28-ACT with Pharmamedia. Compounds with unique retention times and UV-Vis spectrum are highlighted.

The HPLC chromatograms of the crude extracts of P46 cultured in SG2 are shown in Figure 32. Unlike the profiles of GYM and pharmamedia, the P46-C12ACT mutant produced alkylresorcinol that was not produced by P46 WT under the same cultivation conditions. Significant amounts of compound **a** was produced by P46-C28ACT mutant that was not present in the culture of P46 WT while compound **b** was similarly produced by the P46-C28ACT mutant in this media.

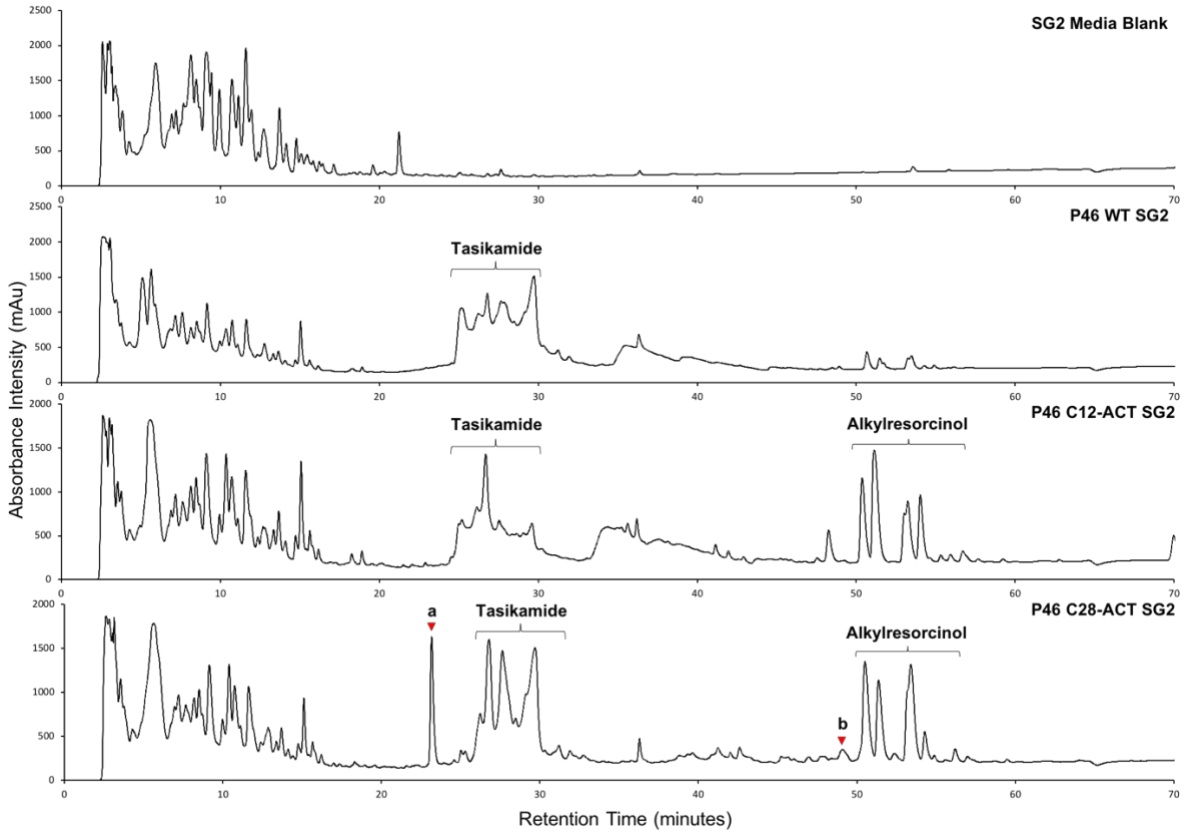


Figure 32. Comparison of HPLC chromatograms of P46 WT, P46 C12-ACT, P46 C28-ACT with SG2. Compounds with unique retention times and UV-Vis spectrum are highlighted.

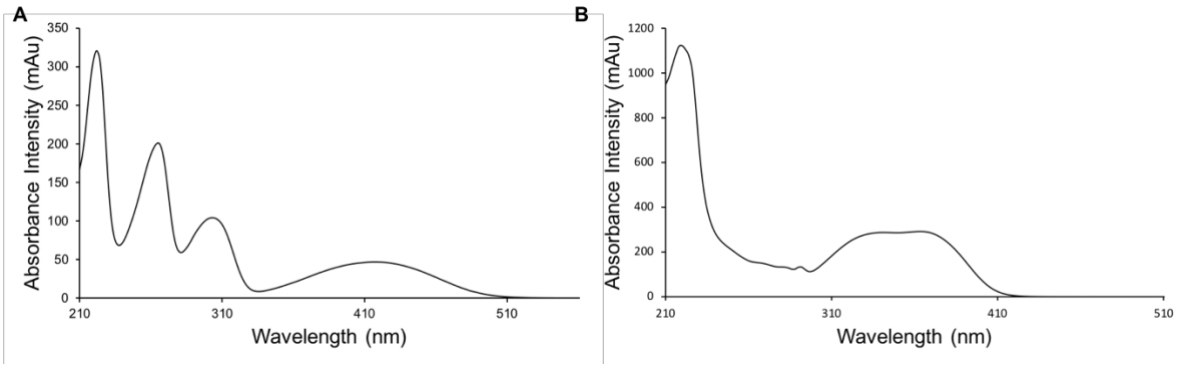


Figure 33. UV spectra of compound **(A) a** and **(B) b**, isolated from P46 C28-ACT Mutants in GYM, Pharmamedia and SG2.

3.8 Comparison of gene expression levels by RT-qPCR

Quantitative RT-PCR (RT-qPCR) was utilised to compare the expression level of the downstream genes as a result of an overexpression of our desired regulator. Based on the identified operon boundaries in **3.5**, a gene within each operon is selected for RT-qPCR studies. Expression levels were tested from the RNA of both the WT and the mutants after fermentation conditions of 10 d. The conditions used were similar to when the culture was tested for potential secondary metabolites. Each set of PCR comprises five samples – no template control, negative control (20× diluted RNA),

positive control (genomic DNA), WT cDNA and mutant cDNA in chronological order (Figures 34 – 36). Every PCR was accompanied by an internal control, the *hrdB* gene from *Streptomyces* sp. P46. Five genes were tested in the P46-C12ACT mutant – ORF2214, 2234, 2237, 2242 and 2243 while four genes were tested in the P46-C28ACT mutant ORF8145, 8146, 8151 and 8153.

PCR results for P46-C12ACT illustrated that after the overexpression of the regulator gene ORF2243 (Figure 35), the regulator gene ORF2243 itself that was initially not expressed in the WT was expressed in the mutant. However, downstream genes such as ORF2234 and ORF2242 witnessed the opposite change. These genes that were originally expressed in the WT were no longer expressed in the mutant. Genes ORF2214 and 2237 witnessed no expression in both the WT and the mutant (Figures 34 – 35).

PCR results for P46-C28ACT illustrated that after the overexpression of the regulator gene ORF8145 (Figure 36), the regulator gene ORF8145 itself witnessed a minor increase in expression levels in the mutant compared to the wildtype. Similarly, downstream genes such as ORF8146, ORF8151 and ORF8153 experienced an increase in expression levels in the mutant compared to the wildtype.

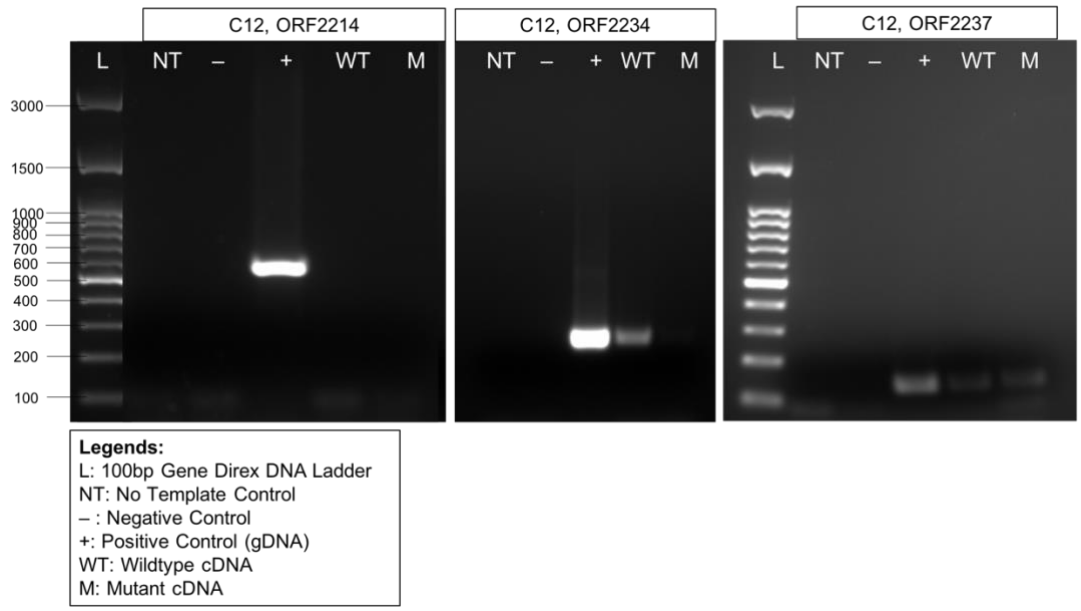


Figure 34. RT-PCR results depicting the level of expression of genes in the mutant P46-C12ACT compared to P46-WT. Genes tested are ORF2214, 2234 and 2237. Primers bind at 515 bp for gene 2214, 237 bp for gene 2234 and 115 bp for gene 2237.

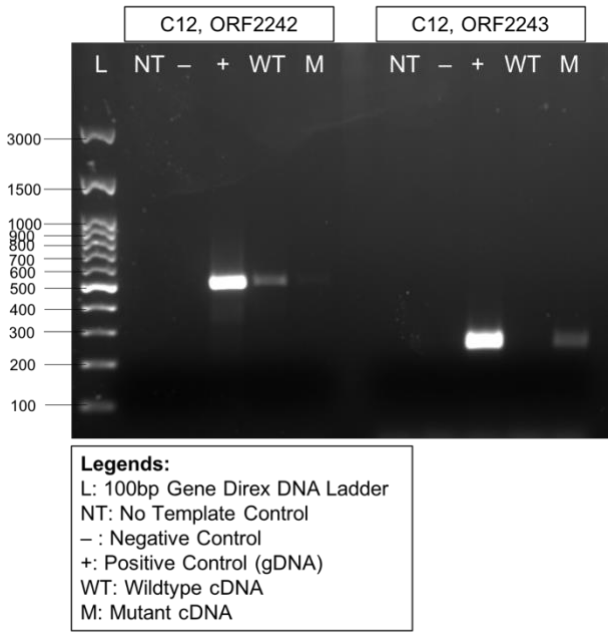


Figure 35. RT-PCR results depicting the level of expression of genes in the mutant P46-C12ACT compared to P46-WT. Genes tested are ORF2242 and 2243 (continued). Primers bind at 501 bp for gene 2242, 239 bp for gene 2243

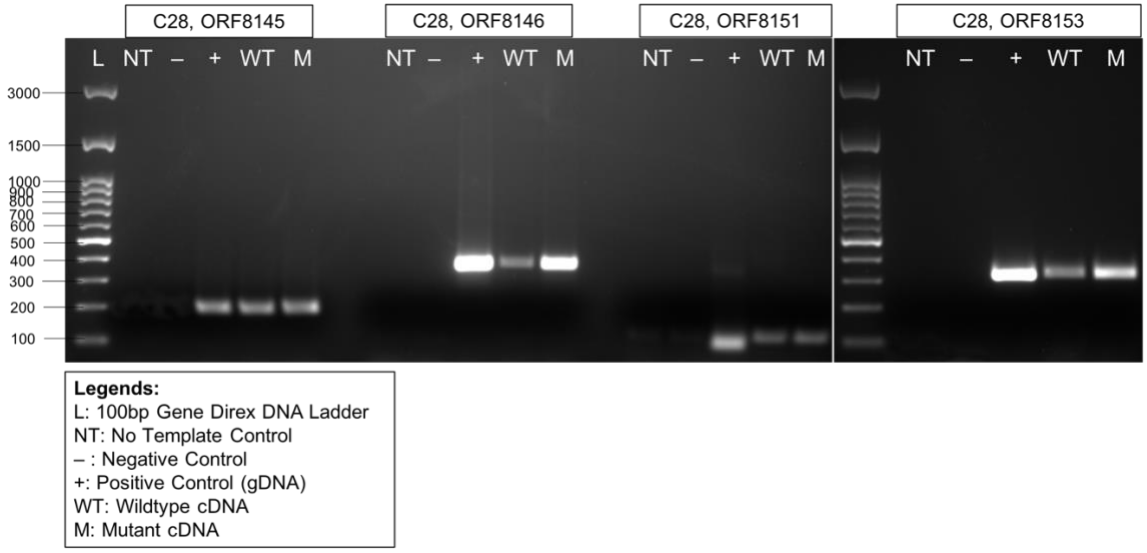


Figure 36. RT-PCR results depicting the level of expression of genes in the mutant P46-C28ACT compared to P46-WT. Genes tested are ORF8145, 8146, 8151 and 8153. Primers bind at 180 bp for gene 8145, 344 bp for gene 8146, 102 bp for gene 8151 and 310 bp for gene 8153.

The overexpression plasmid increased the expression of the targeted genes – ORF2243 and ORF8145. Although the overexpression of the regulator gene ORF8145 triggered an increase in expression of downstream genes such as ORF8146, ORF8151 and ORF8153, the overexpression of the regulator gene ORF2243 failed to trigger an increase in expression of downstream genes such as ORF2214, ORF2234, ORF2237 and ORF2242.

CHAPTER 4: DISCUSSION

4.1 General features of the P46 genome

Actinomycetes have proven to be a rich source of bioactive compounds with diverse applications in the pharmaceutical industries. In a survey of 100 most essential secondary metabolites elucidated from bacteria or fungi, 68% are produced by *Streptomyces* species, while 15% by other actinomycetes (Katz & Baltz, 2016). Additionally, of the 100 secondary metabolites surveyed, more than 60% were synthesised by NRPS, PKS or NRPS-PKS hybrid systems (Katz & Baltz, 2016). These studies have suggested that amongst the actinomycetes, there are specific strains which are termed as 'gifted' (Baltz, 2017). This term suggests that these strains possess the coding capacity to produce more secondary metabolites than others. BGCs of NRPS origins often contain *mbtH* homologs that are responsible for the expression of non-enzymatic chaperone proteins that can enhance adenylation reactions during peptide assembly. The number of *mbtH* homologs has proven to be a good estimate of the number of NRPS clusters. Actinomycetes encode on average 1.1 *mbtH* homologs per genome while *Streptomyces* species 2.6 per genome (Baltz, 2011). It was shown that both the number of *mbtH* homologs and secondary metabolite coding capacity of a species are positively correlated with the genome size. Secondary metabolite coding capacity increases about 20-fold from about 0.15 to 3.0 Mb per genome and actinomycetes with genomes of larger than 8 Mb are considered to be 'gifted' (Baltz, 2017). Genome annotation by antiSMASH analysis discovered that P46 has a large genome size of 9.8 Mb and it possesses 2 *mbtH* homologs in its genome. This seems to suggest that P46 has the potential to be a 'gifted' strain with high coding capacity for secondary metabolites production, that could additionally be novel. Such discovery definitely spurs our efforts for genome mining and novel compound elucidation from P46.

This study has demonstrated that with genome annotation, as well as secondary metabolite and BGC profiling, the types of natural products P46 can potentially produce can be more accurately predicted. BGC profiling of P46 was consistent with preliminary findings that actinomycetes possess an important untapped resource of valuable natural products (Schorn et al., 2016). Having analysed the HPLC chromatograms of P46 WT, we have identified two compounds that possess unique

retention times and UV-Vis spectra that were not present in the media – 1-deoxypentalenic acid and alkylresorcinol. The m/z of alkylresorcinol was calculated by LC-MS to be 321.4167 $[M+H]_+$ and confirmed by the m/z of alkylresorcinols base ion of 316 $[M+H]_+$ (Wierzbicka et al., 2015). The enzymes and proteins involved in the biosynthetic pathway of alkylresorcinol correlates with BlastP analysis of predicted protein products of genes present in BGC 2. Additionally antiSMASH KnownClusterBlast analysis has revealed 100% homology of BGC 2 with *Streptomyces griseus* NBRC 13350. Hence, alkylresorcinol is deduced to be the secondary metabolite produced by genes present in BGC 2.

Additionally, 1-deoxypentalenic acid was analysed by LC-MS to have a calculated m/z = 235.162 $[M+H]_+$ and is postulated to be, a co-metabolite of the antibiotic pentalenolactone which calculated m/z = 262 $[M+H]_+$ (You et al., 2007), supported by preliminary 1H NMR data. Similarly, correlating the enzymes and proteins that are involved in the biosynthetic pathway of both 1-deoxypentalenic acid and pentalenolactone from antiSMASH, 1-deoxypentalenic acid is deduced to be the secondary metabolite produced by genes in BGC 1. Correlating known compounds isolated from P46 WT subcultured in various media through the OSMAC approach and our antiSMASH table of predicted secondary metabolites (Table 2), it confirmed the presence of a large number of cryptic BGCs in P46. Hence, our efforts to activate them could potentially lead to the discovery of novel bioactive natural products that are clinically valuable.

Understanding P46's phylogeny could also add insights to our genome mining approach (Ziemert & Jensen, 2012). 16S ribosomal RNA (rRNA) sequences and DNA-DNA hybridisation (DDH) have been conventionally adopted as a method to define bacterial classification (Kim et al., 2014). The threshold of 16S rRNA gene sequence similarity is >95% for two strains to belong to a distinct genus and >98.65% to belong to the same species. DDH is additionally used as confirmation when the similarity percentage of 16S rRNA is >97% (Busse et al., 2010). However, genome annotation has revealed that P46 has six copies of 16S rRNA and because of the existence of intragenomic variety, each copy has a different similarity readout. Although according to 16S rRNA evaluation, it revealed the closest relative of P46 is *Streptomyces reticuli* TUE45 with a similarity of 100% (according to NCBI

sequences) (Pang, unpublished) and multiple other relatives with the same similarity scores. This has been an issue with 16S rRNA evaluation where it has a poor resolution power at a species level (Janda & Abbott, 2007), hence, can only accurately assign the genus of P46 to be *Streptomyces*. AntiSMASH analysis of BGCs in P46 revealed that most clusters are instead similar to *Streptomyces davaonensis*. *S. davaonensis* has 31 BGCs, of which 15 are predicted to be responsible for the production of NRPS, PKS or NRPS-PKS hybrid compounds. Twelve out of 31 BGCs are predicted to produce the same compounds as that of BGCs in P46. Such discovery could possibly be due to the prevalence of horizontal gene transfer (HGT) events between prokaryotes (Poretsky et al., 2014), causing this discrepancy. The limited resolution of the 16S rRNA gene in differentiating closely related species as well as HGT events being unaccounted for, suggests that metaproteomic based techniques like antiSMASH would instead be a more reliable tool in the prediction of genes (Ranjan et al., 2016). Hence, BGCs in *S. davaonensis* can provide a more accurate reference for the predicted protein products that can be produced by the BGCs in P46.

4.2 Tasikamide and its analogues

My lab has identified a potentially novel bioactive compound and named it Tasikamide A. Although it was first isolated from the fermentation culture of P46 WT in the media GYM, through the OSMAC approach, we were able to isolate Tasikamide A from various other media such as pharmamedia, SG-2, GYM agar, Gauze and SCN, in varying amounts. GYM media was chosen as an optimal media for large scale fermentation for the production of Tasikamide A. NMR was utilised to deduce the chemical structure of Tasikamide A and revealed its amino acid composition to be: hydroxyphenylalanine, isoleucine, N-methylphenylalanine, O-methyltryptophan, baba and a 4 carbon linker and its calculated m/z = 854.39 $[M+H]^+$. Characterised as a nonribosomal peptide, the m/z of Tasikamide A matches the m/z of the skeleton of the predicted NRPS protein product of BGC 20 (m/z = 838.971 $[M+H]^+$). BGC 20 predicted to produce a compound with PKS-NRPS hybrid also supports our structural analysis of Tasikamide A's chemical structure where NRPS modules tend to incorporate non-proteinogenic amino acids such as N-methylphenylalanine and BABA into the products (Sieber & Marahiel, 2005).

NMR analysis has revealed the two phenylalanine amino acids present in the structure of Tasikamide A, of which one undergoes N-methylation. BLASTp anaylsis of BGC 20 has predicted the presence of a methyltransferase gene which can possibly support our understanding of the structure of Tasikamide A (Figure 15). BLASTp analysis has also revealed the presence of an O-methyltransferase gene which can account for the O-methylated tryptophan also present in the structure of Tasikamide A (Figure 15). The predicted presence of cytochrome P450 genes, also identified through the BLASTp analysis, is responsible for the beta position hydroxylation of phenylalanine (Fitzpatrick, 2003) in the structure of Tasikamide A (Figure 15). AntiSMASH gene cluster analysis has predicted the presence of 9 adenylation domains present in the NRPS gene cluster in BGC 20 which can also account for the incorporation of the number of amino acids to form Tasikamide A, a nonribosomal peptide. The concurrent study of the chemical structure of Tasikamide A with the cluster analysis of BGC 20 and its homologous clusters can bring us closer to fully comprehending the biosynthetic mechanism of Tasikamide A and confirm our predictions.

Tasikamide A was also predicted to be a novel potential bioactive compound. Bioassays tested Tasikamide A against cancer cell lines MIA PaCa-2 and A549 and its IC_{50} is determined to be 40 µM, showing relatively strong anticancer potential, characterising Tasikamide A to be a moderately active compound (Méndez-Lucio & Medina-Franco, 2017). Also, the analysis of the chemical structure of Tasikamide A, displayed a large absolute molecular size, which could also suggest higher affinity to their binding targets which could translate into higher bioactivity (Schuffenhauer et al., 2006). These pharmacological properties of Tasikamide A can be an area of further study. It is observed in the HPLC profiling that Tasikamide A and its analogues remain consistently produced across multiple culture conditions. Its consistent production suggests that Tasikamide A could be the main product that is produced by P46, enhancing its ecological fitness in its native environment, the quarry.

4.3 Other potential secondary metabolites in P46 WT through OSMAC

The use of dereplication tools such as HPLC, MS and NMR demonstrated that subculturing P46 in various media (CRM, gauze, GYM, HAV, oatmeal, pharmamedia, SCN and YPD) allowed for the identification of potentially novel compounds that are present in the crude extracts. Analysing the HPLC chromatograms of P46 WT, we have identified compounds that have a unique retention time and UV-Vis spectrum that were not present in the media control.

A total of seven compounds (**c, d, e, f, g, h, j**) and their analogues have been isolated from fermentation cultures of P46 WT in the various media mentioned in **3.4**. Although we have isolated these compounds and elicited their UV-Vis spectrum through preliminary HPLC analysis, identifying them would require determining their m/z and structure through LC-MS and NMR respectively. P46 WT was discovered to show a preference towards glucose as their carbon source, and yeast extract and soytone as their nitrogen sources for biomass and secondary metabolite production. Casein was another nitrogen source that was proven by other *Streptomyces* strains that triggered maximum secondary metabolite production but was only moderately effective to P46 (Al-ghazali & Omran, 2017). Culturing P46 on agar media altered the secondary metabolites profiles although cultivation time had to be extended for sufficient compound production yield compared to liquid media. We also discovered that the production of secondary metabolites by P46 WT is extremely sensitive to fermentation conditions. Minor alterations in the composition of specified ingredients in a media would result in changes of secondary metabolites produced. Such sensitivity poses us a challenge to optimise the culture conditions for each media used for the fermentation of P46.

Supplementation of media with trace metals such as zinc, copper and manganese, etc. at low concentrations has also been discovered to influence secondary metabolite production (Locatelli et al., 2016). These trace metals are micronutrients which affect morphological development and secondary metabolism of *Streptomyces* playing the role of effectors of metalloregulators, cofactors of natural product biosynthesis enzymes and form functional complexes with secondary metabolites (Locatelli et al., 2016). Fermentation of P46 with GYM media supplemented with trace metals such as manganese, copper, iron and zinc (L-media) has revealed

potential novel secondary metabolites that are currently being elucidated. This discovery suggests that similar modification of other base media could potentially result in the activation of silent BGCs and production of novel secondary metabolites.

4.4 Overexpression of regulators in P46 mutants

The use of genomic tools helps to circumvent the significant historical problem of the rediscovery of known compounds by focusing on the new and novel BGCs. Having identified known compounds from the antiSMASH analysis of P46, we highlighted cryptic BGCs of NRPS, PKS and NRPS-PKS hybrids origins that were predicted to produce potentially novel compounds. By utilising an integrative-conjugative plasmid optimised for overexpression, we attempted to activate these cryptic BGCs by overexpressing a predicted positive regulator.

From comparing the HPLC profiles of the P46 WT and mutants, successful overexpression of regulator ORF8145 in P46-C28ACT have resulted in the consistent production of compound **a** in all three media (GYM, Pharmamedia and SG2) tested. The second potentially novel compound, **b** was only produced by P46-C28ACT in two media (GYM and SG2). Cross-referencing the m/z and UV-Vis spectrum of compound **a**, it is postulated to be furaquinocin A (Figure S1) (Kumano et al., 2010) of BGC 26. Compound **b**'s m/z and structure is still being confirmed and elucidated.

Conversely, successful overexpression of regulator ORF2243 in P46-C12ACT did not result in the production of any different compounds from that by P46 WT. Compounds produced by P46 WT were instead not detected in the culture media of P46-C12ACT. RT-PCR results confirmed these findings such that in P46-C12ACT, genes downstream witnessed a fall in expression compared to the P46 WT while in P46-C28ACT, genes downstream witnessed an increase in expression compared to the P46 WT. Upon BLASTp analysis, it revealed that ORF2243 is predicted to code for an AsfR or SARP family transcriptional regulator. The SARP family of transcriptional regulators have been classified as a family of pathway-specific transcriptional activators. Their positive regulatory effect has been confirmed by the increase in yield of undecylprodigiosin and actinorhodin biosynthesis through

overexpression of SARP family regulators (Bruheim et al., 2002). However, AsfR encodes a LysR-type regulator protein which could have a negative regulatory effect on downstream genes (Vermeij et al., 1999). This effect could have been the overriding regulation of the downstream genes of P46-C12ACT mutants, resulting in the lowered expression and cleaner HPLC profiles observed. ORF2243 could also instead solely code for an AsfR regulatory protein which would result in an isolated negative regulatory effect.

On the other hand, ORF8145 is predicted to code for a NovG transcriptional regulator. NovG has been discovered to be a positive regulator of novobiocin biosynthesis in *Streptomyces spheroids* (Eustaquio, 2005). NovG could have a similar positive pathway-specific regulatory effect on the genes downstream in P46-C28ACT, which resulted in the increased production of several compounds compared to the wildtype. The overexpression of transcriptional regulator NovG could have resulted in the activation of BGC 28 and stimulated the production of compound **b**. However, further studies on its m/z and structure, and the confirmation through rigorous study of the biosynthetic pathways in BGC 28 would be required.

The yield of compound **a** was also enhanced during the fermentation of P46-C28ACT. Complete molecular genetic studies of the staQ regulatory protein that is predicted to be encoded by ORF8145 have yet to be complete. However, from the observations, we hypothesised that staQ might be able to bind to regulatory genes such as AsfR/SARP family transcriptional regulator (ORF7679 and ORF7680), LuxR family transcriptional regulator (ORF7681) or LacI family transcriptional regulators (ORF7679) present in BGC 26, leading to simultaneous production of both compounds **a** and **b** by P46-C28ACT. Although compound **a** has already been previously discovered in other media, it was observed that there is a significant increase in yield of compound **a** consistently in the three media tested. Under such fermentation conditions, it provides us with an optimised approach in the isolation of compound **a** (furaquinocin A) compared to traditional large scale fermentation. These results demonstrated that the overexpression of a single positive regulator might result in the activation of more than one BGC by affecting other biosynthetic pathways. Hence, this has definitely contributed to our continued understanding of

the complexity of biosynthesis pathways and regulation of secondary metabolites in P46, highlighting considerations for future attempts of genetic engineering.

CHAPTER 5: CONCLUSIONS AND FUTURE WORK

My research project focus on the genome-guided discovery of natural products from a biosynthetically talented *Streptomyces* strain utilising tools from the merging disciplines of microbiology and natural product chemistry. The employment of genome mining, fermentation techniques, and highly efficient analytical tools to identify natural products, is vital for exploring the full biosynthetic potential of the actinomycetes strain. By demonstrating the potential of *Streptomyces* sp. P46 in producing novel natural products, we have also highlighted the importance of tapping on under-explored environments such as quarries for natural product discovery. Through the combination of both the OSMAC approach and a genetic method, genome mining data can be translated into novel secondary metabolites. The chemical analysis, genome prediction and the overexpression of regulators are equally essential in this project. The OSMAC approach is useful for the preliminary understanding of all potential secondary metabolites that P46 can produce (be it already characterised or not). The genome-guided approach can then work on decreasing the rate of rediscovering known natural products from P46 and streamline the discovery process, allowing us to focus our efforts onto novel compounds. Additionally, the overexpression of local regulators in P46 allowed for the discovery of potentially novel secondary metabolites by activating the BGCs that were initially silent or increasing the yield of secondary metabolite produced that was initially undetectable. Moreover, the overexpression of a single regulator could lead to the activation of multiple cryptic BGCs, suggesting cross activation of BGCs. This complexity can be deeper understood through continuous studying of the biosynthetic mechanism involved in the production of secondary metabolites by the BGCs of interest. All three approaches work best in tandem in pursuit of novel natural products produced by P46.

One of our goals is to uncover bioactive natural products to fill current drug-lead pipelines. Hence, the compounds we discovered (e.g. Tasikamide A) and the others that are in the process of being characterised will be subjected to various assays such as antifungal, antibiofilm and anticancer assays to assess their bioactivity and safety further. The compounds produced by P46 can be tested through a Minimum Inhibitory Concentration (MIC) assay against *Candida albicans* to measure antifungal

characteristics. Additionally, the compounds can be tested against *Pseudomonas aeruginosa* PA01 to assess their ability to inhibit biofilm formation. All these bioactivity assays will then yield potentially interesting results beyond simply the structure of the compounds into possible commercially valuable properties. The secondary metabolites produced by P46, albeit yet to be fully characterised, will contribute to the current natural product space as brand new families of bioactive compounds or new analogues of known molecules.

In addition, we will continue with the ongoing efforts to activate BGC14 and other cryptic BGCs that are predicted to be of NRPS, PKS, or NRPS-PKS hybrid origins which could potentially produce novel bioactive compounds. Clusters 8, 21, 32 and 33 are such clusters that are worth pursuing in the future. The genetic tools validated in my study will be helpful for future students to activate cryptic BGCs by other genetic methods such as promoter replacement and refactoring of the BGCs. Qualitative RT-PCR results was an essential tool for a preliminary analysis of the effect of overexpressing the local regulators of interest. Quantitative RT-PCR will be pursued to offer us a more accurate measure of the extent of gene overexpression. RT-PCR results has allowed us to confirm the BLASTp prediction of functions of local regulators of interest. However, since the method of overexpressing the local regulator to activate BGC 12 was not successful, we are pursuing another activation method through the insertion of a strong and constitutive promoter upstream of the NRPS or PKS gene of interest using a CRISPR-Cas9 system. We are currently working towards generating mutants to compare the HPLC profiles and gene products of the activated mutants and the wildtype. This CRISPR-Cas9 system can also serve as an alternative method to activate BGC 14 as the initial methods of generating activated mutants of BGC 14 were not successful.

Lastly, the growing and readily accessible public genome databases are highly valuable treasure troves for the discovery of novel microbial natural products from many microbial strains from all over the world. The genome-guided approach discussed in my thesis is widely applicable and should be adopted by researchers in Singapore for the expedited discovery of structurally novel bioactive natural products.

REFERENCES

Abbas, A. S., & Edwards, C. (1990). Effects of metals on *Streptomyces coelicolor* growth and actinorhodin production. *Applied and Environmental Microbiology*, *56*(3), 675–680. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16348141

Al-ghazali, L. H., & Omran, R. (2017). Optimization of medium composition for antibacterial metabolite production from *Streptomyces* sp. *Asian Journal of Pharmaceutical and Clinical Research*, *10*(9), 381. https://doi.org/10.22159/ajpcr.2017.v10i9.19813

Alduina, R., & Gallo, G. (2012). Artificial chromosomes to explore and to exploit biosynthetic capabilities of Actinomycetes. *Journal of Biomedicine and Biotechnology*, *2012*, 1–10. https://doi.org/10.1155/2012/462049

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, *12*(1), 402. https://doi.org/10.1186/1471-2164-12-402

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Arul Jose, P., Sivakala, K. K., & Jebakumar, S. R. D. (2013). Formulation and statistical optimization of culture medium for improved production of antimicrobial compound by *Streptomyces* sp. JAJ06. *International Journal of Microbiology*, *2013*, 526260. https://doi.org/10.1155/2013/526260

Bachmann, B. O., Van Lanen, S. G., & Baltz, R. H. (2014). Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *Journal of Industrial Microbiology & Biotechnology*, *41*(2), 175–184. https://doi.org/10.1007/s10295-013-1389-9

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., Miao, J., … Zhang, L. (2015). Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proceedings of the National Academy of Sciences*, *112*(39), 12181–12186. https://doi.org/10.1073/pnas.1511027112

Baltz, R. H. (2011). Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *Journal of Industrial Microbiology & Biotechnology*, *38*(11), 1747–1760. https://doi.org/10.1007/s10295-011-1022-8

Baltz, R. H. (2016). Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. *Journal of Industrial Microbiology & Biotechnology*, *43*(2–3), 343–370. https://doi.org/10.1007/s10295-015-1682-x

Baltz, R. H. (2017). Gifted microbes for genome mining and natural product discovery. *Journal of Industrial Microbiology & Biotechnology*, *44*(4–5), 573–588. https://doi.org/10.1007/s10295-016-1815-x

Banerjee, P., Erehman, J., Gohlke, B.-O., Wilhelm, T., Preissner, R., & Dunkel, M. (2015). Super Natural II--a database of natural products. *Nucleic Acids Research*, *43*(Database issue), D935-9. https://doi.org/10.1093/nar/gku886

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J. P., … van Wezel, G. P. (2016). Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and Molecular Biology Reviews : MMBR*, *80*(1), 1–43. https://doi.org/10.1128/MMBR.00019-15

Bibb, M. J. (2005). Regulation of secondary metabolism in streptomycetes. *Current*

Opinion in Microbiology, *8*(2), 208–215. https://doi.org/10.1016/j.mib.2005.02.016
Bode, H. B., Bethe, B., Höfs, R., & Zeeck, A. (2002). Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem*, *3*(7), 619. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9
Boettger, D., & Hertweck, C. (2013). Molecular diversity sculpted by fungal PKS-NRPS hybrids. *ChemBioChem*, *14*(1), 28–42. https://doi.org/10.1002/cbic.201200624
Brakhage, A. A. (2013). Regulation of fungal secondary metabolism. *Nature Reviews Microbiology*, *11*(1), 21–32. https://doi.org/10.1038/nrmicro2916
Brakhage, A. A., & Schroeckh, V. (2011). Fungal secondary metabolites – strategies to activate silent gene clusters. *Fungal Genetics and Biology*, *48*(1), 15–22. https://doi.org/10.1016/j.fgb.2010.04.004
Bruheim, P., Sletta, H., Bibb, M. J., White, J., & Levine, D. W. (2002). High-yield actinorhodin production in fed-batch culture by a *Streptomyces lividans* strain overexpressing the pathway-specific activator gene actII-ORF4. *Journal of Industrial Microbiology & Biotechnology*, *28*(2), 103–111. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12074050
Busse, H.-J., Tindall, B. J., Ludwig, W., Rosselló-Móra, R., & Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology*, *60*(1), 249–266. https://doi.org/10.1099/ijs.0.016949-0
Calvo, A. M., Wilson, R. A., Bok, J. W., & Keller, N. P. (2002). Relationship between secondary metabolism and fungal development. *Microbiology and Molecular Biology Reviews*, *66*(3), 447. https://doi.org/10.1128/MMBR.66.3.447-459.2002
Castaldo, G., Zucko, J., Heidelberger, S., Vujaklija, D., Hranueli, D., Cullum, J., … Long, P. F. (2008). Proposed arrangement of proteins forming a bacterial type II polyketide synthase. *Chemistry & Biology*, *15*(11), 1156–1165. https://doi.org/10.1016/j.chembiol.2008.09.010
Chater, K. F. (2006). *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *361*(1469), 761–768. https://doi.org/10.1098/rstb.2005.1758
Chater, K. F., & Chandra, G. (2006). The evolution of development in *Streptomyces* analysed by genome comparisons. *FEMS Microbiology Reviews*, *30*(5), 651–672. https://doi.org/10.1111/j.1574-6976.2006.00033.x
Chen, G., Wang, G., Li, X., Waters, B., & Davies, J. (2000). Enhanced production of microbial metabolites in the presence of dimethyl sulfoxide. *The Journal of Antibiotics*, *53*(10), 1145–1153. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11132960
Chen, X., Ji, R., Jiang, X., Yang, R., Liu, F., & Xin, Y. (2014). Iterative type I polyketide synthases involved in enediyne natural product biosynthesis. *IUBMB Life*, *66*(9), 587–595. https://doi.org/10.1002/iub.1316
Connolly, D. J., Cusack, D., O'Sullivan, T. P., & Guiry, P. J. (2005). Synthesis of quinazolinones and quinazolines. *Tetrahedron*, *61*(43), 10153–10202. https://doi.org/10.1016/J.TET.2005.07.010
Cox, R. J., & Simpson, T. J. (2009). Fungal type I polyketide synthases. *Methods in Enzymology*, *459*, 49–78. https://doi.org/10.1016/S0076-6879(09)04603-5
Cruz-López, O., Conejo-García, A., Núñez, M. C., Kimatrai, M., García-Rubiño, M.

E., Morales, F., … Campos, J. M. (2011). Novel substituted quinazolines for potent EGFR tyrosine kinase inhibitors. *Current Medicinal Chemistry*, *18*(7), 943–963. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21254978

Dai, H.-Q., Wang, J., Xin, Y.-H., Pei, G., Tang, S.-K., Ren, B., … Zhang, L.-X. (2010). Verrucosispora sediminis sp. nov., a cyclodipeptide-producing actinomycete from deep-sea sediment. *International Journal of Systematic and Evolutionary Microbiology*, *60*(8), 1807–1812. https://doi.org/10.1099/ijs.0.017053-0

Dairi, T., Motohira, Y., Kuzuyama, T., Takahashi, S., Itoh, N., & Seto, H. (2000). Cloning of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from terpenoid antibiotic-producing Streptomyces strains. *Molecular & General Genetics : MGG*, *262*(6), 957–964. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10660057

Daley, D. K., Brown, K. J., & Badal, S. (2017). Fungal metabolites. *Pharmacognosy: Fundamentals, Applications and Strategy*, 413–421. Elsevier Inc. https://doi.org/10.1016/B978-0-12-802104-0.00020-2

Dalmas, F. R., Astarita, L., Defilippis, L., Magel, E., Fiedler, H.-P., Bauer, R., & Hampp, R. (2013). Growth inhibition of an *Araucaria angustifolia* (Coniferopsida) fungal seed pathogen, *Neofusicoccum parvum*, by soil streptomycetes. *BMC Microbiology*, *13*, 168. https://doi.org/10.1186/1471-2180-13-168

Dashti, Y., Grkovic, T., Abdelmohsen, U. R., Hentschel, U., & Quinn, R. J. (2017). Actinomycete metabolome induction/suppression with *N*-acetylglucosamine. *Journal of Natural Products*, *80*(4), 828–836. https://doi.org/10.1021/acs.jnatprod.6b00673

Deng, Z. X., Kieser, T., & Hopwood, D. A. (1987). Activity of a *Streptomyces* transcriptional terminator in *Escherichia coli*. *Nucleic Acids Research*, *15*(6), 2665–2675. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3031607

Denis, F., & Brzezinski, R. (1992). A versatile shuttle cosmid vector for use in *Escherichia coli* and Actinomycetes. *Gene*, *111*(1), 115–118. https://doi.org/10.1016/0378-1119(92)90611-r

Donadio, S., Monciardini, P., & Sosio, M. (2007a). Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Natural Product Reports*, *24*(5), 1073. https://doi.org/10.1039/b514050c

Donadio, S., Monciardini, P., & Sosio, M. (2007b). Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Natural Product Reports*, *24*(5), 1073. https://doi.org/10.1039/b514050c

Dreier, J., & Khosla, C. (2000). Mechanistic analysis of a type II polyketide synthase. role of conserved residues in the β-ketoacyl synthase−chain length factor heterodimer. https://doi.org/10.1021/BI992121L

Edwards, C. (1993). Isolation properties and potential applications of thermophilic Actinomycetes. *Applied Biochemistry and Biotechnology*, *42*(2–3), 161–179. https://doi.org/10.1007/BF02788050

Erlanger, B. F., & Goode, L. (1960). Antibacterial activity of acyclic decapeptide analogs of Gramicidin S. *Science*, *131*(3401), 669–670. https://doi.org/10.1126/science.131.3401.669

Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., & Jhoti, H. (2016). Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery*, *15*(9), 605–619. https://doi.org/10.1038/nrd.2016.109

Eustaquio, A. S. (2005). NovG, a DNA-binding protein acting as a positive regulator of novobiocin biosynthesis. *Microbiology*, *151*(6), 1949–1961.

https://doi.org/10.1099/mic.0.27669-0
Feng, Z., Wang, L., Rajski, S. R., Xu, Z., Coeffet-LeGal, M. F., & Shen, B. (2009). Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts. *Bioorganic & Medicinal Chemistry*, *17*(6), 2147–2153. https://doi.org/10.1016/j.bmc.2008.10.074
Finking, R., & Marahiel, M. A. (2004). Biosynthesis of nonribosomal peptides. *Annual Review of Microbiology*, *58*(1), 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615
Fitzpatrick, P. F. (2003). Mechanism of aromatic amino acid hydroxylation. *Biochemistry*, *42*(48), 14083–14091. https://doi.org/10.1021/bi035656u
Flinspach, K., Kapitzke, C., Tocchetti, A., Sosio, M., & Apel, A. K. (2014). Heterologous expression of the thiopeptide antibiotic GE2270 from *Planobispora rosea* ATCC 53733 in *Streptomyces coelicolor* requires deletion of ribosomal genes from the expression construct. *PLoS ONE*, *9*(3), e90499. https://doi.org/10.1371/journal.pone.0090499
Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M., & Smith, M. C. M. (2014). New applications for phage integrases. *Journal of Molecular Biology*, *426*(15), 2703–2716. https://doi.org/10.1016/j.jmb.2014.05.014
Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., & Ramos, J. L. (1997). Arac/XylS family of transcriptional regulators. *Microbiology and Molecular Biology Reviews : MMBR*, *61*(4), 393–410. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9409145
Genilloud, O. (2017). Actinomycetes: still a source of novel antibiotics. https://doi.org/10.1039/c7np00026j
Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318
Gokhale, R. S., Tsuji, S. Y., Cane, D. E., & Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science*, *284*(5413), 482–485. https://doi.org/10.1126/science.284.5413.482
Gomez-Escribano, J. P., Castro, J. F., Razmilic, V., Chandra, G., Andrews, B., Asenjo, J. A., & Bibb, M. J. (2015). The *Streptomyces leeuwenhoekii* genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics*, *16*(1), 485. https://doi.org/10.1186/s12864-015-1652-8
Goodfellow, M., & Williams, S. T. (1983). Ecology of Actinomycetes. *Annual Review of Microbiology*, *37*(1), 189–216. https://doi.org/10.1146/annurev.mi.37.100183.001201
Gregory, M. A., Till, R., & Smith, M. C. M. (2003). Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *Journal of Bacteriology*, *185*(17), 5320–5323. https://doi.org/10.1128/jb.185.17.5320-5323.2003
Gressler, M., Zaehle, C., Scherlach, K., Hertweck, C., & Brock, M. (2011). Multifactorial induction of an orphan PKS-NRPS gene cluster in *Aspergillus terreus*. *Chemistry & Biology*, *18*(2), 198–209. https://doi.org/10.1016/j.chembiol.2010.12.011
Griffiths, G. L., Sigel, S. P., Payne, S. M., & Neilands, J. B. (1984). Vibriobactin, a siderophore from *Vibrio cholerae*. *The Journal of Biological Chemistry*, *259*(1), 383–385. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6706943
Gross, H., Stockwell, V. O., Henkels, M. D., Nowak-Thompson, B., Loper, J. E., &

Gerwick, W. H. (2007). The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chemistry & Biology*, *14*(1), 53–63. https://doi.org/10.1016/j.chembiol.2006.11.007

Guo, C.-J., & Wang, C. C. C. (2014). Recent advances in genome mining of secondary metabolites in *Aspergillus terreus*. *Frontiers in Microbiology*, *5*, 717. https://doi.org/10.3389/fmicb.2014.00717

Gupta, A., Bedre, R., Thapa, S. S., Sabrin, A., Wang, G., Dassanayake, M., & Grove, A. (2017). Global awakening of cryptic biosynthetic gene clusters in *Burkholderia thailandensis*. *ACS Chemical Biology*, *12*(12), 3012–3021. https://doi.org/10.1021/acschembio.7b00681

Hackl, S., & Bechthold, A. (2015). The Gene *bldA* , a regulator of morphological differentiation and antibiotic production in *Streptomyces*. *Archiv Der Pharmazie*, *348*(7), 455–462. https://doi.org/10.1002/ardp.201500073

Hadjithomas, M., Chen, I.-M. A., Chu, K., Huang, J., Ratner, A., Palaniappan, K., … Ivanova, N. N. (2017). IMG-ABC: new features for bacterial secondary metabolism analysis and targeted biosynthetic gene cluster discovery in thousands of microbial genomes. *Nucleic Acids Research*, *45*(D1), D560–D565. https://doi.org/10.1093/nar/gkw1103

Hann, M. M., Leach, A. R., & Harper, G. (2001). Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences*, *41*(3), 856–864. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11410068

Hanson, J. R. (2003). *Natural Products: The Secondary Metabolites* (Vol. 17). Cambridge: Royal Society of Chemistry. https://doi.org/10.1039/9781847551535

Haydon, D. J., & Guest, J. R. (1991). A new family of bacterial regulatory proteins. *FEMS Microbiology Letters*, *63*(2–3), 291–295. https://doi.org/10.1016/0378-1097(91)90101-f

He, J., & Hertweck, C. (2003). Iteration as programmed event during polyketide assembly; molecular analysis of the aureothin biosynthesis gene cluster. *Chemistry & Biology*, *10*(12), 1225–1232. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14700630

Helfrich, E. J. N., & Piel, J. (2016). Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports*, *33*(2), 231–316. https://doi.org/10.1039/C5NP00125K

Henkel, T., Brunne, R. M., Müller, H., & Reichel, F. (1999). Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angewandte Chemie International Edition*, *38*(5), 643–647. https://doi.org/10.1002/(SICI)1521-3773(19990301)38:5<643::AID-ANIE643>3.0.CO;2-G

Hopwood, D. A. (2009). Complex enzymes in microbial natural product biosynthesis. Part B, Polyketides, aminocoumarins, and carbohydrates. *Elsevier/Academic Press*.

Horsman, G. P., Chen, Y., Thorson, J. S., & Shen, B. (2010). Polyketide synthase chemistry does not direct biosynthetic divergence between 9- and 10-membered enediynes. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11331–11335. https://doi.org/10.1073/pnas.1003442107

Hoshino, S., Okada, M., Wakimoto, T., Zhang, H., Hayashi, F., Onaka, H., & Abe, I. (2015). Niizalactams A–C, multicyclic macrolactams isolated from combined culture of *Streptomyces* with mycolic acid-containing bacterium. *Journal of*

Natural Products, *78*(12), 3011–3017. https://doi.org/10.1021/acs.jnatprod.5b00804
Huang, H., Hou, L., Li, H., Qiu, Y., Ju, J., & Li, W. (2016). Activation of a plasmid-situated type III PKS gene cluster by deletion of a *wbl* gene in deepsea-derived *Streptomyces somaliensis* SCSIO ZH66. *Microbial Cell Factories*, *15*(1), 116. https://doi.org/10.1186/s12934-016-0515-6
Jakubiec-Krzesniak, K., Rajnisz-Mateusiak, A., Guspiel, A., Ziemska, J., & Solecka, J. (2004). Secondary metabolites of Actinomycetes and their antibacterial, antifungal and antiviral properties. *Polish Journal of Microbiology* (Vol. 67). Polskie Towarzystwo Mikrobiologów. Retrieved from https://www.exeley.com/polish_journal_of_microbiology/doi/10.21307/pjm-2018-048
Janda, J. M., & Abbott, S. L. (2007, September). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of Clinical Microbiology*. https://doi.org/10.1128/JCM.01228-07
Jensen, K., Niederkrüger, H., Zimmermann, K., Vagstad, A. L., Moldenhauer, J., Brendel, N., … Piel, J. (2012). Polyketide proofreading by an acyltransferase-like enzyme. *Chemistry & Biology*, *19*(3), 329–339. https://doi.org/10.1016/j.chembiol.2012.01.005
Jiang, W., Zhao, X., Gabrieli, T., Lou, C., Ebenstein, Y., & Zhu, T. F. (2015). Cas9-assisted targeting of chromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nature Communications*, *6*(1), 8101. https://doi.org/10.1038/ncomms9101
Kabera J., Semana E., Mussa A., & He X. (2014). Plant secondary metabolites: biosynthesis, classification, function and pharmacological properties. *Journal of Pharmacy and Pharmacology*.
Kalan, L., Gessner, A., Thaker, M. N., Waglechner, N., Zhu, X., Szawiola, A., … Zechel, D. L. (2013). A cryptic polyene biosynthetic gene cluster in *Streptomyces calvus* is expressed upon complementation with a functional *bldA* gene. *Chemistry & Biology*, *20*(10), 1214–1224. https://doi.org/10.1016/J.CHEMBIOL.2013.09.006
Kallifidas, D., Jiang, G., Ding, Y., & Luesch, H. (2018). Rational engineering of *Streptomyces albus* J1074 for the overexpression of secondary metabolite gene clusters. *Microbial Cell Factories*, *17*(1), 25. https://doi.org/10.1186/s12934-018-0874-2
Katz, L., & Baltz, R. H. (2016). Natural product discovery: past, present, and future. *Journal of Industrial Microbiology & Biotechnology*, *43*(2–3), 155–176. https://doi.org/10.1007/s10295-015-1723-5
Keating, T. A., Marshall, C. G., & Walsh, C. T. (2000). Reconstitution and characterization of the *Vibrio cholerae* vibriobactin synthetase from VibB, VibE, VibF, and VibH. *Biochemistry*, *39*(50), 15522–15530. https://doi.org/10.1021/bi0016523
Keller, N. P., Turner, G., & Bennett, J. W. (2005). Fungal secondary metabolism — from biochemistry to genomics. *Nature Reviews Microbiology*, *3*(12), 937–947. https://doi.org/10.1038/nrmicro1286
Keravala, A., & Calos, M. P. (2008). Site-specific chromosomal integration mediated by ϕC31 integrase. *Methods in molecular biology (Clifton, N.J.)* (Vol. 435, pp. 165–173). https://doi.org/10.1007/978-1-59745-232-8_12
Kessler, N., Schuhmann, H., Morneweg, S., Linne, U., & Marahiel, M. A. (2004). The linear pentadecapeptide gramicidin is assembled by four multimodular

nonribosomal peptide synthetases that comprise 16 modules with 56 catalytic domains. *Journal of Biological Chemistry*, *279*(9), 7413–7419. https://doi.org/10.1074/jbc.M309658200
Khaw, L. E., Böhm, G. A., Metcalfe, S., Staunton, J., & Leadlay, P. F. (1998). Mutational biosynthesis of novel rapamycins by a strain of *Streptomyces hygroscopicus* NRRL 5491 disrupted in rapL, encoding a putative lysine cyclodeaminase. *Journal of Bacteriology*, *180*(4), 809–814. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9473033
Khosla, C., Gokhale, R. S., Jacobsen, J. R., & Cane, D. E. (1999). Tolerance and specificity of polyketide synthases. *Annual Review of Biochemistry*, *68*(1), 219–253. https://doi.org/10.1146/annurev.biochem.68.1.219
Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., Hopwood, D. A., & John Innes Foundation. (2000). Practical Streptomyces genetics. *Norwich: The John Innes Foundation*. Retrieved from https://www.nhbs.com/practical-streptomyces-genetics-book
Kim, H. U., Blin, K., Lee, S. Y., & Weber, T. (2017). Recent development of computational resources for new antibiotics discovery. *Current Opinion in Microbiology*, *39*, 113–120. https://doi.org/10.1016/j.mib.2017.10.027
Kim, J. H., Feng, Z., Bauer, J. D., Kallifidas, D., Calle, P. Y., & Brady, S. F. (2010). Cloning large natural product gene clusters from the environment: piecing environmental DNA gene clusters back together with TAR. *Biopolymers*, *93*(9), 833–844. https://doi.org/10.1002/bip.21450
Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 2), 346–351. https://doi.org/10.1099/ijs.0.059774-0
Kim, S.-Y., Zhao, P., Igarashi, M., Sawa, R., Tomita, T., Nishiyama, M., & Kuzuyama, T. (2009). Cloning and heterologous expression of the cyclooctatin biosynthetic gene cluster afford a diterpene cyclase and two P450 hydroxylases. *Chemistry & Biology*, *16*(7), 736–743. https://doi.org/10.1016/j.chembiol.2009.06.007
Klementz, D., Döring, K., Lucas, X., Telukunta, K. K., Erxleben, A., Deubel, D., … Günther, S. (2016). StreptomeDB 2.0--an extended resource of natural products produced by *Streptomycetes*. *Nucleic Acids Research*, *44*(D1), D509-14. https://doi.org/10.1093/nar/gkv1319
Klopries, S., Sundermann, U., & Schulz, F. (2013). Quantification of *N* -acetylcysteamine activated methylmalonate incorporation into polyketide biosynthesis. *Beilstein Journal of Organic Chemistry*, *9*(1), 664–674. https://doi.org/10.3762/bjoc.9.75
Koju, D., Maharjan, S., Dhakal, D., Yoo, J. C., & Sohng, J. K. (2012). Effect of different biosynthetic precursors on the production of nargenicin A1 from metabolically engineered *Nocardia* sp. CS682. *Journal of Microbiology and Biotechnology*, *22*(8), 1127–1132. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22713990
Koshla, O., Lopatniuk, M., Rokytskyy, I., Yushchuk, O., Dacyuk, Y., Fedorenko, V., … Ostash, B. (2017). Properties of *Streptomyces albus* J1074 mutant deficient in tRNALeuUAA gene *bldA*. *Archives of Microbiology*, *199*(8), 1175–1183. https://doi.org/10.1007/s00203-017-1389-7
Kouprina, N., & Larionov, V. (2016). Transformation-associated recombination (TAR)

cloning for genomics studies and synthetic biology. *Chromosoma*, *125*(4), 621–632. https://doi.org/10.1007/s00412-016-0588-3
Kuhstoss, S., Richardson, M. A., & Rao, R. N. (1991). Plasmid cloning vectors that integrate site-specifically in *Streptomyces* spp. *Gene*, *97*(1), 143–146. https://doi.org/10.1016/0378-1119(91)90022-4
Lancini, G. C., Thiemann, J. E., Sartori, G., & Sensi, P. (1967). Biogenesis of rifamycins. The conversion of rifamycin B into rifamycin Y. *Experientia*, *23*(11), 899–900. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6056996
Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., & Aigle, B. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proceedings of the National Academy of Sciences*, *108*(15), 6258–6263. https://doi.org/10.1073/pnas.1019077108
Lautru, S., Deeth, R. J., Bailey, L. M., & Challis, G. L. (2005). Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nature Chemical Biology*, *1*(5), 265–269. https://doi.org/10.1038/nchembio731
Lawlor, E. J., Baylis, H. A., & Chater, K. F. (1987). Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). *Genes & Development*, *1*(10), 1305–1310. https://doi.org/10.1101/GAD.1.10.1305
Lech, K., & Jarosz, M. (2011). Novel methodology for the extraction and identification of natural dyestuffs in historical textiles by HPLC-UV-Vis-ESI MS. Case study: chasubles from the Wawel Cathedral collection. *Analytical and Bioanalytical Chemistry*, *399*(9), 3241–3251. https://doi.org/10.1007/s00216-010-4591-x
Lechevalier, H., & Lechevalier, M. P. (1965). Classification of aerobic Actinomycetes based on their morphology and their chemical composition. *Annales de l'Institut Pasteur*, *108*(5), 662–673. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5838643
Li, J., Li, L., Tian, Y., Niu, G., & Tan, H. (2011). Hybrid antibiotics with the nikkomycin nucleoside and polyoxin peptidyl moieties. *Metabolic Engineering*, *13*(3), 336–344. https://doi.org/10.1016/j.ymben.2011.01.002
Lipmann, F., Gevers, W., Kleinkauf, H., & Roskoski, R. (1971). Polypeptide synthesis on protein templates: the enzymatic synthesis of gramicidin S and tyrocidine. *Advances in Enzymology and Related Areas of Molecular Biology*, *35*, 1–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4131410
Locatelli, F. M., Goo, K.-S., & Ulanova, D. (2016). Effects of trace metal ions on secondary metabolism and the morphological development of *Streptomycetes*. *Metallomics*, *8*(5), 469–480. https://doi.org/10.1039/C5MT00324E
Lombó, F., Velasco, A., Castro, A., de la Calle, F., Braña, A. F., Sánchez-Puelles, J. M., … Salas, J. A. (2006). Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine Actinomycete and its expression in two *Streptomyces* species. *ChemBioChem*, *7*(2), 366–376. https://doi.org/10.1002/cbic.200500325
Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., & Voigt, C. A. (2012). Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nature Biotechnology*, *30*(11), 1137–1142. https://doi.org/10.1038/nbt.2401
Loub, W. D., Farnsworth, N. R., Soejarto, D. D., & Quinn, M. L. (1985). NAPRALERT: computer handling of natural product research data. *Journal of Chemical Information and Computer Sciences*, *25*(2), 99–103. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4008538

Lünse, C. E., & Mayer, G. (2017). Reporter gene-based screening for TPP riboswitch activators (pp. 227–235). *Humana Press, New York, NY*. https://doi.org/10.1007/978-1-4939-6634-9_13

Lussier, F.-X., Colatriano, D., Wiltshire, Z., Page, J. E., & Martin, V. J. J. (2012). ENGINEERING MICROBES FOR PLANT POLYKETIDE BIOSYNTHESIS. *Computational and Structural Biotechnology Journal*, *3*(4), e201210020. https://doi.org/10.5936/csbj.201210020

Maddocks, S. E., & Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology*, *154*(12), 3609–3623. https://doi.org/10.1099/mic.0.2008/022772-0

Malla, S., Niraula, N. P., Singh, B., Liou, K., & Sohng, J. K. (2010). Limitations in doxorubicin production from *Streptomyces peucetius*. *Microbiological Research*, *165*(5), 427–435. https://doi.org/10.1016/j.micres.2009.11.006

Mayfield, C. I., Williams, S. T., Ruddick, S. M., & Hatfield, H. L. (1972). Studies on the ecology of Actinomycetes in soil IV. Observations on the form and growth of streptomycetes in soil. *Soil Biology and Biochemistry*, *4*(1), 79–91. https://doi.org/10.1016/0038-0717(72)90045-4

Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., … Glöckner, F. O. (2015). Minimum information about a biosynthetic gene cluster. *Nature Chemical Biology*, *11*(9), 625–631. https://doi.org/10.1038/nchembio.1890

Méndez-Lucio, O., & Medina-Franco, J. L. (2017). The many roles of molecular complexity in drug discovery. *Drug Discovery Today*. Elsevier Ltd. https://doi.org/10.1016/j.drudis.2016.08.009

Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., … Santi, D. V. (2005). Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nature Biotechnology*, *23*(9), 1171–1176. https://doi.org/10.1038/nbt1128

Monciardini, P., Iorio, M., Maffioli, S., Sosio, M., & Donadio, S. (2014). Discovering new bioactive molecules from microbial sources. *Microbial Biotechnology*, *7*(3), 209–220. https://doi.org/10.1111/1751-7915.12123

Mor, A. (2000). Peptide-based antibiotics: A potential answer to raging antimicrobial resistance. *Drug Development Research*, *50*(3–4), 440–447. https://doi.org/10.1002/1098-2299(200007/08)50:3/4<440::AID-DDR27>3.0.CO;2-4

Motallebi-Veshareh, M., Balzer, D., Lanka, E., Jagura-Burdzy, G., & Thomas, C. M. (1992). Conjugative transfer functions of broad-host-range plasmid RK2 are coregulated with vegetative replication. *Molecular Microbiology*, *6*(7), 907–920. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1376390

Müller, C. A., Oberauner-Wappis, L., Peyman, A., Amos, G. C. A., Wellington, E. M. H., & Berg, G. (2015). Mining for nonribosomal peptide synthetase and polyketide synthase genes revealed a high level of diversity in the Sphagnum Bog metagenome. *Applied and Environmental Microbiology*, *81*(15), 5064–5072. https://doi.org/10.1128/AEM.00631-15

Newman, D. J., & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055

Nguyen, Q.-T., Merlo, M. E., Medema, M. H., Jankevics, A., Breitling, R., & Takano, E. (2012). Metabolomics methods for the synthetic biology of secondary metabolism. *FEBS Letters*, *586*(15), 2177–2183.

https://doi.org/10.1016/j.febslet.2012.02.008
Nikaido, H. (2009). Multidrug resistance in bacteria. *Annual Review of Biochemistry*, *78*, 119–146. https://doi.org/10.1146/annurev.biochem.78.082907.145923
Noskov, V. N., Karas, B. J., Young, L., Chuang, R.-Y., Gibson, D. G., Lin, Y.-C., … Weyman, P. D. (2012). Assembly of large, high G+C bacterial DNA fragments in yeast. *ACS Synthetic Biology*, *1*(7), 267–273. https://doi.org/10.1021/sb3000194
Oh, S.-Y., Shin, J.-H., & Roe, J.-H. (2007). Dual role of OhrR as a repressor and an activator in response to organic hydroperoxides in *Streptomyces coelicolor*. *Journal of Bacteriology*, *189*(17), 6284–6292. https://doi.org/10.1128/JB.00632-07
Onaka, H., Mori, Y., Igarashi, Y., & Furumai, T. (2011). Mycolic acid-containing bacteria induce natural-product biosynthesis in Streptomyces species. *Applied and Environmental Microbiology*, *77*(2), 400–406. https://doi.org/10.1128/AEM.01337-10
Paranthaman, S., & Dharmalingam, K. (2003). Intergeneric conjugation in *Streptomyces peucetius* and *Streptomyces* sp. strain C5: chromosomal integration and expression of recombinant plasmids carrying the *chiC* gene. *Applied and Environmental Microbiology*, *69*(1), 84–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12513980
Pascolutti, M., & Quinn, R. J. (2014). Natural products as lead structures: chemical transformations to create lead-like libraries. *Drug Discovery Today*, *19*(3), 215–221. https://doi.org/10.1016/J.DRUDIS.2013.10.013
Paulus, C., Rebets, Y., Tokovenko, B., Nadmid, S., Terekhova, L. P., Myronovskyi, M., … Luzhetskyy, A. (2017). New natural products identified by combined genomics-metabolomics profiling of marine *Streptomyces* sp. MP131-18. *Scientific Reports*, *7*(1), 42382. https://doi.org/10.1038/srep42382
Pelludat, C., Rakin, A., Jacobi, C. A., Schubert, S., & Heesemann, J. (1998). The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. *Journal of Bacteriology*, *180*(3), 538–546. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9457855
Pérez, J., Muñoz-Dorado, J., Braña, A. F., Shimkets, L. J., Sevillano, L., & Santamaría, R. I. (2011). *Myxococcus xanthus* induces actinorhodin overproduction and aerial mycelium formation by *Streptomyces coelicolor*. *Microbial Biotechnology*, *4*(2), 175–183. https://doi.org/10.1111/j.1751-7915.2010.00208.x
Pettit, R. K. (2011). Small-molecule elicitation of microbial secondary metabolites. *Microbial Biotechnology*, *4*(4), 471–478. https://doi.org/10.1111/j.1751-7915.2010.00196.x
Pfeifer, B. A., & Khosla, C. (2001). Biosynthesis of polyketides in heterologous hosts. *Microbiology and Molecular Biology Reviews : MMBR*, *65*(1), 106–118. https://doi.org/10.1128/MMBR.65.1.106-118.2001
Pfennig, F., Schauwecker, F., & Keller, U. (1999). Molecular characterization of the genes of actinomycin synthetase I and of a 4-methyl-3-hydroxyanthranilic acid carrier protein involved in the assembly of the acylpeptide chain of actinomycin in *Streptomyces*. *Journal of Biological Chemistry*, *274*(18), 12508–12516. https://doi.org/10.1074/jbc.274.18.12508
Pigac, J., & Schrempf, H. (1995). A simple and rapid method of transformation of *Streptomyces rimosus* R6 and other *Streptomycetes* by electroporation. *Applied and Environmental Microbiology*, *61*(1), 352–356. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16534915

Ramos, J. L., Martinez-Bueno, M., Molina-Henares, A. J., Teran, W., Watanabe, K., Zhang, X., … Tobes, R. (2005). The TetR family of transcriptional repressors. *Microbiology and Molecular Biology Reviews*, *69*(2), 326–356. https://doi.org/10.1128/MMBR.69.2.326-356.2005

Rebets, Y. V., Ostash, B. O., Fukuhara, M., Nakamura, T., & Fedorenko, V. O. (2006). Expression of the regulatory protein LndI for landomycin E production in *Streptomyces globisporus* 1912 is controlled by the availability of tRNA for the rare UUA codon. *FEMS Microbiology Letters*, *256*(1), 30–37. https://doi.org/10.1111/j.1574-6968.2005.00087.x

Reen, F. J., Romano, S., Dobson, A. D. W., & O'Gara, F. (2015). The sound of silence: activating silent biosynthetic gene clusters in marine microorganisms. *Marine Drugs*, *13*(8), 4754–4783. https://doi.org/10.3390/md13084754

Reimer, D., & Bode, H. B. (2014). A natural prodrug activation mechanism in the biosynthesis of nonribosomal peptides. *Natural Product Reports*, *31*(2), 154–159. https://doi.org/10.1039/c3np70081j

Richard, T., Temsamani, H., Cantos-Villar, E., & Monti, J. P. (2013). Application of LC-MS and LC-NMR techniques for secondary metabolite identification. In *Advances in Botanical Research* (Vol. 67, pp. 67–98). Academic Press Inc. https://doi.org/10.1016/B978-0-12-397922-3.00002-2

Richter, T. K. S., Hughes, C. C., & Moore, B. S. (2015). Sioxanthin, a novel glycosylated carotenoid, reveals an unusual subclustered biosynthetic pathway. *Environmental Microbiology*, *17*(6), 2158–2171. https://doi.org/10.1111/1462-2920.12669

Rigali, S., Titgemeyer, F., Barends, S., Mulder, S., Thomae, A. W., Hopwood, D. A., & van Wezel, G. P. (2008). Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Reports*, *9*(7), 670–675. https://doi.org/10.1038/embor.2008.83

Romano, S., Jackson, S., Patry, S., Dobson, A., Romano, S., Jackson, S. A., … Dobson, A. D. W. (2018). Extending the "one strain many compounds" (OSMAC) principle to marine microorganisms. *Marine Drugs*, *16*(7), 244. https://doi.org/10.3390/md16070244

Rutledge, P. J., & Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology*, *13*(8), 509–523. https://doi.org/10.1038/nrmicro3496

Sánchez, S., Chávez, A., Forero, A., García-Huante, Y., Romero, A., Sánchez, M., … Ruiz, B. (2010). Carbon source regulation of antibiotic production. *The Journal of Antibiotics*, *63*(8), 442–459. https://doi.org/10.1038/ja.2010.78

Schuffenhauer, A., Brown, N., Selzer, P., Ertl, P., & Jacoby, E. (2006). Relationships between molecular complexity, biological activity, and structural diversity. *Journal of Chemical Information and Modeling* (Vol. 46, pp. 525–535). https://doi.org/10.1021/ci0503558

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. https://doi.org/10.1073/pnas.1400019111

Sharma, S. V., Tong, X., Pubill-Ulldemolins, C., Cartmell, C., Bogosyan, E. J. A., Rackham, E. J., … Goss, R. J. M. (2017). Living GenoChemetics by hyphenating synthetic biology and synthetic chemistry in vivo. *Nature Communications*, *8*(1), 229. https://doi.org/10.1038/s41467-017-00194-3

Sieber, S. A., & Marahiel, M. A. (2005). Molecular mechanisms underlying

nonribosomal peptide synthesis: approaches to new antibiotics. *Chemical Reviews*, *105*(2), 715–738. https://doi.org/10.1021/cr0301191
Siezen, R. J., & Khayatt, B. I. (2008). Natural products genomics. *Microbial Biotechnology*, *1*(4), 275–282. https://doi.org/10.1111/j.1751-7915.2008.00044.x
Slee, A. M., Wuonola, M. A., McRipley, R. J., Zajac, I., Zawada, M. J., Bartholomew, P. T., … Forbes, M. (1987). Oxazolidinones, a new class of synthetic antibacterial agents: in vitro and in vivo activities of DuP 105 and DuP 721. *Antimicrobial Agents and Chemotherapy*, *31*(11), 1791–1797. https://doi.org/10.1128/aac.31.11.1791
Smanski, M. J., Yu, Z., Casper, J., Lin, S., Peterson, R. M., Chen, Y., … Shen, B. (2011). Dedicated ent-kaurene and ent-atiserene synthases for platensimycin and platencin biosynthesis. *Proceedings of the National Academy of Sciences*, *108*(33), 13498–13503. https://doi.org/10.1073/pnas.1106919108
Song, L., Barona-Gomez, F., Corre, C., Xiang, L., Udwary, D. W., Austin, M. B., … Challis, G. L. (2006). Type III Polyketide synthase β-Ketoacyl-ACP starter unit and ethylmalonyl-CoA extender unit selectivity discovered by *Streptomyces coelicolor* genome mining. *Journal of the American Chemical Society*, *128*(46), 14754–14755. https://doi.org/10.1021/ja065247w
Supong, K., Suriyachadkun, C., Tanasupawat, S., Suwanborirux, K., Pittayakhajonwut, P., Kudo, T., & Thawai, C. (2013). *Micromonospora sediminicola* sp. nov., isolated from marine sediment. *International Journal of Systematic and Evolutionary Microbiology*, *63*(Pt 2), 570–575. https://doi.org/10.1099/ijs.0.041103-0
Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., & Julien, B. (2000). Cloning and heterologous expression of the epothilone gene cluster. *Science*, *287*(5453), 640–642. https://doi.org/10.1126/science.287.5453.640
Tercero, J. A., Espinosa, J. C., & Jiménez, A. (1998). Expression of the *Streptomyces alboniger* pur cluster in *Streptomyces lividans* is dependent on the bldA-encoded tRNALeu. *FEBS Letters*, *421*(3), 221–223. https://doi.org/10.1016/s0014-5793(97)01564-0
Tokuoka, M., Seshime, Y., Fujii, I., Kitamoto, K., Takahashi, T., & Koyama, Y. (2008). Identification of a novel polyketide synthase–nonribosomal peptide synthetase (PKS–NRPS) gene required for the biosynthesis of cyclopiazonic acid in *Aspergillus oryzae*. *Fungal Genetics and Biology*, *45*(12), 1608–1615. https://doi.org/10.1016/j.fgb.2008.09.006
Tong, Y., Charusanti, P., Zhang, L., Weber, T., & Lee, S. Y. (2015). CRISPR-Cas9 based engineering of Actinomycetal genomes. *ACS Synthetic Biology*, *4*(9), 1020–1029. https://doi.org/10.1021/acssynbio.5b00038
Trivella, D. B. B., & de Felicio, R. (2018). The tripod for bacterial natural product discovery: genome mining, silent pathway induction, and mass spectrometry-based molecular networking. *MSystems*, *3*(2). https://doi.org/10.1128/mSystems.00160-17
Tseng, M., Liao, H. C., Chiang, W. P., & Yuan, G. F. (2011). *Isoptericola chiayiensis* sp. nov., isolated from mangrove soil. *International Journal of Systematic and Evolutionary Microbiology*, (61), 1667–1670. https://doi.org/10.1099/ijs.0.022491-0
Van Mellaert, L., Mei, L., Lammertyn, E., Schacht, S., & Anne, J. (1998). Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. *Microbiology*, *144*(12), 3351–3358. https://doi.org/10.1099/00221287-144-12-3351

Vermeij, P., Wietek, C., Kahnert, A., Wuest, T., & Kertesz, M. A. (1999). Genetic organization of sulphur-controlled aryl desulphonation in Pseudomonas putida S-313. *Molecular Microbiology*, *32*(5), 913–926. https://doi.org/10.1046/j.1365-2958.1999.01398.x

Walsh, C. T., Chen, H., Keating, T. A., Hubbard, B. K., Losey, H. C., Luo, L., … Patel, H. M. (2001). Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Current Opinion in Chemical Biology*, *5*(5), 525–534. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11578925

Wang, H., Fewer, D. P., Holm, L., Rouhiainen, L., & Sivonen, K. (2014). Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proceedings of the National Academy of Sciences*, *111*(25), 9259–9264. https://doi.org/10.1073/pnas.1401734111

Wang, Huan, & van der Donk, W. A. (2012). Biosynthesis of the Class III Lantipeptide Catenulipeptin. *ACS Chemical Biology*, *7*(9), 1529–1535. https://doi.org/10.1021/cb3002446

Wang, J., Schully, K. L., & Pettis, G. S. (2009). Growth-regulated expression of a bacteriocin, produced by the sweet potato pathogen *Streptomyces ipomoeae*, that exhibits interstrain inhibition. *Applied and Environmental Microbiology*, *75*(5), 1236–1242. https://doi.org/10.1128/AEM.01598-08

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., … Medema, M. H. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, *43*(W1), W237–W243. https://doi.org/10.1093/nar/gkv437

Wierzbicka, R., Wu, H., Franek, M., Kamal-Eldin, A., & Landberg, R. (2015). Determination of alkylresorcinols and their metabolites in biological samples by gas chromatography–mass spectrometry. *Journal of Chromatography B*, *1000*, 120–129. https://doi.org/10.1016/j.jchromb.2015.07.009

Williams, S. T., Locci, R., Beswick, A., Kurtböke, D. I., Kuznetsov, V. D., Le Monnier, F. J., … West, M. (1993). Detection and identification of novel Actinomycetes. *Research in Microbiology*, *144*(8), 653–656. https://doi.org/10.1016/0923-2508(93)90069-E

Winn, M., Fyans, J. K., Zhuo, Y., & Micklefield, J. (2016, February 1). Recent advances in engineering nonribosomal peptide assembly lines. *Natural Product Reports*. Royal Society of Chemistry. https://doi.org/10.1039/c5np00099h

Winter, J. M., Moffitt, M. C., Zazopoulos, E., McAlpine, J. B., Dorrestein, P. C., & Moore, B. S. (2007). Molecular basis for chloronium-mediated meroterpene cyclization. *Journal of Biological Chemistry*, *282*(22), 16362–16368. https://doi.org/10.1074/jbc.M611046200

Wohlleben, W., & Pühler, A. (1987). The *Streptomyces ghanaensis* low copy plasmid pSG2 and its use for vector construction. *Archives of Microbiology*, *148*(4), 298–304. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3318753

Wong, F. T., Jin, X., Mathews, I. I., Cane, D. E., & Khosla, C. (2011). Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. *Biochemistry*, *50*(30), 6539–6548. https://doi.org/10.1021/bi200632j

Xi, L., Zhang, L., Ruan, J., & Huang, Y. (2011). *Nonomuraea maritima* sp. nov., isolated from coastal sediment. *International Journal of Systematic and Evolutionary Microbiology*, *61*(11), 2740–2744. https://doi.org/10.1099/ijs.0.028266-0

You, Z., Omura, S., Ikeda, H., & Cane, D. E. (2007). Pentalenolactone biosynthesis:

molecular cloning and assignment of biochemical function to PtlF, a short-chain dehydrogenase from *Streptomyces avermitilis*, and identification of a new biosynthetic intermediate. *Archives of Biochemistry and Biophysics*, *459*(2), 233–240. https://doi.org/10.1016/j.abb.2006.11.016

YPD media. (2010). *Cold Spring Harbor Protocols*, *2010*(9), pdb.rec12315-pdb.rec12315. https://doi.org/10.1101/pdb.rec12315

Yu, D., Xu, F., Zeng, J., & Zhan, J. (2012). Type III polyketide synthases in natural product biosynthesis. *IUBMB Life*, *64*(4), 285–295. https://doi.org/10.1002/iub.1005

Zaburannyi, N., Rabyk, M., Ostash, B., Fedorenko, V., & Luzhetskyy, A. (2014). Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics*, *15*(1), 97. https://doi.org/10.1186/1471-2164-15-97

Zotchev, S. B. (2014). Genomics-based insights into the evolution of secondary metabolite biosynthesis in Actinomycete bacteria. *Evolutionary Biology: Genome Evolution, Speciation, Coevolution and Origin of Life* (pp. 35–45). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-07623-2_2

APPENDICES

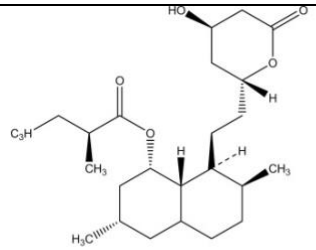
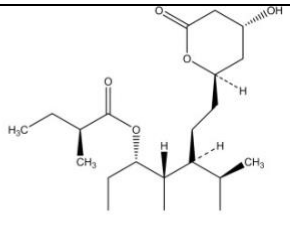
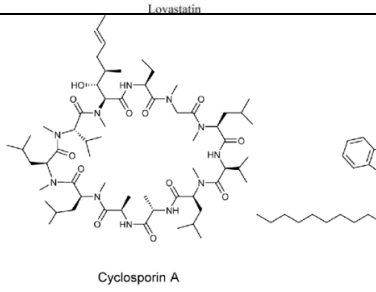
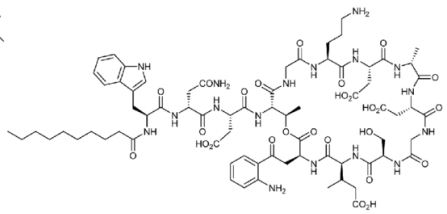
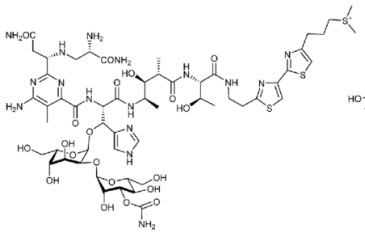
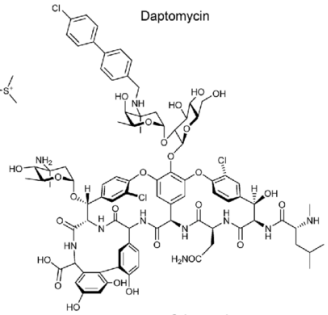
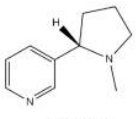
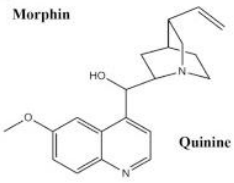
Primer	Sequence (5' → 3')
2243-F	AACGACGACCTAGGGGTTTTGCATGGACATCGAAGTGCTCGGC
2243-R	CGGGGCTCTTCACGTCCAGATCAGTACCGGCCCGGCG
2694-F	AACGACGACCTAGGGGTTTTGCATGGTGTTATCAGGTGACAGCG
2694T4-F	GGTTTTGCGGCCGCATGGTGTTATCAGGTGACAGCG
2694-R	CGGGGCTCTTCACGTCCAGATCAGTCCGGCCACGCC
8145-F	AACGACGACCTAGGGGTTTTGCATGTCCGACTCACGGCCC
8145-R	CGGGGCTCTTCACGTCCAGATCAGGCGCCGGAGGAGAG
Terminator-F	TCTGGACGTGAAGAGCCCCG
Terminator-R	TAGGTCTACGTAACTCATTAATTAACTCGCCCATTGAGAAGCCCCGTCAG
***bldA*-F**	CTGACGGGGCTTCTCAATGGGCGAGTTAATTAATGAGTTACGTAGACCTA
***bldA*-R**	ACGAATTCGATATCGCGCGCGCAGGACGAAAGCCCATACCCAGGTCGCCA
PSK152-F	TTGGGTAACGCCAGGGTTTT
PSK152-R	CACTCATTAGGCACCCCAGG
ΦC31-F	CCAGGTGCGAATAAGGGACA
ΦC31-R	TCAACTACCGCCATCTGGAC

Table S1. Sequence of primers for constructing the P46 mutants.

Primer	Sequence (5' → 3')
INTCONTROL-F	TACCTCAAGCAGATCGGCAA
INTCONTROL-R	CTCCAGGAGGTGGTTCTTGG
2214-F	CTATCGCAACATCTCCGGCA
2214-R	CGTCGCTGAGCAGATAGACC
2234-F	ATGGACCGGAGTTCCTCCC
2234-R	GTGGAGCAGATTGCCCAAAG
2237-F	GGTGTCGTCAGCCAGAAGAA
2237-R	TTCATGGGCTTGACCTGCTT
2242-F	CCGGACAGCAAAGGCGTAT
2242-R	CACGGTGTTGGAGATGTTGG
2243-F	CCCAAAGGCATATCAGGCGA
2243-R	GCATGCAGAGTGATGCAGAC
8145-F	ACGAAGACGATTACGGCTGG
8145-R	TCAGTGACATCCGTAGCGAG

8146-F	ATCCACAGGTAGCGGTTGTC
8146-R	ACCCAACTCGTCCTCAATGG
8151-F	GCTCATCGTGATGGAGGACG
8151-R	CGTCTGACCATGCCTTCGAT
8153-F	GCTCGTCACCAGCTATGAGT
8153-R	CACCGGTACCGAGTAGTCC

Table S2. Sequence of primers for RT-PCR.

Secondary Metabolite Class	**Representing Compounds**	**Source**
Polyketides and fatty acid-derived substances	 Lovastatin  Pravastatin	(Daley et al., 2017)
Ribosomal and non-ribosomal peptides	 Cyclosporin A  Daptomycin  Bleomycin A2  Oritavancin	(Winn et al., 2016)
Alkaloids	 Caffeine  Morphin  Nicotine  Strychnine  Quinine	(Richard et al., 2013)

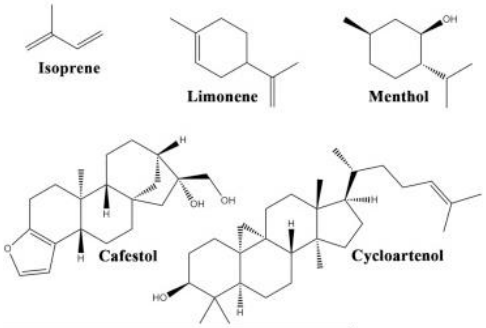
Terpenoids		(Richard et al., 2013)

Table S3. Secondary metabolite classes and their respective representing compounds.

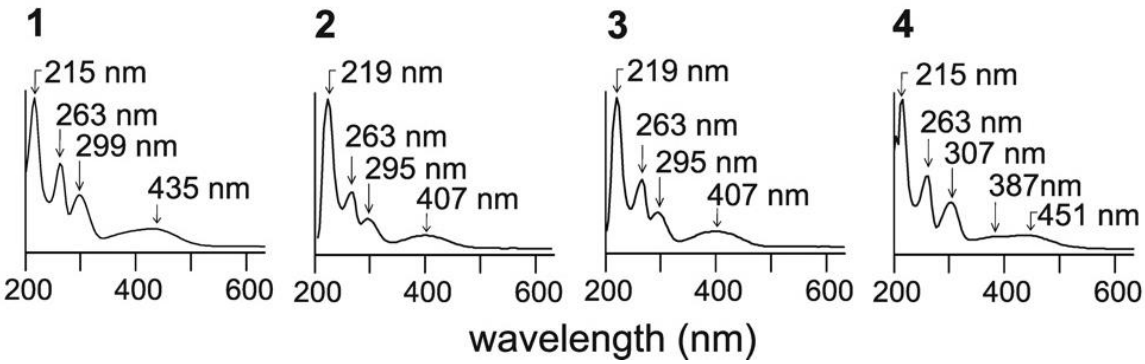


Figure S1. UV-Vis Spectra of furaquinocin A (Kumano, 2006).

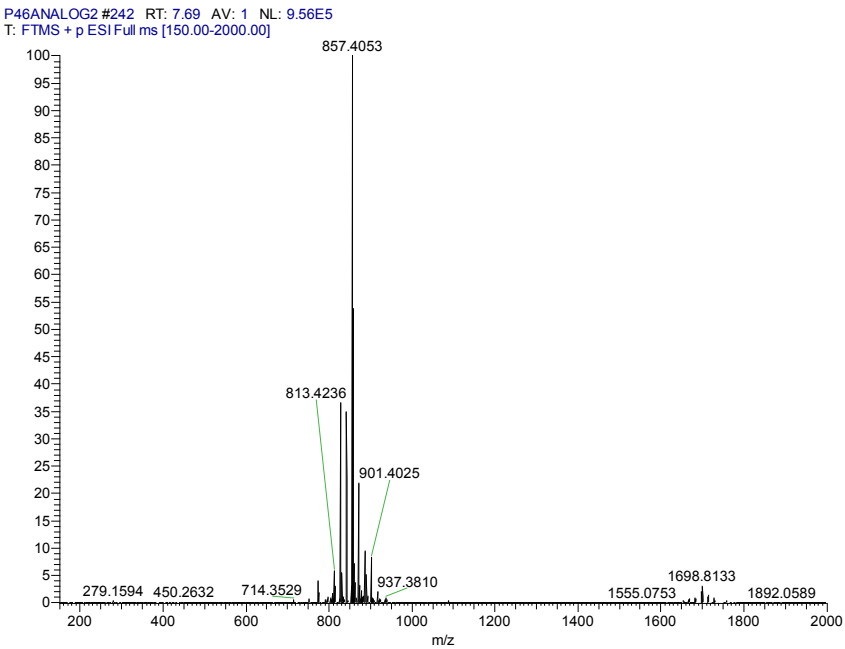


Figure S2. High resolution electrospray ionization mass spectrometry (HR-ESI-MS) positive ion mode profile for Tasikamide A (**T2**).