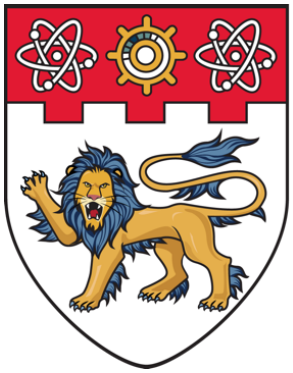


NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

CHARACTERIZATION AND OPTIMIZATION OF PROPRIETARY HETEROLOGOUS *STREPTOMYCES* CHASSIS

SEAN LEE QIU EN

SCHOOL OF BIOLOGICAL SCIENCES

2024

CHARACTERIZATION AND OPTIMIZATION OF PROPRIETARY HETEROLOGOUS *STREPTOMYCES* CHASSIS

SEAN LEE QIU EN

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2024

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.

19/01/2024

. . . . . .

Date

. . . . . .

SEAN LEE QIU EN

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.

20/01/2024

Date

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.

19/01/2024

. . . . . .

Date

. . . . . .

SEAN LEE QIU EN

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Liang Zhao Xun for his support, guidance, and great patience during the four years of my PhD program. Being able to pursue a special interest in actinomycete chassis was a great privilege and opportunity that I did not expect to be able to undertake in Singapore. Furthermore, the encouragements he gave helped me build resilience in overcoming adversities, especially during the disruptive COVID-19 outbreak. I especially appreciated the guidance that was given to help improve the quality of my scientific thinking, writing, and communication – essential for going beyond mere technician to growing into the role of a scientist. Undertaking this crucial and pivotal training in his lab has truly been a great and memorable experience.

I would also like to express my gratitude to Dr. Hartono Candra for his invaluable help and guidance for the better part of my project, and always being available for rendering insightful advice. Many an experiment had gone most smoothly by being supported with the depth of his experience, for which I am indeed grateful.

Additionally, I would like to express thanks to Dr. Ma Guang lei and Dr. Wang Xue Jiao for their assistance in chemistry related topics and analysis, particularly with the NMR based structure elucidation of compounds.

I am also grateful to the support and assistance from the other colleagues in the lab – Srashti Khandelwal, Liao Yang Hui, and Huang Hua Wei, for making the

overall experience of the lab far more enjoyable, especially during the team building events of lab maintenance duties.

Also, I would like to express a special thanks to Hoa Tran and Rachel Andrea Chea Yuen Fong for helping me settle into the lab in the early days of the program. Their support during the early days made the transition and onboarding into a new phase of my scientific journey much easier to handle.

Finally, I would like to thank the undergraduate students whom I had the great pleasure of mentoring, who in turn aided in my own growth and understanding - Teh Ker Na and Nur Dini Binte Mohamad Faisal Abnass, and especially Clara Lie Kai Kee, whose immense industriousness and positivity was a great blessing and encouragement.

Table of Contents

List of Figures

List of Supplementary Figures

List of Supplementary Tables

Summary

Microbial natural products are secondary metabolites synthesized by enzymes encoded by biosynthetic gene clusters (BGCs). Over the past three decades, genome sequencing efforts have unveiled the presence of multiple BGCs in many actinomycete strains. A large portion of these BGCs remains dormant or silent, concealing the enigmatic products that are yet to be identified. The profusion of BGCs within each strain, coupled with the widespread distribution of actinomycete strains in terrestrial and marine environments, bestows a still under-tapped reserve for discovering potential drug candidates. An approach that is being exploited to activate silent cryptic BGCs involves the expression of cryptic BGCs in a genetically manipulable, fast-growing heterologous chassis. Identifying and developing a panel of actinomycete hosts spanning the phylogenetic spectrum is imperative for increasing the success rate of heterologous expression of BGCs.

In this thesis, my research objective is to systematically identify, characterize, and optimize proprietary *Streptomyces spp.* strains to establish them as robust heterologous production chassis.

In Chapter 2, I will discuss the characterization and evaluation of the newly isolated *Streptomyces sungeiensis* SD3 as a promising microbial chassis for heterologous production of secondary metabolites. *S. sungeiensis* SD3 possesses several advantageous traits as a microbial chassis, including genetic tractability, rapid growth, susceptibility to antibiotics, and metabolic capability supporting secondary metabolism. Its distinctive phylogenetic position differentiates *S. sungeiensis* SD3 from the *Streptomyces* chassis strains in use today, designating it

as the preferred platform for the expression of BGCs within its phylogenetic clade. Successful expression of pathways from a closely related yet slow-growing strain underscores the advantages of *S. sungeiensis* SD3 as a heterologous expression chassis. Genomic and transcriptomic sequencing unveiled the primary metabolic capabilities and secondary biosynthetic pathways of *S. sungeiensis* SD3, including a previously unidentified pathway responsible for the biosynthesis of Streptazone B1. Validation of CRISPR/Cas9-assisted genetic tools for chromosomal deletion and insertion paved the way for strain improvement by rational genome editing. The inclusion of *S. sungeiensis* SD3 into the heterologous chassis toolkit will facilitate the discovery and production of microbial secondary metabolites.

In Chapter 3, I will introduce the *Streptomyces albidoflavus* MD102 strain as a promising microbial chassis for the heterologous production of secondary metabolites. This strain, closely related to the widely used *S. albidoflavus* J1074, exhibits a compact genome, exceptional genetic tractability, rapid growth, and susceptibility to antibiotics. Whole-genome sequencing has revealed the metabolic capabilities of *S. albidoflavus* MD102, highlighting its versatility in supporting the production of diverse secondary metabolites. Employing CRISPR/Cas9-assited genome editing tools, we successfully created mutant strains with a reduced genome and simplified chromatographic background to enhance the strain's utility as a heterologous host. A distinctive feature of *S. albidoflavus* MD102 that distinguishes it from J1074 and other *Streptomyces* chassis is the presence of metabolic genes in its genome dedicated to the metabolism of aromatic compounds. This characteristic may enable the strain to grow on aromatic substrates, indicating

its potential as a heterologous chassis for the conversion of petrogenic polycyclic aromatic hydrocarbons (PAHs), heterocyclics, and substituted aromatics into valuable secondary metabolites.

In summary, the findings from my Ph.D. research work indicate that *S. sungeiensis* SD3 and *S. albidoflavus* MD102 possess significant potential as microbial chassis for heterologous expression applications. These two proprietary strains will contribute to the progress of synthetic biology and the implementation of microbial fermentation in industrial applications.

Chapter 1 Introduction

1.1 Actinomycetes are prolific producers of specialized metabolites

Natural products are chemical compounds synthesized by living organisms, encompassing plants, animals, and microorganisms (Sorokina and Steinbeck, 2020). These compounds can be categorized into primary and secondary metabolites. Primary metabolites participate in fundamental metabolic pathways crucial for the sustenance of life, whereas secondary metabolites are not integral to core anabolic or catabolic processes (Swallah et al., 2020). Recognized for their structural diversity, secondary metabolites are frequently produced by organisms as chemical defenses against predators or competitors (Kumar et al., 2020; Li and Rebuffat, 2020). Additionally, secondary metabolites play multifaceted roles, such as mediating cell-to-cell communication, responding to environmental stimuli, and shielding organisms from UV light (Schmidt et al., 2019; Yang et al., 2020).

The enduring significance of natural products in contributing to human well-being lies in their role as rich sources of bioactive compounds with medicinal potential (Pham et al., 2019). Over millennia, spanning diverse medical traditions from China to Greece, these natural products have been leveraged for their therapeutic efficacy in addressing a wide range of ailments (Luo et al., 2019; Sorokina and Steinbeck, 2020). Despite advances in synthetic chemistry, natural products maintain their prominence. Notably, derivatives or analogues inspired by natural products account for almost half of all commercially available pharmaceutical agents over the past four decades (Newman and Cragg, 2020).

The enzymatic synthesis inherent to natural products imparts them with unique chemical scaffolds that are challenging to achieve through combinatorial chemistry (Davison and Brimble, 2019). Shaped by evolutionary forces for heightened bioactivity, these compounds possess disproportionate medical relevance whilst exploring a smaller chemical space compared to synthetic compounds (Zabolotna et al., 2021). The continual discovery of natural products remains a valuable source of drug and therapeutic agents due to the novel chemical scaffolds they offer (da Rosa et al., 2020). This may result in the production of more complex natural product compounds characterized by heightened specificity and a concomitant reduction in toxicity and off-target reactions. Considering the long history of natural product utilization, commercial and biomedical interest in these compounds has experienced fluctuations over recent decades, with a current on-going resurgence (Li et al., 2019).

The initiation of the golden age of antibiotic discovery in 1929, marked by Alexander Fleming's discovery of penicillin, led to the identification of various antimicrobial compounds, including vancomycin, tetracycline, erythromycin, and the anti-cancer drug doxorubicin (Dias et al., 2012). However, this fruitful era ended due to challenges associated with the constant rediscovery of known natural products, impeding overall commercial viability (Atanasov et al., 2021). Over the ensuing decades, dwindling support for the labor-intensive processes involved in sample preparation and compound structure elucidation contributed to an overall decline in interest in natural products (Baltz, 2019). Issues related to the isolation of compounds from diverse biological sources with low native production levels

further hindered efficiency and viability. Efforts to circumvent this issue via high-throughput screening proved lackluster, for example with around 500,000 compounds screened by Glaxo Smith Kline (GSK) and 65 high-throughput screening projects initiated by Astra Zeneca yielding no candidates against drug resistant bacteria (Hutchings et al., 2019; Payne et al., 2007).

However, a new resurgence in natural product discovery is currently underway, buoyed by advancements in genome sequencing, mining, and modification tools, and coupled with a shift towards exploring organisms from underexplored environments (Baltz, 2017; Challis, 2008; Clardy et al., 2006). The decline in genome sequencing costs enables the exploration of novel or less studied species using bioinformatic tools, predicting potentially novel secondary natural products and simultaneously dereplicating known compounds (Gomez-Escribano et al., 2016). This approach streamlines the exploration process, avoiding laborious and resource-intensive compound extraction and isolation steps (Rajwani et al., 2021). This also allows for more targeted approaches in prioritizing the organism to be explored, and with the available tools for interrogating even novel species, a revival in natural product discovery is once again taking place. In particular, one genus that germanely illustrates this narrative is the prodigious producer of natural products, actinomycetes (Donald et al., 2022).

Actinomycetes are gram positive bacteria that are among some of the most prolific producers of bioactive natural product compounds (Quinn et al., 2020). In the realm of microbial-derived natural products, actinomycetes play a prominent role, contributing approximately 30% of the reported 70,000 compounds.

Impressively, *Streptomyces* genus alone accounts for 76% of this output (Bérdy, 2012; Donald et al., 2022). As the largest genus within the actinomycetes, *Streptomyces* holds significant importance, as nearly two-thirds of all clinically relevant antimicrobials originate from this group, including chloramphenicol, kanamycin, streptomycin, candicidin and fungichromin (Alam et al., 2022; Bérdy, 2012; Subramani and Sipkema, 2019). Beyond their sheer quantity, *Streptomyces*-derived compounds exhibit a remarkable diversity of bioactivities. Notably, these compounds contribute to the development of anti-viral drugs such as virantmycin and xiamycin D, anti-cancer drugs such as doxorubicin and bleomycin, anti-parasitic drugs exemplified by ivermectin, and immunosuppressive drugs like manumycin and tautomycetin (Alam et al., 2022; Chater et al., 2010; Krause et al., 2020; R. Liu et al., 2018). This multifaceted bioactivity underscores the pivotal role of *Streptomyces*-derived compounds in pharmaceutical advancements.

Streptomyces exhibit a ubiquitous presence across a diverse array of environments, encompassing both terrestrial and aquatic realms, ranging from soil and swamps to oceans, and extending to various plants and animals (Quinn et al., 2020). This acclimatization to a broad spectrum of competitive environments necessitates the development of a diverse array of secondary metabolites to contend with the challenges posed by a wide range of biotic competitors and abiotic factors (Essarioui et al., 2016; Schlatter and Kinkel, 2014).

The synthesis of secondary metabolites in *Streptomyces* is intricately linked to their complex life cycle. Typically, when colonizing a new solid surface, *Streptomyces* initiate their life cycle with germinating spores, followed by the

formation and branching of hyphae (Yagüe et al., 2013). Initial hyphae development involves the formation of compartmentalized mycelia, during the stage generally known as the MI phase (Yagüe et al., 2012). This is then followed by the MII phase, during which the hyphae transition into a multi-nucleated form. It is in this MII phase and the subsequent developmental phase where aerial mycelia is formed, that the majority of secondary metabolic production occurs (Manteca and Yagüe, 2018). The differentiation into late-stage mycelia and the formation of hydrophobic aerial mycelia are typically triggered by nutrient deprivation, which can have pleiotropic effects which leads to the production of different secondary metabolites.

The analysis of *Streptomyces* genomes provides valuable insight into their capacity to synthesize a remarkably diverse range of natural products, even within a single species. *Streptomyces* genomic DNA is typified by a high guanine-cytosine (GC) content, and overall constitutes some of the largest prokaryotic genomes, usually 8-12 megabases (Mb) in size (Lee et al., 2020). Genes involved in the production of secondary metabolites are often organized into biosynthetic gene clusters (BGCs), within which they may fulfill various roles, including synthesis, regulation, and resistance (Lee et al., 2020). These genes result in a significant proportion of the *Streptomyces* genome being dedicated to secondary metabolism (around 5 to 10%), with an average *Streptomyces* genome harboring 20-40 BGCs (Baltz, 2008; Culp et al., 2019). When the first *Streptomyces* genomes were sequenced, it was of great surprise that the number of BGCs was greater than predicted from their observed metabolome (Katz and Baltz, 2016). It has now been established that less than 10% of *Streptomyces* BGCs are adequately expressed

under routine lab culture and fermentation conditions to be observed and identified (Baltz, 2016; Katz and Baltz, 2016).

Hence, the full potential of *Streptomyces* natural products remains to be fully explored. Leveraging the widespread distribution of *Streptomyces* species across diverse environments, rare and diverse species isolated through bioprospecting efforts can be sequenced and subjected to bioinformatic analysis (Donald et al., 2022; Li et al., 2019). Within each *Streptomyces* species, a considerable number of silent BGCs could be awakened using an expanding repertoire of genome editing tools, with recent advancements facilitated by CRISPR technology. Consequently, a substantial reservoir of untapped *Streptomyces* natural products with potential for valuable and medically relevant bioactivity still exists, awaiting exploration (Lacey and Rutledge, 2022; Nepal and Wang, 2019).

1.2 Targeting the cryptic biosynthetic genes clusters of actinomycetes to discover novel specialized metabolites

As highlighted in the preceding chapter, a significant impediment arises from the prevalence of silent BGCs (Liu et al., 2021; Xia et al., 2020). These clusters could be subject to strict regulatory control and suppression, or alternatively, expressed albeit weakly under standard laboratory conditions (Katz and Baltz, 2016). Such weakly expressed clusters stymie efforts in detecting novel compounds during High-Performance Liquid Chromatography (HPLC) analysis, and also may not yield sufficient material for downstream purification and structural identification. Therefore, the activation of these silent clusters becomes imperative

to unlock the full potential of *Streptomyces* natural products. To address this, various approaches have thus far been developed to activate dormant BGCs.

One effective approach for activating silent BGCs is the "one strain many compounds" (OSMAC) strategy. This strategy capitalizes on the observation that altering culture conditions is often sufficient to generate markedly different metabolic profiles (Pan et al., 2019; Romano et al., 2018). By simply manipulating the growth conditions of a microbe, a variety of different compounds can be produced from a single species. Modifying physical or chemical perimeters such as media composition, shaking speed, aeration, temperature, and pH, can simulate the signals necessary required for effective activation of silent clusters.

For instance, changing media compositions by altering the availability of the carbon, nitrogen, sulfur, phosphorous, trace salts and trace metals in either liquid or solid media can lead to diverse outcomes. In one example within in the same *Streptomyces* species, the use of glucose as the carbon source led to the production of chaxamycins A and B, whilst substituting it with glycerol instead led to the production of chaxamycins C and D (Rateb et al., 2011; Scott Zarins-Tutt et al., 2016). Nutrient depletion can also relieve negative regulation of secondary metabolites. For example, glucose has been implicated in the negative regulation of beta-lactam antibiotic production due to carbon catabolite repression (Sánchez et al., 2010).

These factors may influence changes in growth conditions and entry into different phases of the *Streptomyces* life cycle, or provide the necessary metabolic precursors required to divert metabolic flux to the production of secondary

metabolites (Liao et al., 2015). Secondary metabolite production is often induced upon nutrient exhaustion, whilst conversely the supplementation of specific limiting elements or metals during biosynthesis can remedy low production of specialized compounds (Romano et al., 2018). Additionally, supplementation of chemical elicitors is another viable way for activating silent clusters (Pinedo-Rivilla et al., 2022). These chemicals mimic the signalling molecules found in the *Streptomyces* natural environment and activate the relevant signalling pathways controlling secondary metabolite production. For example, N-acetylglucosamine (GlcNAc) is a monomer of chitin and is able to serve as an activator by inactivating DasR, a pleiotropic transcriptional repressor involved in repressing transcription of secondary metabolite genes (Abdelmohsen et al., 2015; Tomm et al., 2019). Interestingly, global control of this activation/repression is dependent on the richness of the media constitutions, and release of DasR repression upon GlcNAc elicitation only occurs under poor media conditions (Świątek et al., 2012). Other chemical elicitors such as γ-butyrolactone, DMSO, heavy metals, and sub-inhibitory antibiotics can also be utilised (Zong et al., 2022). Overall, the OSMAC strategy is a straightforward and a priori approach that does not require deep understanding of the complex regulatory networks that govern the native strains' metabolism. However, having such a multiplicity of factors and parameters involved and often in an unpredictable manner means that execution in screening can be laborious and resource intensive.

1.3 Heterologous expression and chassis development

In addition to manipulating fermentation conditions, an effective strategy for activating silent BGCs involves genetic element refactoring, as demonstrated by (Li et al., 2021). This approach encompasses the modification of promoter regions within biosynthetic genes, along with the targeted overexpression or deletion of regulator genes, providing a rational and precise means to refactor BGCs (Liu et al., 2021). However, a substantial challenge arises from the genetic intractability observed in the majority of wildtype *Streptomyces spp*., limiting their susceptibility to modification using conventional genetic and synthetic biology tools (Myronovskyi and Luzhetskyy, 2019; Xu et al., 2022). This challenge is even more pronounced in less common actinomycete strains, such as Micromonospora (Parra et al., 2023).

Crafting customized protocols for each genetically intractable species discovered in the wild proves both time-consuming and challenging to scale up. Furthermore, numerous wildtype strains exhibit slow growth rates under laboratory culture conditions, posing an additional hurdle to effective strain engineering (Liu et al., 2021; Xu et al., 2022). The combined impact of these obstacles renders genetic refactoring approaches generally impractical for many strains.

To surmount the challenges posed by genetic intractability and sluggish growth rates in wildtype strains, employing heterologous expression of BGCs in a well-characterized chassis or host emerges as a widely adopted and robust alternative (Ahmed et al., 2020). In this methodology, the genomic DNA containing the target BGC can be cloned into suitable vectors, such as bacterial artificial

chromosomes (BACs), which carry partition genes essential for the stable propagation of large constructs (ranging from 20 to 200 kb) during the cloning process (Hwang et al., 2021; Kang and Kim, 2021). Alternatively, cosmid library constructions may be utilized, even incorporating environmental DNA to establish a comprehensive BGC library. Subsequently, the vectors housing the desired BGC are introduced into the heterologous chassis, often integrating them to ensure the stable maintenance of exogenous elements during subsequent procedures.

The efficacy of this heterologous expression approach crucially relies on the judicious selection of the chassis. Commonly employed heterologous hosts in industrial settings encompass well-established organisms such as *Escherichia coli*, *Bacillus subtilis*, and *Saccharomyces cerevisiae* - chosen for their rapid growth, extensive prior characterization, and the presence of large research communities with developed tools for their manipulation (Hwang et al., 2021; Yang et al., 2022). Notably, in the context of actinomycete-derived natural products, successful heterologous expression has been accomplished for specific compounds. Instances include the expression of BGCs for erythromycin and tautomycetin in *E. coli* and 6-dEB in *B. subtilis* (Gao et al., 2021). These achievements underscore the adaptability and utility of heterologous expression systems in facilitating the production of diverse natural compounds.

However, both context specific and phylogenetic challenges prevent widespread adoption of these commonly used chassis for *Streptomyces* biosynthetic pathway expression. The substantial phylogenetic distance between these commonly used hosts and *Streptomyces spp*. introduces significant interspecies

differences. Issues such as promoter-sigma factor compatibility, the high GC content of *Streptomyces* genes, codon bias, and issues related to protein folding all present challenging obstacles to successful heterologous expression (Myronovskyi and Luzhetskyy, 2019; Yue et al., 2023). Additionally, distinct native regulatory networks in different species hinder the correct expression of *Streptomyces* genes (Kang and Kim, 2021). Failures in the proper folding and assembly of multienzyme complexes that are sensitive to perturbation further impede correct biosynthesis (Hwang et al., 2023). The lack of an appropriate supportive enzymatic and metabolic background may also hinder the successful heterologous expression of natural products (Hwang et al., 2021).

Critical post-translational modifications required for the activation of nonribosomal peptide synthetases and polyketide synthases, such as the transfer of the pantetheine prosthetic group by 4'-phosphopantetheinyl transferases (PPTase) to convert apo-ACP to the activated holo-ACP form, also pose challenges (Otani et al., 2022). The sufficient availability of malonyl-CoA, acetyl-CoA and NADPH is vital for supporting metabolic flux through natural product biosynthesis. Native levels of less common metabolic precursors like methylmalonyl-CoA and non-canonical amino acids must also be sufficiently supportive enough for high levels of production (Bu et al., 2021). For instance, *E. coli* not being a native polyketide producer, often struggles with low methylmalonyl-CoA availability and the lack of functional expression of *Streptomyces* polyketide genes, especially when post-translational modification is required (Liu et al., 2022; Zhan et al., 2023). Upon successful biosynthesis of the exogenous product, accumulation of metabolites

could also result in cell toxicity (LeBlanc and Charles, 2022). In combination, all these difficulties paint a challenging picture for heterologous expression of *Streptomyces* natural products in commonly established non-*Streptomyces* hosts.

1.4 *Streptomyces* as heterologous hosts

To surmount the challenges arising from phylogenetic distance, *Streptomyces spp*. emerge as suitable and effective hosts for heterologous expression (Kang and Kim, 2021; Myronovskyi and Luzhetskyy, 2019; Nepal and Wang, 2019). Defined by their intrinsic capacity to be prolific producers of natural products, *Streptomyces spp*. exhibit excellence as hosts for heterologous production - successfully expressing *Streptomyces'* natural products, including the activation of silent BGCs. This proficiency is attributed to their inherent possession of native BGCs and a diverse array of metabolic pathways capable of supplying essential precursors (Hwang et al., 2021; Li et al., 2021). The presence of relevant post-translational modifications required for the biosynthesis of structurally complex secondary metabolites further augments the compatibility of *Streptomyces spp*. with the expression of natural products (Myronovskyi and Luzhetskyy, 2019).

Beyond the accomplishment of successful natural product expression, *Streptomyces spp*. demonstrate heightened tolerances to bioactive and toxic natural products (Hwang et al., 2021). For instance, the production of botryococcene, a terpene-based biofuel commonly extracted from *Botryococcus braunii* cultures, was notably enhanced in *Streptomyces reveromyceticus* SN-593, yielding 212 ± 20 mg/L, surpassing yields of 60 mg/L in yeast and 110 mg/L in native *Botryococcus braunii* (Khalid et al., 2017). This underscores the suitability of *Streptomyces* hosts

not only for natural product expression but also for their robustness in handling and producing bioactive compounds.

Despite the complexity of their life cycles and high GC content genomes, recent years have witnessed significant advancements in rendering the handling and manipulation of *Streptomyces spp.* more accessible. The repertoire of genetic tools available for manipulating *Streptomyces* has expanded, particularly with the development of several CRISPR-Cas9 genome editing vectors (Tao et al., 2018; Zhao et al., 2020). Additionally, the increased availability of 3rd generation sequencing methods, such as Pacbio SMRT or Oxford Nanopore MinION, has effectively addressed challenges related to high GC content and the presence of repetitive regions in *Streptomyces* genomes (Heng et al., 2023).

In light of these improvements, several *Streptomyces* hosts have been developed, including *S. coelicolor*, *S. lividans*, *S. avermitilis*, *S. albidoflavus*, *S. chattanoogensis*, *S. griseofuscus*, *S. atratus* and *S. venezuelae*, as documented in the following studies (Ahmed et al., 2020; Gomez-Escribano and Bibb, 2011; Gren et al., 2021; Hwang et al., 2021; Liu et al., 2016; Myronovskyi et al., 2018; Yang et al., 2022).

1.4.1 *Streptomyces coelicolor*

S. coelicolor A3(2) stands out as the model strain for studying actinomycete genetics, owing to its extensively characterized 8.7 Mbp genome, making it a traditionally choice as a heterologous host (Gomez-Escribano and Bibb, 2014; Hopwood, 1999; Schlimpert and Elliot, 2023). One of the notable features of *S. coelicolor*, which led to interest in studying it in the first place, is its production of

colored pigments actinorhodin and undecyleprodigiosin (Kang and Kim, 2021). This provides an excellent platform for interrogating genetic regulatory systems governing secondary metabolite production in *Streptomyces*, and additionally aids in chromogenic selection of genetic mutants. Significant work has been done to engineer *S. coelicolor* to generate improved and optimised strains like *S. coelicolor* M1154 (Δact Δred Δcpk Δcda rpoB[C1298T] rpsL[A262G]) (Gomez-Escribano and Bibb, 2014; Hwang et al., 2021). Based on the initial *S. coelicolor* M145 strain which lacks its native SCP1 and SCP2 plasmids, three endogenous BGCs (*act*, *red, cda* and *cpk*) were deleted and point mutations were introduced in RNA polymerase β-subunit and ribosomal protein S12 genes *rpoB* and *rpsL* respectively that increase secondary metabolite production with no notable impairment to growth, thereby generating the final *S. coelicolor A3(2)* M1154 strain. Notable enhancements to the yields of heterologously expressed BGCs in *S. coelicolor* M1154 was observed, for instance production of chloramphenicol was found to be 20-40 times greater as compared to that of the parent *S. coelicolor* M145 strain. Another effort to engineer *S. coelicolor* include the deletion of all 10 PKS and NRPS BGCs in *S. coelicolor* M145, resulting in the genome minimized strain *S. coelicolor* ZM12 which produced higher levels of actinorhodin (Lee et al., 2021). Given this extensive pedigree, a diverse variety of BGCs, even from phylogenetically diverse actinobacteria have been successfully expressed in *S. coelicolor* including actinoallolides, enterocin, cacibiocin, vancoresmycin, fluostatin, GE2270, lyngbyatoxin and epothilone (Gomez-Escribano and Bibb, 2014; Myronovskyi and Luzhetskyy, 2019).

However, certain limitations stymie the use of *S. coelicolor* as a universal chassis. Its slower growth rate compared to strains like *S. albidoflavus*, and less desirable morphological features, such as clumpy and less dispersed mycelia in liquid media, pose challenges (Myronovskyi and Luzhetskyy, 2019). More significantly, the levels of heterologous production in *S. coelicolor* are generally low, which impedes practical downstream compound detection and isolation efforts (Kang and Kim, 2021).

1.4.2 *Streptomyces lividans*

S. lividans is a close relative of *S. coelicolor*, with a slightly smaller 8.3 Mbp genome. It was initially identified during studies of the phage ΦC31, of which the phage components were later extensively developed into a large series of *Streptomyces* integrative vectors (Hopwood, 1999; Hopwood et al., 1983; Nepal and Wang, 2019). A distinctive difference that *S. lividans* possesses is it lacks methylation-dependent restriction, which allows it to readily accept methylated DNA, thereby increasing its ease of transformation with foreign DNA (Hwang et al., 2021). In contrast to *S. coelicolor*, the *act*, *red*, and *cda* BGCs are transcriptionally silent in *S. lividans* (Peng et al., 2018). *S. lividans* also exhibits low endogenous protease activity, making it suitable for the production of recombinant proteins or peptide natural products, in particular nonribosomal peptides (NRP) and ribosomally synthesized and post-translationally modified peptides (RiPP) (Myronovskyi and Luzhetskyy, 2019). Moreover, its efficient secretion of heterologous proteins into the culture media further enhances its utility (Hamed et al., 2018).

The *S. lividans* TK24 strain, featuring a mutation in the ribosomal protein S12 gene rpsL (K88E) which induces expression of the silent *act* BGC, is commonly used as a heterologous host (Peng et al., 2018). Several genetically engineered *S. lividans* strains have hence been developed, utilizing *S. lividans* TK24 as a starting strain. For instance, *S. lividans* SBT5 was developed by deleting the *act*, *red*, and *cda* BGCs from *S. lividans* TK24, and incorporating the positive global regulator gene *afsRS* from *S. clavuligerus* (Xu et al., 2016). Building on this, *S. lividans* LJ1018 was then developed by deleting the negative global regulator gene *wblA* from *S. lividans* SBT5, and inserting two codon-optimized multi-drug efflux pump genes *lmrA* and *mdfA* (Whitford et al., 2021). This led to the drastic improvement of the yields of the type II polyketide murayaquinone to 10.6 mg/L, 96 times higher than *S. lividans* SBT5 (Peng et al., 2018). Another developed strain, *S. lividans* ΔYA11, involved the deletion of 11 endogenous totaling 228.5 kb, generating a genome minimized strain with a cleaner and simplified metabolic background for ease of product identification and isolation, and also exhibited slightly better biomass accumulation (Ahmed et al., 2020). Additionally, it also had the introduction of two additional phage ΦC31 *attB* sites for integration of integrative vectors containing complementary phage ΦC31 *attP* sites, which allows additional copies of BGCs to be quickly transformed into the host. These modifications led to a fourfold increase in the production of the lanthipeptide deoxycinnamycin. However, the production yield of the nucleoside antibiotic tunicamycin remained similar between TK24 and ΔYA11, whilst the aromatic polyketide griseorhodin was not produced at all, and only observed in *S. albidoflavus* instead.

1.4.3 *Streptomyces albidoflavus*

S. albidoflavus J1074, recently reclassified as *Streptomyces albidoflavus* from the originally designated *Streptomyces albus*, is a fast-growing *Streptomyces* strain that is one of the most widely used heterologous host for secondary metabolites (Myronovskyi and Luzhetskyy, 2019). Its fast growth rate is supported by one of the smallest genomes (6.8 Mb) within the *Streptomyces* genus and seven rRNA operons (Zaburannyi et al., 2014). Derived from the original *S. albidoflavus* G strain, *S. albidoflavus* J1074 is defective in its SalI restriction-modification system, thereby enabling ready transformation of exogenous genes. Additionally, the silent state of most of its endogenous BGCs provides a naturally simplified metabolic background.

S. albidoflavus J1074 has demonstrated exceptional efficacy as a host for polyketide compounds, particularly aromatic polyketides. Various BGCs, such as those responsible for fredericamycin, grecocyline, griseorhodin, holomycin, isomigrastatin, landomycin, napyradiomycin and thiocoraline, have all been successfully expressed in *S. albidoflavus* J1074 (Myronovskyi and Luzhetskyy, 2019). Importantly, in many instances heterologous expression of BGCs in *S. albidoflavus* J1074 resulted in higher yields as compared to other hosts. For instance, isomigrastatin was produced in *S. albidoflavus* J1074 two times higher as *S. coelicolor* and *S. lividans* and 10 times higher than *S. avermitilis* SUKA (Yang et al., 2011). In another example, the integration of cryptic BGCs did not result in any expression in other hosts, including *S. coelicolor* M1152, but only *S. albidoflavus* J1074 yielded the production of landepoxcins A and B (Owen et al., 2015). *S.*

albidoflavus J1074 contains an additional pseudo attB site that is functional, allowing for an additional integration of ΦC31 based vectors at high frequency (Myronovskyi et al., 2018). This could provide an explanation for the higher yields observed during heterologous production compared to other hosts (Bilyk and Luzhetskyy, 2014).

Several engineered *S. albidoflavus* J1074 strains have thus far been developed. One of the more extensively engineered strains is *S. albidoflavus* Del14, generated from the systematic deletion of 15 predicted native BGCs, resulting in a 500 kb genome minimized cluster-free strain, which displayed an even cleaner metabolic background (Myronovskyi et al., 2018). Additionally, a cryptic BGC from *Frankia alni* ACN14a was succesfully expressed, leading to the discovery of a novel compound fralnimycin. This strain then formed the basis for subsequent modifications, including the addition of two additional ΦC31 *attB* sites resulting in a total of four ΦC31 *attB* sites in the resultant *S. albidoflavus* B4 strain. Heterologous production of didesmethylmensacarcin and didemethoxyaranciamycinone was higher in the modified *S. albidoflavus* strains as compared to *S. coelicolor* (M1154 and M1152), whilst the production of tunicamycin B2, moenomycin M, griseorhodin A and pyridinopyrone A was successful in the *S. albidoflavus* strains, while not even being detected in the *S. coelicolor* strains. However, cinnamycin was only expressed in *S. coelicolor* strains but not in the *S. albidoflavus* ones. Additionally, NRPS or RiPPs BGCs generally have a lower success rate of expression in *S. albidoflavus* (Myronovskyi and Luzhetskyy, 2019).

1.4.4 *Streptomyces avermitilis*

S. avermitilis is best known for its role in the industrial production of the anthelmintic avermectin (Komatsu et al., 2013). Its 9.03 Mbp genome was the second streptomyces genome to be reported. One of the main attractive features of *S. avermitilis* is its short terminal inverted repeats (TIRs) at only 49 bp, in contrast to the 10-100 kb TIRs found in other *Streptomyces spp.*, thereby contributing to a more stable chromosome.

A series of iterative improvements to *S. avermitilis* has generated a panel of *S. avermitilis* SUKA mutants. A large 1.51 Mbp deletion of chromosomal region and subsequent cre-loxP based BGC deletion resulted in *S. avermitilis* SUKA17 and SUKA22 strains that have the native avermectins, filipins, oligomycins, goesmin, neopentalenolactone and carotenoid BGCs deleted (Komatsu et al., 2010). This resulted in a 81.46% genome size strain as compared to the original *S. avermitilis* strain. Subsequently, *S. avermitilis* SUKA variants were further created upon the introduction of the positive transcriptional regulator gene rapH and promiscuous Sfp-type phosphopantetheinyl transferase (Kim et al., 2018; Kudo et al., 2020).

S. avermitilis SUKA hosts have been extensively validated for their utility, including the expression of various *Streptomyces* BGCs producing aureothin, chloramphenicol, holomycin, kasugamycin, leptomycin, novobiocin, oxytetracycline, ribostamycin, and streptomycin (Komatsu et al., 2013; Myronovskyi and Luzhetskyy, 2019). Notably, the production of streptomycin in *S. avermitilis* SUKA5 at 180 mg/L was higher even compared to the native *S. griseus*

IFO13350 strain (30 mg/L). Moreover, terpenoids from both microbial and plant origins have successfully been produced in *S. avermitilis* SUKA17, including pentalenolactone from *S. exfoliatus* UC5319 and *S. arenae* TÜ469, and amorpha-1,4-diene using a codon optimised sesquiterpene synthase from *Artemisia annua* (Komatsu et al., 2013; Tarasova et al., 2023).

1.4.5 Other *Streptomyces* hosts

S. venezuelae stands out with the shortest reported doubling time among *Streptomyces* (~1 hr) and high transformation efficiency, and hence has been used as a heterologous host (Kim et al., 2015). *S. venezuelae* ATCC15439 has a genome of 9.05 Mbp and similarly to *S. albidoflavus,* possesses 7 rRNAs. A number of *S. venezuelae* mutant strains have been developed, including *S. venezuelae* YJ003 with the *des* BGC deleted, *S. venezuelae* DHS2001 with the *pik* genes deleted, and *S. venezuelae* YJ028 with both the *pik* and *des* BGC deleted (Kim et al., 2015). More recently *S. venezuelae* ATCC 10712, a strain able to produce the angucycline antibiotics jadomycins when exposed to ethanol shock was engineered to utilize the ethanol triggered regulatory logic for heterologous production (Qiu et al., 2024). *afsQ1*, *bldD*, *bldA*, and *miaA* were identified to be involved in this regulatory logic, and utilised as an ethanol-induced orthogonal amplification system to increase the titer of jadomycins by 2-fold. A number of streptomyces natural products have been successfully expressed in *S. venezuelae*, including thaxtomins, doxorubins, oxytetracycline, actinorhodin, oxytetracycline, kanamycins A-D, and pseudoribostamycin (Kurumbang et al., 2011; Park et al., 2011; Qiu et al., 2024; Yin et al., 2016). Most notably, despite the typical difficulties in expressing plant

biosynthetic genes in prokaryotic hosts, a number of plant natural products have been expressed in *S. venezuelae*, including apigenin, chrycin, naringenin, pinocembrin, pinosylvin and resveratrol (Park et al., 2009).

S. chattanoogensis L10 is an industrial strain used for the production of natamycin, and was investigated for its ability to be developed as a heterologous host, resulting in the engineered strain *S. chattanoogensis* L321 (Bu et al., 2019). Genome minimization was carried out using Cre/loxP recombination to delete 0.7 Mbp of putative non-essential genomic regions closer to the ends of the linear chromosome arms. This led to improvements in cleaner metabolic background, increased intracellular ATP and NADPH/NADP+ levels, and 2-3 fold increase in production of indigoidine and actinorhodin. Intriguingly, the genome minimized L321 strain was also able to be transformed with CRISPR-Cas9 vectors that were not able to be introduced into the original L10 strain, which the authors proposed the loss of the endogenous type I-E CRISPR-Cas9 system as the explanation.

S. griseofuscus DSM 40191 was initially isolated as a producer of antibiotic compounds lankacidins A and C (Gren et al., 2021). It was recently investigated as a potential heterologous host given its relatively fast growth and ease of genetic manipulation. Its phenotype regarding the ability to utilise various carbon, nitrogen, sulphur and phosphorus sources was characterized, and genome mining was carried out to ascertain its potential as a host for secondary metabolite production. A panel of genetic tools was tested including integrative and replicative vectors, CRISPR-cas9 mediated gene deletion, and CRISPR-cBEST base editor. Strain improvement was also carried out, involving the curing of two of its three native plasmids as well

as the deletion of pentamycin and an unknown NRPS BGCs to generate *S. griseofuscus* DEL2, corresponding to a 500 kb reduction in genome size.

S. atratus SCSIO ZH16 is a deep-sea streptomyces strain recently investigated as a production host with an emphasis on marine derived BGCs (Yang et al., 2022). It was initially isolated from the South China Sea at a depth of –3536m. It's lack of fastidious culture requirements for a marine streptomyces, and fast growth and sporulation led it to be investigated for its potential as a chassis. Phylogenetic analysis revealed that it occupied a separate clade compared to other commonly used Streptomyces hosts, and it was successful in heterologously expressing 7 BGCs, 6 of which were marine derived. Double cross-over knock out approach was used to remove the production of the native atratumycin and ilamycins, generating the strain MGCEP 1.0.

1.5 Aims of the Ph.D. project

In spite of significant advancements in optimizing model *Streptomyces* strains as heterologous chassis, the quest for an ideal strain capable of expressing diverse BGC types with high success rate remains elusive (Ahmed et al., 2020; Hwang et al., 2021; R. Liu et al., 2018; Yang et al., 2022). The realization of a universal "super host" for *Streptomyces* is impeded by the challenges associated with the inherent and intractable genetic and metabolic differences between the heterologous host and native producer, posing challenges in predicting the compatibility of a host with a specific BGC (Hwang et al., 2021). Experimental observations indicate that production titers and the success of heterologous expression are highly unpredictable across various model *Streptomyces* strains (Ahmed et al., 2020; Hwang et al., 2021). Consequently, employing a diverse panel of heterologous hosts is imperative to increase the success rate of heterologous expression, necessitating the expansion of available *Streptomyces* strains for use as heterologous hosts (Hwang et al., 2021; Liu et al., 2016; Myronovskyi and Luzhetskyy, 2019).

Hence, there is a need to develop strains with characteristics that are phylogenetically different from currently existing *Streptomyces* chassis strains to augment the diversity of the available hosts (Figure 1). Such diversification can be achieved by exploring *Streptomyces* strains isolated from disparate environments that poses evolutionary selective pressure to generate distinct genetic makeup and metabolic background, such as those found in marine or swamp environments (Yang et al., 2022).

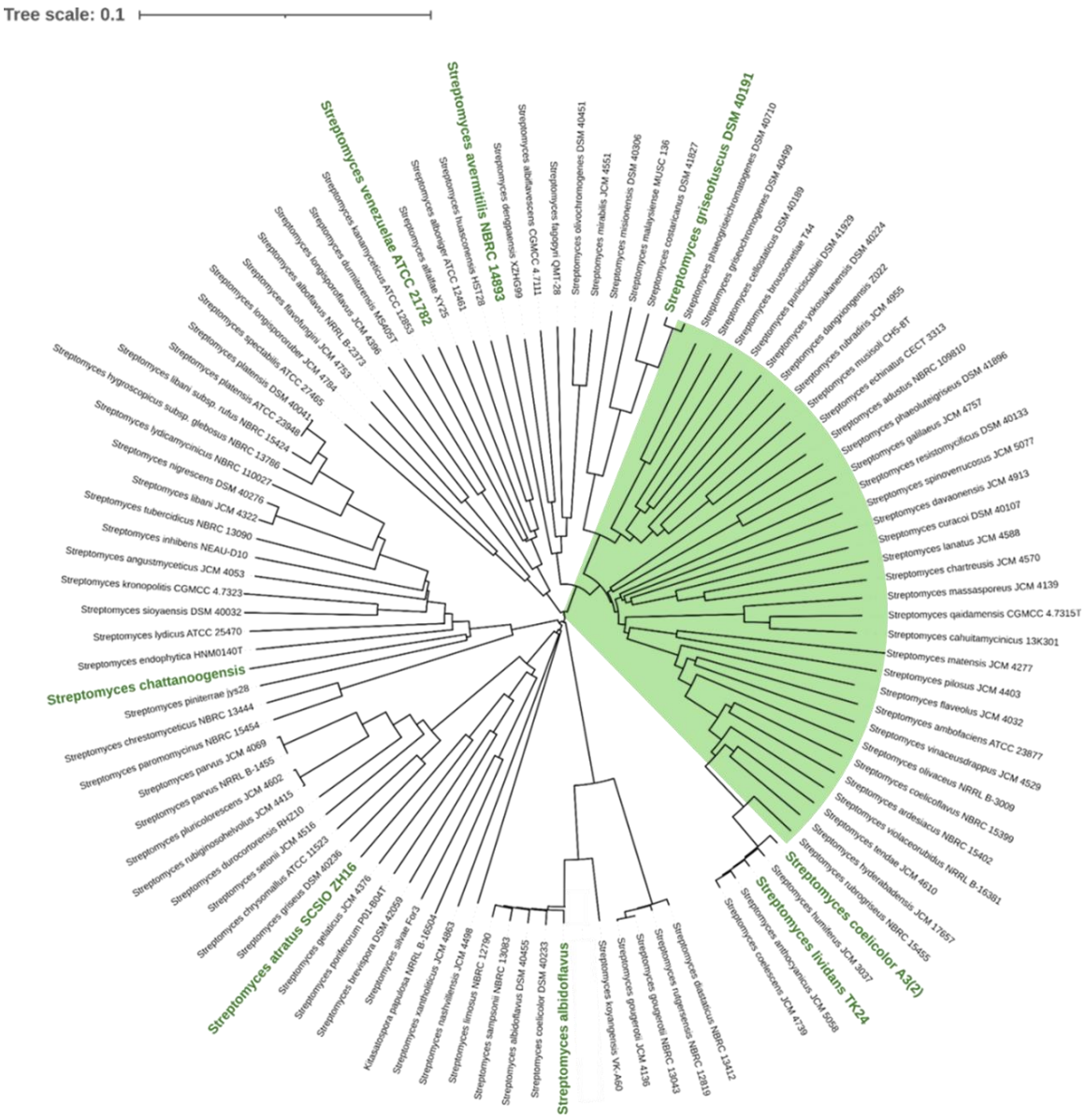


Figure 2. Phylogenetic tree of commonly used *Streptomyces* chassis. The currently available *Streptomyces* chassis are highlighted in green. The green arc denotes the phylogenetic space that could be filled in by *S. sungeiensis* SD3. The phylogenetic tree was constructed using TYGS (Meier-Kolthoff and Göker, 2019) to calculate a Genome BLAST Distance Phylogeny (GBDP) and visualized using Iterative Tree Of Life iTOL (Letunic and Bork, 2021).

The development of novel *Streptomyces* strains as heterologous chassis for natural product production necessitates their compatibility with the Design-Build-Test-Learn (DBTL) cycle, leveraging contemporary automation and bio-foundry capabilities for iterative improvements (Whitford et al., 2021). This is important due to the additional necessities of refactoring BGCs for heterologous expression, since otherwise success rates remain low, reaching 30% (Gren et al., 2021; Hwang et al., 2021; Myronovskyi and Luzhetskyy, 2019). Hence, potential host strains should possess fast growth to facilitate both genetic modification and culturing work, especially given the slower growth rates of streptomyces compared to other commonly used microbial chassis like yeast or *E coli*, and in particular if the native streptomyces producer strain is slow growing (Gren et al., 2021; Liu et al., 2016; Yang et al., 2022). Additionally, the potential strain should be amenable and easy to genetically modify, which is vital given that many *Streptomyces spp*. are poorly genetically tractable. Commonly used integrative vectors for transforming heterologous BGCs, along with genome editing tools for refactoring BGCs like CRISPR-Cas9 should be explored and validated for use (Ahmed, 2019; Gren et al., 2021). The strain should also have high conjugation efficiency for these commonly used vectors (Yang et al., 2022). Since the main purpose of the strain is for the discovery and production of natural products, the ability to easily identify new metabolites via HPLC is advantageous. Hence, the strain should have a relatively clean metabolite profile to simplify the detection and subsequent purification steps of the compound of interest (Ahmed, 2019; Ahmed et al., 2020; Myronovskyi et al., 2018).

Our laboratory recently isolated more than 200 actinomycete strains derived from diverse, previously underexplored habitats from around Singapore. Employing a combination of methodologies, we utilized specialized media containing materials sourced from native environments and implemented various plating techniques, including traditional plating and in-situ cultivation, to enhance the diversity and rarity of the isolated strains (Low et al., 2018). These strains were then screened for their inhibitory bioactivities against various common pathogens like *Pseudomonas aeruginosa*, methicillin-resistant *Staphylococcus aureus*, *Bacillus subtilis* and *Klebsiella pneumoniae*, to identify novel biomedically relevant natural products. This screen can also form the basis from which potentially suitable *Streptomyces* strains could be identified to be explored as novel heterologous chassis. Notably, being newly isolated, these strains would have no intellectual property limitations commonly associated with currently employed academic or industrial strains. Furthermore, this provides the unique advantage of selecting strains based on desired properties suited for heterologous chassis development, circumventing the constraints imposed by legacy academic strains and their inherited artefacts.

The objective of my Ph.D. research project is to systematically identify, characterize, and optimize proprietary *Streptomyces spp*. strains from our microbial collection, intending to establish them as robust heterologous production chassis for producing microbial natural products. The initiative will leverage the rich diversity within our local *Streptomyces* strains, employing a methodical screening

process to identify candidates that fulfill the criteria of resilient and efficient microbial hosts.

The first phase of the project involves the screening of the local *Streptomyces* strains to identify those exhibiting desirable traits as potential heterologous hosts. These criteria may include robust growth, high productivity, and compatibility with genetic manipulation techniques. Following the selection of promising candidates, a thorough characterization process will be initiated, encompassing in-depth analyses of their genomes and metabolic features. This characterization phase aims to unravel key genetic and biochemical attributes that contribute to their suitability as expression hosts for *Streptomyces* natural products. Subsequently, practical protocols will be developed for the manipulation and handling of the selected strains. This entails the establishment of standardized procedures for genetic engineering, fermentation, and other relevant processes to ensure reproducibility and scalability. Practical considerations, such as growth conditions, media composition, and cultivation techniques, will be optimized to enhance the strains' performance as efficient heterologous expression hosts. The project's final phase involves iterative improvements to the selected strains, employing a feedback-driven approach to enhance their capabilities continually. This optimization process will focus on refining the strains' genetic and physiological features to elevate their performance in expressing *Streptomyces* natural products. Through this iterative refinement, the aim is to establish a versatile and potent heterologous expression platform with expanded capabilities, laying the foundation for broader applications in the production of valuable natural products.

Chapter 2 Characterization and development of *Streptomyces sungeiensis* SD3 as a heterologous host

2.1 Introduction

Microbial natural products constitute a significant proportion of pharmaceutical agents utilized in the treatment of a wide spectrum of human afflictions (Newman and Cragg, 2020). Among the microbial producers, actinomycetes stand out as highly prolific sources, producing many clinically important antibacterial, anti-fungal, anti-parasitic, anticancer, and immunosuppressive therapeutics (Krause et al., 2020; R. Liu et al., 2018). Microbial natural products are secondary metabolites synthesized by the enzymes encoded by biosynthetic gene clusters (BGCs) (Lee et al., 2020). Over the past three decades, genome sequencing efforts have unveiled the presence of multiple BGCs in many actinomycete strains. A large portion of these BGCs remains dormant or silent, concealing the enigmatic products that are yet to be identified (Xia et al., 2020). The profusion of BGCs within each strain, coupled with the widespread distribution of actinomycete strains in terrestrial and marine environments, bestows a still under tapped reserve for discovering potential drug candidates (Lacey and Rutledge, 2022; Nepal and Wang, 2019).

Many cryptic BGCs remain dormant under typical laboratory culture conditions due to repressed gene expression, limited precursor availability, or other inhibitory factors (Liu et al., 2021). Activating silent BGCs to disclose their cryptic secondary metabolites is often unfruitful, with many reported successes owing to serendipitous discovery of elicitors or fermentation conditions. Rational genetic

approaches such as BGC refactoring suffers from low success rate, mainly due to the lack of understanding of the multi-layered genetic regulatory networks governing gene expression (Myronovskyi and Luzhetskyy, 2019). BGC activation in the native microbial producers is further complicated because of the slow growth rates and limited genetic manipulability that afflict many non-domesticated actinomycete strains (Myronovskyi and Luzhetskyy, 2019). The commonly employed genetic tools and vectors developed for model actinomycetes, primarily *Streptomyces* species, are often incompatible with wild-type actinomycetes (Xu et al., 2022).

An alternative approach that has been exploited to activate silent cryptic BGCs involves the expression of cryptic BGCs in a genetically manipulable, fast-growing heterologous chassis (Ahmed et al., 2020). Several model Streptomyces strains have been rendered to genome engineering and strain optimization, encompassing *S. coelicolor*, *S. lividans*, *S. avermitilis*, *S. albidoflavus*, *S. chattanoogensis*, *S. griseofuscus*, *S. atratus*, and *S. venezuelae* (Ahmed et al., 2020; Gomez-Escribano and Bibb, 2011; Hwang et al., 2021; Liu et al., 2016; Myronovskyi et al., 2018; Whitford et al., 2021; Yang et al., 2022). Initially, there was optimism regarding the prospect of identifying a universal *Streptomyces* heterologous host capable of expressing the BGCs from phylogenetically diverse actinomycetes (Ahmed et al., 2020; Hwang et al., 2021; R. Liu et al., 2018; Yang et al., 2022). However, this aspiration turned out to be insurmountable, likely because of the underestimated differences in regulatory mechanism and cellular environment between the host chassis and native producers. At present, whether a

BGC can be successfully expressed in a heterologous *Streptomyces* host to produce the anticipated secondary metabolite remains largely unpredictable (Hwang et al., 2021; Myronovskyi and Luzhetskyy, 2019). Nevertheless, it has been recognized that the likelihood of successful production of secondary metabolites in a heterologous host increases if the heterologous chassis shares a closer phylogenetic relationship with the native producer (Yue et al., 2023). Consequently, the strategy of identifying and developing a panel of *Streptomyces* hosts spanning the phylogenetic spectrum has emerged as a more pragmatic approach, as opposed to the pursuit of a single universal heterologous *Streptomyces* host (Hwang et al., 2021; Liu et al., 2016; Myronovskyi and Luzhetskyy, 2019).

In this study, we identified the *Streptomyces sungeiensis* SD3 strain that displays the desired properties of a microbial heterologous chassis, including rapid growth and sporulation on widely employed solid and liquid culture media. The strain displays susceptibility to antibiotics commonly employed in molecular biology laboratories, rendering it amenable to genetic manipulation. *S. sungeiensis* SD3 can be readily transformed with genetic material, including large BGC-containing artificial chromosomes, using readily available *E. coli-Streptomyces* shuttling vectors. The genetic tractability of *S. sungeiensis* SD3 makes it well suited for optimization and progression through the Design-Build-Test-Learn (DBTL) cycle. One notable feature of *S. sungeiensis* SD3 is its unique placement within a distinct clade in the phylogenetic tree. It stands apart from all other Streptomyces chassis strains previously developed by other researchers. This distinctive phylogenetic placement positions *S. sungeiensis* SD3 as the preferred chassis for

the expression of BGCs derived from *Streptomyces* strains within the same or neighbouring clades.

2.2 Methods

General

NMR data were collected on an Avance NEO 400 MHz or Avance III 600 MHz spectrometer at 298 K. Chemical shifts are expressed in δ (ppm) and referenced to the residual solvent signals. ESI-MS spectra were acquired on a Thermo LTQ XL spectrometer with ESI source. Flash column chromatography (CC) was performed using silica gel (230−400 mesh, Merck, Darmstadt, Germany). HPLC seperation and purficaiton of the metabolites was conducted using an Agilent 1200 series HPLC-DAD system with an ODS column (Pursuit XRs: diphenyl, 250 mm × 10 mm, 5 μm; Cosmosil: cholester, 250 mm × 10 mm, 5 μm) and a Shimadzu liquid chromatography system equipped with an ODS column (ACE: C18-HL, 250 mm × 10 mm, 5 μm). UV-Vis spectra were collected on a Denovix spectrophotometer. All reagents and materials were purchased from Sigma-Aldrich unless otherwise indicated.

Bacterial strains and culture conditions

The *E. coli* strains Top10, DH10B (Thermo Fisher Scientific) and ET12567 were cultured at 37 °C in LB liquid medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.2). Selective antibiotics were added when necessary and were used at 50 µg/mL apramycin; 25 µg/mL chloramphenicol; 50 µg/mL kanamycin; 25 µg/mL nalidixic acid (Sigma-Aldrich). Routine culturing of *Streptomyces* strains on solid media was carried out on mannitol soy (MS) agar (20 g/L soya flour, 20 g/L mannitol and 20 g/L BactoTM agar), while culturing in liquid media was done in

tryptic soy broth (TSB) (30 g/L TSB powder) (Merck Millipore, Germany). Incubation of *Streptomyces* strains on solid media was done at 30c incubators whilst for liquid culture 30c shaking incubators at 180 rpm was used.

Media supplementation used in this study were described: Trace element solution (per liter: 40 mg $ZnCl_2$, 200 mg $FeCl_3$•$6H_2O$, 10 mg $CuCl_2$•$2H_2O$, 10 mg $MnCl_2$•$4H_2O$, 10 mg $Na_2B_4O_7$•$10H_2O$, 10 mg $(NH_4)6Mo_7O_{24}$•$4H_2O$), and Trace salt solution (per liter: 1g $FeSO_4$•$7H_2O$, 1g $MnCl_2$•$4H_2O$, 1g $ZnSO_4$•$7H_2O$), $FeSO_4$ solution (0.01g $FeSO_4$ in 1ml H_20), ZnSO4 solution (0.01g ZnSO4 in 1ml H_20). Media used in this study for fermentation were as described: GYMose media (per liter: 20g Glucose, 10g Maltose, 5g Yeast Extract), Oatmeal media (per liter: 25g Oatmeal Quaker), Media A (per liter: 5g Mannitol, 5g soyabean flour, 10ml Glycerol (86%), 9g yeast extract, 1 mL trace element solution), Media B (per liter: 10g Glucose, 1g asparagine, 1g K_2HPO_4, 10mg $MnCl_2$•$4H_2O$, 1ml $ZnSO_4$ solution, 1ml $FeSO_4$ solution), Media C (per liter: 80g Sucrose, 10g soybean powder, 2g yeast extract, 1g meat extract (beef), 0.3g K_2HPO_4, 0.3g $MgSO_4$•$7H_2O$, 5ml $FeSO_4$ solution, 1g $CaCO_3$), Media D (per liter: 20g Soluble starch, 1g KNO_3, 0.5g NaCl, 0.5g $MgSO_4$·$7H_2O$, 0.5g K_2HPO_4, 1ml $FeSO_4$ solution), Media E (per liter: 5g Glucose, 5g corn steep powder, 10g oatmeal, 10ml oil, 1g K_2HPO_4, 1g $MgSO_4$•$7H_2O$, 1 ml trace element solution), GYM media (per liter: 10g Malt extract, 10g agar, 4g yeast extract, 4g glucose, 2g $CaCO_3$), ISP4 media (per liter: 10g soluble potato starch, 1g $MgSO_4$•$7H_20$, 1g NaCl, 2g $(NH_4)_2SO_4$, 2g $CaCO_3$, 20g agar, 1ml Trace salts solution), ISP5 media (per liter: 1g L-asparagine ,10ml Glycerol, 1g K_2HPO_4, 1ml Trace salts solution, 20g Agar), ISP6 media (per liter:

5g proteose-peptone, 0.5g Ammonium iron (II) citate, 1g K_2HPO_4, 15g agar, 1g yeast extract), ISP7 media (per liter: 15g Glycerol, 0.5g L-tyrosine, 1g L-asparagine, 0.5g K_2HPO_4, 0.5g $MgSO_4 \cdot 7H_20$, 0.5g NaCl, 1ml $FeSO_4$ solution, 20g agar).

Transformation of *S. sungeiensis* SD3 with BGC-containing PACs

pESAC13A containing BGCs from *S. spp.* P46 was obtained from BioS&T. Three sets of primers were used to confirm the correct insertion of BGCs into the pESAC13A vectors, corresponding to the first, last and middle genes of each cluster delineated by antiSMASH prediction. To integrate the BGC containing pESAC13A vectors into *S. sungeiensis* SD3, triparental conjugation was first carried out utilizing donor DH10B containing pESAC13A::BGC, DH10B/pR9604 helper, and ET12567 *E. coli* recipient strains. ET12567 cells harboring the correct pESAC13A::BGC was then conjugated with *S. sungeiensis* SD3. For conjugation of vectors into *S. sungeiensis* SD3, spores were germinated by heat-shock at 45 ºC for 10 minutes and mixed with E. coli ET12567/pUZ8002 harboring the correct vector at a ratio of 1:5. The mixture was spread on MS agar supplemented with 20 mM $CaCl_2$ and 20 mM $MgCl_2$ and incubated for 20 hours at 30oC. Following which, the plates were overlaid with 50 µg/mL apramycin and 25 µg/mL nalidixic acid.

Standard DNA manipulations

Standard PCR reactions were carried out using Q5 High Fidelity Polymerase (NEB) in accordance with the manufacturer's instructions. DNA was purified from 1% agarose gels with SYBR safe (ThermoFisher Scientific) staining. The molecular weight standard used for agarose gel electrophoresis was 1 kb Plus DNA Ladder (ThermoFisher Scientific). The purification of plasmids was carried out using Qiagen Spin Miniprep, and gel extractions using Zymo Research Zymoclean Gel DNA Recovery Kit, according to the manufacturer's instructions. Restriction endonucleases and DNA-modifying enzymes were from New England Biolabs (NEB). Molecular cloning was carried out using restriction cloning or Gibson Assembly (NEB) to transfer PCR amplified fragments to the vector of choice.

Whole genome sequencing

S. sungeiensis SD3 was isolated from a sediment sample collected from Sungei Buloh mangrove wetland reserve, Singapore. The complete genome of *S. sungeiensis* SD3 was obtained as a linear chromosome of 10,292,314 base pairs using Single Molecule Real Time (SMRT) (Pacific Biosciences, California, USA) sequencing and Illumina sequencing and CLC Genomics Workbench (CLC bio, Denmark).

Phylogenetic and Comparative Genomics Analysis

The phylogenetic relationship of *S. sungeiensis* SD3 with other *Streptomyces* species was established via whole genome comparisons using the TYGS server (https://tygs.dsmz.de) and visualised by the online software iTOL

(https://itol.embl.de/). The ANI calculator from EZBioCloud was used to ascertain the OrthoANIu value between *S. sungeiensis* SD3 and *S. chartreusis*. The open reading frame prediction and genome annotation was carried out using RAST (Rapid Annotation using Subsystem Technology). The predicted proteins obtained by RAST was used in OrthoVenn to search for orthologous protein clusters. Secondary metabolite biosynthetic gene clusters were identified using the antiSMASH database.

Fermentation and metabolite profiling

Spores of *S. sungeiensis* SD3 wild-type or the knockout strain were plated on mannitol-soy agar (2% mannitol, 2% pollen soy powder, 2% agar) and incubated for 4 days at 30ºC. Starter cultures were prepared by inoculating the mycelia to 50 mL GYM broth (0.4% glucose, 0.4% yeast extract, 1% malt extract) in 250 mL conical flask and shaken for 5-7 days at 170 rpm until dispersed. 100 μL starter cultures were used for spreading on PM3 agar (2% oatmeal, 0.25% glycerol, 0.000004% $ZnCl_2$, 0.00002% $FeCl_3 \cdot 6H_2O$, 0.000001% $CuCl_2 \cdot 2H_2O$, 0.000001% $MnCl_2 \cdot 4H_2O$, 0.000001% $Na_2B_4O_7 \cdot 10H_2O$, 0.000001% $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 1% agarose) and incubated at 30 ºC. After two weeks fermentation, the resulting agar plates were mashed and extracted with MeOH. The filtered MeOH extract was concentrated under vacuum and re-dissolved in 500 μL MeOH to obtain the crude sample. HPLC analysis was performed on Agilent1200 HPLC-DAD system equipped with an ODS column (Pursuit XRs: 250 mm × 4.6 mm, 5 μm) using a linear gradient of CH_3CN in H_2O with 0.1% formic acid (0–5 min, 10%–20%

CH_3CN; 5–35 min, 20%–70% CH_3CN; 35–50 min, 70%–90% CH_3CN; 50–60 min, 90%–100% CH_3CN; 60–70 min, 100% CH_3CN) at a flow rate of 1.0 mL/min.

Isolation and characterization of endogenously produced secondary metabolites

The liquid PM3 medium with better production titer was chosen for the large-scale fermentation of *S. sungeiensis* SD3 to obtain the main secondary metabolites. Briefly, eight 2.0 L baffled flasks, each containing 800 mL production medium, were inoculated with 40 mL seed culture and incubated for 4 days at 30 °C, 180 rpm. The resulting cultures were centrifuged and separated into supernatant and cell pellet, followed by extraction with ethyl acetate (EA, 1:1, v/v) and MeOH (3 × 1 L), respectively. Extracts were combined and evaporated under vacuum to yield the total crude residue, which was then subjected to the flash column separation (CC) over silica gel using a stepwise gradient of hexane/EA/MeOH (20\1\0, 10\1\0, 5\1\0, 1\1\0, 0\1\0, 0\20\1, 0\10\1, 0\3\1, 0\0\1) to yield nine fractions (Fr.1-Fr.9). Further purification was guided by HPLC-UV analysis, which showed the products were mainly in Fr.4 and Fr.5. Fr.4 was purified by repeated semi-preparative RP-HPLC with a high-load ACE ODS column [CH_3CN-H_2O (containing 0.1% formic acid, v/v) 30:70, v/v; flow rate, 4.7 mL/min] to furnish streptazone B1 (**3**, 2.0 mg, t_R = 12.5 min) and obscurolide A2 (**1**, 7.2 mg, t_R = 16.5 min). Fr.5 that contains chartreusin peaks was applied to the same RP-HPLC system [ACE ODS column, CH_3CN-H_2O (containing 0.1% formic acid, v/v) 55:45, v/v; flow rate, 4.7 mL/min], and the collected chartreusin-containing fraction was further purified by a Pursuit XRs Diphenyl column [CH_3CN-H_2O (containing 0.1% formic acid, v/v) 60:40, v/v;

flow rate, 3.0 mL/min] to afford pure chartreusin (**2**, 8.5 mg, t_R = 15.2 min). As for the two other metabolites produced by a PAC-transformed *S. sungeiensis* SD3, the fermentation of SD3 mutant and the broth extraction and isolation procedures were conducted with the same protocol as aforementioned, and the two NMR-pure samples (adipostations A (**4**) and B (**5**), t_R = 14.5 min and 17.0 min) were finally achieved by preparative RP-HPLC with a Cosmosil Cholester column [CH_3CN-H_2O (containing 0.1% formic acid, v/v) 88:12, v/v; flow rate, 3.0 mL/min].

By comparison of the observed spectroscopic data and physicochemical properties with those reported in the literatures, the above isolated compounds were identified to be obsurolide A2 (**1**), chartreusin (**2**), streptazone B1 (**3**), adipostatins A (**4**) and B (**5**), respectively. The MS, ^1H and ^{13}C NMR spectra can be found in the Supporting Information.

RNA isolation and sequencing

S. sungeiensis SD3 was grown in TSB media for either 12hrs or 96hrs in baffled flasks. Cells were collected and homogenized using a pestle. 1mg/ml working concentration of lysozyme was used for incubation at 37 ºC for half an hour. Cells were freeze-thawed using liquid nitrogen for 5-10 cycles. TRIzol Reagent was added according to the manufacturer's instructions (Invitrogen). Finally, total RNA was extracted and purified using the Purelink RNA Mini kit according to manufacturer's instructions (Thermo Fisher Scientific). DNAse-I (Thermo Fisher Scientific) was added in accordance with the Purelink RNA Mini kit protocol.

The resulting RNA samples were then sent to Novogene Co., Ltd. for Illumina paired-end 150bp sequencing. Agilent 2100 Bioanalyzer was used to measure the concentration and quality of total RNA, and samples with integrity values >7.0 was used for library construction.

Trimmomatic was used to carry out read trimming to remove low quality bases and adapter sequences. HISAT2 v 2.2.1 was used for mapping the RNA-seq to the *S. sungeiensis* SD3 genome. FeatureCounts was used to quantify reads based on the "CDS" feature characteristics of the annotation file generated by RAST for *S. sungeiensis* SD3. DESeq2 was used for the normalization of reads and detection of differentially expressed genes. Integrative genomics viewer (IGV) v 2.16.2 was used for simultaneously visualizing RNA sequencing data and genomic annotations.

Measurement of promoter strength

Streptomyces spores were inoculated into 50ml TSB liquid media starter cultures and shaken for two days, before subculturing into 50ml TSB liquid media for another three days with the addition of 30 5-mm glass beads. 10ml of culture was extracted and washed and resuspended with 1x PBS. 100ul of cells from 10ml of 1x PBS cell suspension was used for measurement of OD_{600} and normalization of cell concentration to 0.5 OD_{600}. Fluorescence measurements were carried out using a Tecan Infinite® 200 PRO Plate Reader. For EGFP, the excitation and emission wavelengths used were 485 nm and 515 nm respectively. Readings were done in triplicates for each sample.

CRISPR/Cas9-assisted gene deletion

pCRISPR-Cas9 from Tong et al. and pQS-idgS from Wang et al. were utilized in the genome modification efforts, with the deletion strain successfully obtained from the use of pQS-idgS. The construction of pQS-idgS containing the sgRNA and homology arms were as described by Wang et al. In brief, the short oligonucleotide corresponding to the sgRNA was cloned into the pQS-idgS vector after *Nco*I and *Xba*I digestion. 2kb homologous recombination templates were introduced to this vector via Gibson assembly, upon linearization by *Stu*I. The Gibson mix was then transformed into chemically competent TOP10 *E. coli*. Correctly assembled vectors were then selected using colony PCR and DNA sequencing and subsequently transformed into E. coli ET12567/pUZ8002 for conjugation with *S. sungeiensis* SD3 using conjugation conditions as described above. After 5 days of incubation at 30oC, exconjugants were transferred to MS agar supplemented with 5 μg/mL thiostrepton for activation of the *tipA* promoter controlling the *Streptomyces* codon-optimized nuclease gene *scas9*. Curing of plasmid was done by culturing in TSB liquid media and subsequent plating onto MS agar plates, both without selective antibiotics. Correct mutants were then screened with PCR of deletion regions.

Isolation and characterization of secondary metabolites produced by heterologous pathways

For heterologous expression of Tsk BGC in SD3: As tasikamides generally feature with similar characteristic UV absorptions and we have tasikamides as authentic samples in hand, the reality of tasikamide production by the Tsk BGC-

transformed strain SD3::BGC23 was readily confirmed though the HPLC-UV and MS analysis of the organic extract of the culture broth.

For heterologous expression of furaquinocin BGC in SD3: A noticeable secondary metabolite was observed when P46_BGC28 was heterologously expressed within SD3 host, and a routine HPLC-guided isolation and purification for the extract of 4L culture broth of SD3::P46_BGC23 leads to the collection of the target compound with bright yellow colour. Based on the fact that BGC28 is predicted to produce a furaquinocin-type product coupled with the observed characteristic UV absorption for the isolated compound (**6**), this yellow metabolite was deduced to be a furaquinocin analogue. The detail of MS and NMR-based structural characterization is described below: the molecular formula of compound **6** was established as $C_{21}H_{24}O_6$ on the basis of HRESIMS data (*m/z*, 373.1631 $[M+H]^+$, calcd for $C_{21}H_{25}O_6$ 373.1645), which requires 10 degrees of unsaturation. The ^{13}C NMR data of **6** displayed resonances for two carbonyls at δ 182.6 and 179.9 ppm and ten olefinic/aromatic carbons between δ 107.9 and δ 162.1 ppm. The ^1H NMR spectrum accordingly showed resonances attributable to two aromatic proton (δ 7.01, 6.05), one methoxy (δ 3.77), four methyl groups (δ 1.35, 1.47, 1.49, 1.62), three methines (δ 5.15, 4.51, 3.67), and one methylene (δ 2.12/2.17) (Figure S10-1). The above NMR data (Table S6) were highly identical to those of furaquinocin D, with the only difference being that the Me-8 substitution in furaquinocin D was replaced by a proton in compound **6** (Figure S3). This realization was further confirmed by an extensive 2D NMR analysis (Figure S10), especially for the key HMBC correlations from H-8 (δ 6.05) to C-6 (δ 179.9)/C-9 (δ 182.6)/9a (δ 107.9).

Given the structure of compound **6**, we determined it is a new member of the furaquinocins family and named as furaquinocin M (**6**, Figure S3).

For heterologous expression of chlorostreptazone B1 (**11**) in SD3: P46_BGC13 encodes a marinopyrrole pathway that contains six putative halogenase genes. Although the *S. sungeiensis* SD3 host transformed with P46_BGC13 did not produce the anticipated marinopyrroles, the strain produced a new compound **11** that features a characteristic UV absorption similar to those of streptazones (Puder et al., 2000). We isolated and characterized the compound to establish its identify as a hitherto unreported compound. The molecular formula of compound **11** was shown to be $C_{10}H_{10}ClNO$ by HREIMS analysis. Its ^{1}H NMR spectrum (Figure S11-1) displayed one exchangeable proton (δ 8.05), two olefinic protons (δ 6.36 and 6.29), four aliphatic protons and one methyl protons (δ 2.21). The ^{13}C NMR spectrum (Figure S11-1) of **11** indicated the presence of one C-methyl, two aliphatic methylenes, one hetero-substituted and two olefinic methines, one carbonyl, and three other quaternary carbons. On the basis of HMBC correlations (Figure S11-4), biosynthetic considerations, and comparison of its ^{1}H and ^{13}C NMR data (Table S7) with structurally related known streptazone B1 (Puder et al., 2000). Compound **11** was assigned to be a chlorinated derivative of streptazone B1 at carbon C-7 and named as chlorostreptazone B1 (**11**, Figure S3).

2.3 Results

2.3.1 Selection of *S. sungeiensis* SD3 for host development

The use of fast-growing microbial chassis is crucial for efficient genetic modification and culturing work, particularly in the case of *Streptomyces* strains, which tend to grow slower than microbial chassis like *E. coli*. To identify potential chassis strains, we screened a collection of actinomycete strains isolated from the soil and aquatic environments of Singapore. The strains were evaluated based on their growth rate and sporulation on common media, including MS and GYM solid agar media, and TSB liquid media. After filtering out slow-growing or non-sporulating strains, twelve strains were identified with growth rates comparable to or higher than the model *Streptomyces* strains *S. coelicolor* M1154 and *S. lividans* TK24 (Figure S1). The twelve strains exhibited readily observable sporulation on MS solid agar within 3-4 days, which is essential for long-term storage and DNA transformation via *E. coli-Streptomyces* conjugation.

For a microbial strain to serve as a reliable heterologous chassis in the production of secondary metabolites, it is imperative for the strain to be amenable to genetic manipulation and proficient in the uptake of exogenous DNA vectors. This ability to accept foreign genetic material is not only essential for introducing BGCs into the chassis but also crucial in strain improvement through targeted gene knock-ins and knock-outs. Considering that most wild-type or undomesticated *Streptomyces* spp. are inherently genetically recalcitrant and unresponsive to conventional transformation methods, genetic tractability stands out as a pivotal criterion when choosing a suitable heterologous chassis. As many DNA vectors

employed for genetic manipulation of *Streptomyces* employed today harbour Apr^r, Cm^r, or Kan^r markers, the susceptibility to the three antibiotics is an important requirement of a chassis. We assessed the antibiotic sensitivity of the twelve strains and found that ten of the twelve strains are susceptible to apramycin (50 μg/mL), chloramphenicol (25 μg/mL), and kanamycin (50 μg/mL) on solid agar media (Figure S2). We next evaluated the capability of the strains in accepting DNA vectors carrying large foreign DNA inserts. We tested 14 pESAC13A-derived vectors derived from P1-derived artificial chromosomes (PACs) obtained from BioS&T (https://www.biost.com/pesac_13-libraries). These 14 PACs, which contain long DNA fragments in the range of 80 to 160 Kb, originated from the two genomic libraries from the genomic DNA of *Streptomyces tasikensis* P46 and *Streptomyces* sp. SD50. By introducing the 14 PACs into the strains through *E. coli-Streptomyces* conjugation, we found that only four of the ten strains displayed natural competence in accepting the PACs. Notably, *S. sungeiensis* SD3 exhibited the highest rate of successful conjugation with BGC-containing PACs, with 13 of the 14 PACs resulting in viable exconjugants as confirmed by positive PCR results (data not shown). Based on the number of exconjugants, the transformation efficiency of *S. sungeiensis* is comparable to the widely used hosts *S. coelicolor* M1154 and *S. lividans* TK24.

Subsequently, we subjected *S. sungeiensis* SD3 to further transformation experiments using a variety of DNA vectors, including some common pIJ101 and pSET152-derived conjugative and integrative plasmids (e.g., pIJ12551, pIJ8630, pSEVA28c1) used by academic researchers. These plasmids rely on the phiC31

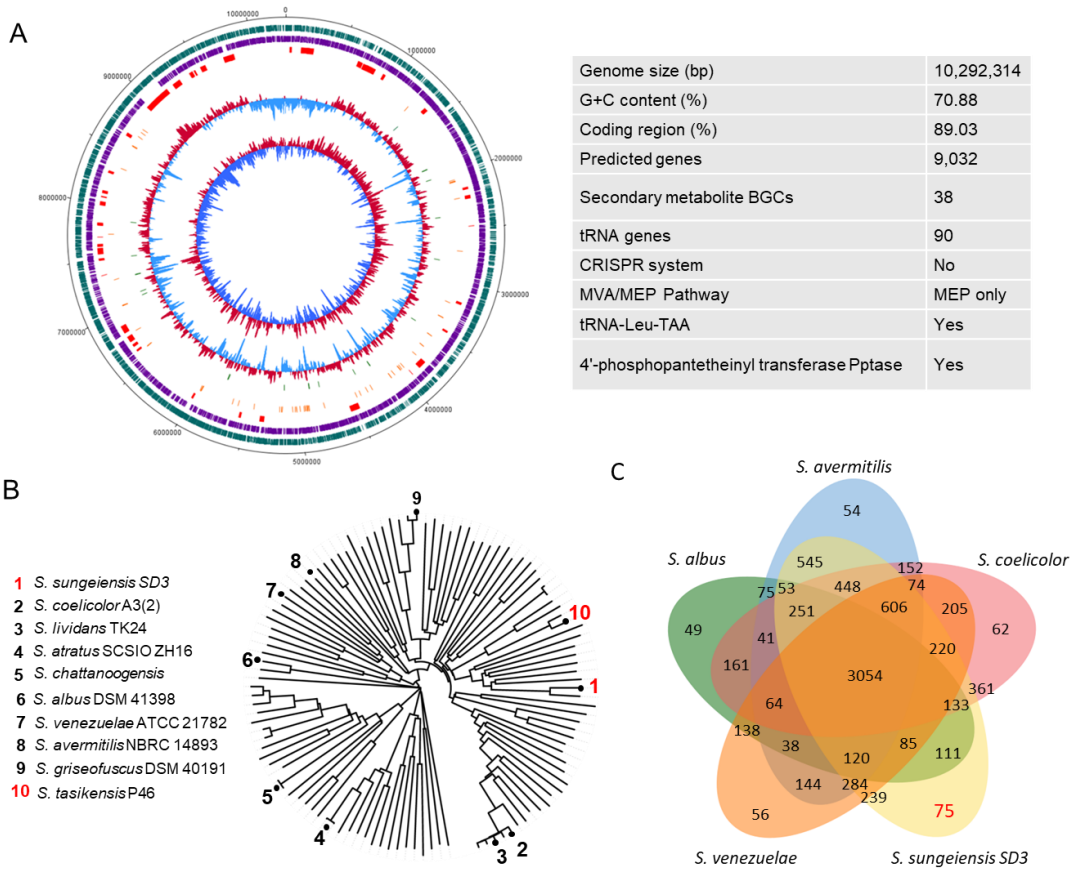
integration system to target the conserved phiC31-attB site present in most actinomycetes. All the pIJ101 and pSET152-derived plasmids tested by us were successfully conjugated into *S. sungeiensis* SD3 and their integration into the chromosome was confirmed by PCR (data not shown). The observations suggest that *S. sungeiensis* SD3 can be readily transformed with the readily available conjugative and/or integrative plasmids, which are often employed for the heterologous expression of individual biosynthetic genes or entire BGCs.

S. sungeiensis SD3 was initially isolated from a sediment sample obtained from the mangrove swamp of Sungei Buloh Wetland Reserve of Singapore. We confirmed that *S. sungeiensis* SD3 displayed compatibility with International *Streptomyces* Project (ISP) media, specifically ISP2-7. No additional media requirements were needed to induce the sporulation of *S. sungeiensis* SD3, as it readily sporulated on MS media within a short period of 3-4 days. Based on these promising outcomes, we decided to further characterize and optimize *S. sungeiensis* SD3 as a heterologous chassis for the expression of exogenous biosynthetic pathways.

2.3.2 Phylogenetic relationship and genomic features of *S. sungeiensis* SD3

To gain a better understanding of the primary and secondary metabolic capability of the strain, we sequenced and assembled the whole genome of *S. sungeiensis* SD3 using a hybrid approach combining Illumina and PacBio sequencing data. The complete genome had a total length of 10,292,314 bp and a guanine-cytosine content (GC) of 70.88% (Figure 2A). To determine its closest

phylogenetic neighbours, the whole genome sequence of *S. sungeiensis* SD3 was used to construct a Genome BLAST Distance Phylogeny (GBDP) tree using the Type Strain Genome Software (TYGS) (Meier-Kolthoff and Göker, 2019).



Genome size (bp)	10,292,314
G+C content (%)	70.88
Coding region (%)	89.03
Predicted genes	9,032
Secondary metabolite BGCs	38
tRNA genes	90
CRISPR system	No
MVA/MEP Pathway	MEP only
tRNA-Leu-TAA	Yes
4'-phosphopantetheinyl transferase Pptase	Yes

Figure 2. Phylogenetic and genomic characterization of *S. sungeiensis* SD3. (A) Chromosomal genome map and genomic features of *S. sungeiensis* SD3. Tracks from outermost circle to the center: (i) Marker for genome size. (ii) Predicted coding DNA sequence (CDS) on forward and reverse stands in green and purple respectively. (iii) Biosynthetic gene clusters predicted by antiSMASH 7.0.0 in red. (iv) tRNA genes in orange. (v) rRNA genes in green. (vi) GC content, the outward portion in red indicating higher than genome-wide average GC content, whilst the inward portion in blue indicating

lower than average GC content. (vii) GC skew, the outward the outward portion in red indicating higher than average GC skew, and the inward portion in dark blue indicating below average GC skew. The genome figure was constructed using DNAPlotter v 18.1.0. (Carver et al., 2009). (B) Phylogenetic relationship between *S. sungeiensis* SD3 and commonly used *Streptomyces* chassis. The phylogenetic tree was constructed using TYGS (Meier-Kolthoff and Göker, 2019) to calculate a Genome BLAST Distance Phylogeny (GBDP) and visualized using Iterative Tree Of Life iTOL (Letunic and Bork, 2021). The most commonly used *Streptomyces* chassis are labelled. (C) Venn diagram of unique and shared orthologous gene clusters amongst *S. sungeiensis* SD3, *S. avermitilis* NBRC 14893 (NC_003155), *S. coelicolor* A3(2) (CP042324), *S. venezuelae* ATCC 21782 (CP029193), and *S. albidoflavus* DSM 41398 (CP010519).

Notably, *S. sungeiensis* SD3 seems to belong to a distinct clade separate from other commonly used *Streptomyces* chassis strains that include *S. albidoflavus* DSM 41398, *S. coelicolor* A3(2), *S. lividans* TK24, *S. avermitilis* NBRC 14893, *S. venezuelae* ATCC 21782, *S. griseofuscus* DSM 40191, *S. chattanoogensis*, and *S. atratus* SCSIO ZH16 (Figure 2B). *S. sungeiensis* SD3 forms a monophyletic lineage with *S. chartreusis* and other closely related strains. The OrthoANIu value obtained from the EzBioCloud database between *S. sungeiensis* SD3 and *S. chartreusis* is 84.03%, which meets the ANI threshold value for establishing a novel species (95%) (Hitch et al., 2021; Yoon et al., 2017). Considering that the success rate of heterologous production of secondary metabolites usually increases with the phylogenetic proximity of the original producer and the heterologous host (Myronovskyi and Luzhetskyy, 2019; Nepal and Wang, 2019), we reasoned that *S.*

sungeiensis SD3 could be utilized as a heterologous chassis to complement the existing *Streptomyces* chassis.

Genome annotation of *S. sungeiensis* SD3 using RAST resulted in the identification of 9,154 protein-coding ORFs and 90 tRNA genes (Aziz et al., 2008). A total of 1,589 genes were categorized into subsystems by RAST, with the highest relative abundance observed in genes involved in amino acid metabolism (29.8%) and carbohydrate metabolism (26.8%). To analyze genomic content differences, OrthoVenn2 was employed to compare the annotated proteins of *S. sungeiensis* SD3 with those of *S. coelicolor*, *S. avermitilis*, *S. albidoflavus*, and *S. venezuelae* (Xu et al., 2019). A total of 6,754 orthologous putative proteins were identified in *S. sungeiensis* SD3, with 3,054 shared among the other four strains, representing the core proteins shared by the *Streptomyces* strains. *S. sungeiensis* SD3 shared the highest number of orthologous protein clusters with *S. avermitilis* (545), followed by *S. coelicolor* (361), *S. venezuelae* (239), and *S. albidoflavus* (111), reflecting their phylogenetic relatedness (Figure 2C). 75 protein clusters are unique to *S. sungeiensis* SD3 and are not found in the other four strains, while only 62, 56, 54, and 49 clusters are unique to *S. coelicolor*, *S. venezuelae*, *S. avermitilis*, and *S. albidoflavus*, respectively. The 75 unique protein clusters are primarily involved in various metabolic processes, including amino acid and carbohydrate metabolism, with the largest gene ontology group associated with DNA binding and regulation (14.28%). *S. sungeiensis* SD3 possesses more enzymes involved in the metabolism of aromatic compounds such as benzoate (75) compared to *S. coelicolor* (42) and *S. avermitilis* (38). Another unique group of proteins found in *S. sungeiensis* SD3

are involved in the biosynthesis of pyrroloquinoline quinone (8%), which is an organic cofactor essential for bacterial dehydrogenases, including methanol dehydrogenases (Matsutani and Yakushi, 2018).

2.3.3 The biosynthetic capability of *S. sungeiensis* SD3

To evaluate the capacity of *S. sungeiensis* SD3 to produce secondary metabolites, we mined its genome for BGCs using antiSMASH 7.0 (Blin et al., 2023). The mining yielded at least 37 BGCs, including nine coding for Type I, II, or III polyketide synthases (PKS), and seven coding for non-ribosomal peptide synthetases (NRPS) (refer to Table S5). Among the 37 BGCs, several BGCs bear resemblance to characterized BGCs from other strains in terms of gene composition and sequence homology. These putative homologous BGCs are anticipated to produce secondary metabolites that either match or share structural similarities with coelichelin, ectoine, ishigamide, albaflavenone, focixins, chartreusin, lagunapyrone A, desferrioxamine, streptazone E, and isocomplestatin (Figure 3A and Table S5). Interestingly, although BGC24 shares similarity with the BGC for streptazone E biosynthesis in *Streptomyces* sp. MSC090213JE08, the modular PKS encoded by the BGC only contain five modules, in contrast to the six modules encoded by the streptazone E BGC (Ohno et al., 2015), indicating a difference in carbon-chain length between the products of the two BGCs. Meanwhile, several BGCs are predicted to synthesize unknown secondary metabolites with novel polyketide or non-ribosomal peptide backbone. The number and diversity of BGCs found in the

genome suggests *S. sungeiensis* SD3 possesses the metabolic networks and pathways to support the synthesis of diverse secondary metabolites.

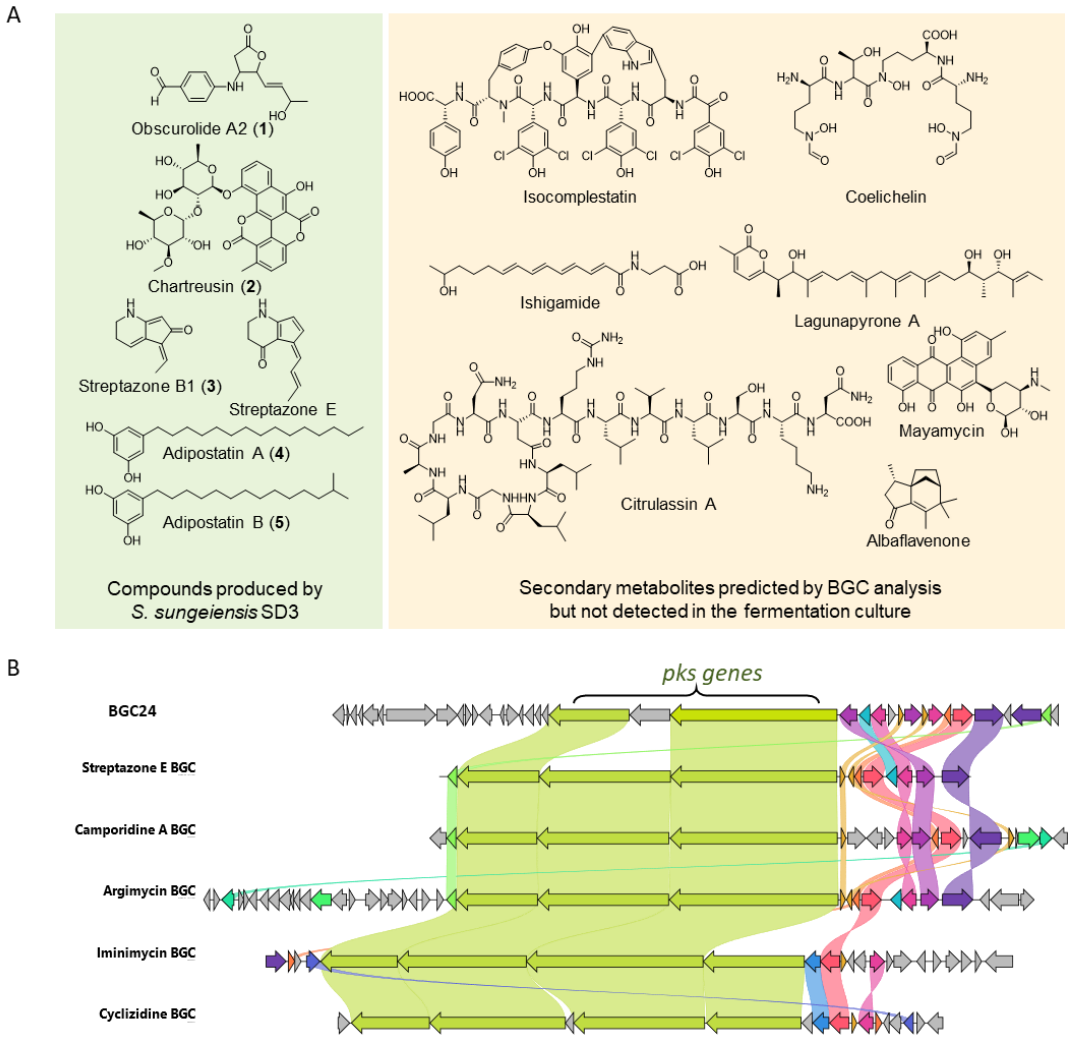


Figure 3. Production of secondary metabolites by *S. sungeiensis* SD3 under laboratory fermentation conditions. (A) Secondary metabolites isolated from the fermentation culture of *S. sungeiensis* SD3 or predicted by AntiSMASH analysis but not detected likely due to repressed gene expression. (B) Comparison between the streptazone B1-synthesizing BGC24 from *S. sungeiensis* SD3 and the streptazone E-synthesizing *stz* BGC and other similar BGCs that produce polyketide-derived piperidine alkaloids.

We cultured *S. sungeiensis* SD3 in 27 different culture media, including 8 liquid and 19 solid media to substantiate the biosynthetic potential and identify the constitutively active BGCs. Following the fermentation, the organic metabolites were extracted from the culture broth and mycelium. The metabolites were subsequently profiled using liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS). Aided by an in-house natural product library and characterization by NMR spectroscopy, we identified two major secondary metabolites produced by *S. sungeiensis* SD3 that include obscurolide A2 (**1**) ($[M + H]^+$, *m/z* 276.08) and chartreusin ($[M + Na]^+$, *m/z* 663.25) (**2**) (Figures 3A, S5, and S6) (Hoff et al., 1992). Obscurolide is a putative butyrolactone derivative that was produced in abundance under several fermentation conditions. The other major product chartreusin is a glycosylated aromatic polyketide known for its dsDNA-binding and cleaving activity (Xu et al., 2005). The production of chartreusin agrees with the presence of the Type II PKS-encoding BGC33 that is highly similar to the chartreusin BGC found in *S. chartreusis* HKI-249.

S. sungeiensis SD3 also produced streptazone B1 ($[M + H]^+$, *m/z* 162.08) (**3**) under several fermentation conditions. Streptazone B1 is a piperidine alkaloid that features a densely packed functionalised scaffold and unknown biosynthetic pathway (Puder et al., 2000) (Figures 3A, S7). As noted earlier, *S. sungeiensis* SD3 contains a Type-I modular PKS-encoding BGC (BGC24) that shares homology with the BGCs that produces streptazone E and other polyketide-derived piperidine alkaloids (Ohno et al., 2015; Puder et al., 2000; Wørmer and Poulsen, 2022). However, BGC24 encodes a five-module PKS system, in contrast to the six-module

PKS system for streptazone E biosynthesis (Figure 3B). With BGC24 assigned as the BGC for streptazone B1 biosynthesis, our studies yielded the surprising finding that streptazone B1 and streptazone E are produced by similar but different pathways.

S. sungeiensis SD3 produced two highly hydrophobic compounds that share the same mass ($[M + H]^+$, *m/z* 321.28) after the strain was transformed with a P1-derived artificial chromosome (PAC) that contains an exogenous DNA fragment. The two compounds were found to be adipostatins A (**4**) and B (**5**) by HRMS and NMR spectroscopy (Figures 3A, S8, and S9). Adipostatins are alkylresorcinols or Type III PKS-generated phenolic lipids with modest antifilarial activity against the nematode parasite *Brugia malayi* that causes elephantiasis (Rateb et al., 2015). Considering that the exogenous BGC does not contain any Type III PKS genes, the production of adipostatin A and B is likely due to the inadvertent activation of the native alkylresorcinols BGC37b.

The results from the AntiSMASH mining analysis underscore the potential of S. sungeiensis SD3 to synthesize a wide array of secondary metabolites, indicating that the genetic makeup of the chassis encompasses a robust metabolic network conducive to the heterologous production of secondary metabolites. Considering that only a small fraction of the predicted secondary metabolites has been identified by the metabolite profiling, many of the endogenous BGCs are likely to be silent under laboratory fermentation conditions. The discrepancy between the large number of secondary metabolites suggested by AntiSMASH analysis and the small

number of metabolites detected underscores the highly regulated production of secondary metabolites in *S. sungeiensis* SD3.

2.3.4 Gene expression profile of *S. sungeiensis* SD3

We performed transcriptomic profiling by RNA sequencing to gain a better understanding of the metabolic activity and gene expression pattern of *S. sungeiensis* SD3. Based on the timing of the production of chartreusin and streptazone B1, RNA sequencing was performed at two time points: pre-BGC expression (12 hours) and at the on-set of BGC expression (96 hours). Total RNA was extracted from biological triplicates at these two time points and subjected to RNA-seq analysis. Analysis of the RNA-Seq results revealed that at 96 hours, 22 out of 41 BGCs exhibited upregulation of expression for at least 50% for the core biosynthetic genes ($log2FC > 1$, $P < 0.05$, DESeq2) (Figure 4). The highest log2FC between 12 and 96 hours was observed for the chartreusin BGC (BGC33, 10.8-11.5 for PEG 8216-8217) and BGC37a (11.1-12.3 for PEG 8652-8655). Interestingly, comparison of the 96hr reads shows comparable transcriptional activity for BGC33 and the BGC37a that is predicted to produce a non-ribosome peptide that resembles isocomplestatin. However, the anticipated product of BGC37a could not be detected by HPLC and LCMS-assisted metabolite profiling, suggesting the BGC remains silent due to post-transcriptional regulation. While for the two cryptic BGC32a and 32b that are predicted to produce novel polyketide natural products, the low gene expression levels of suggest that they are silent BGCs because of repressed gene expression.

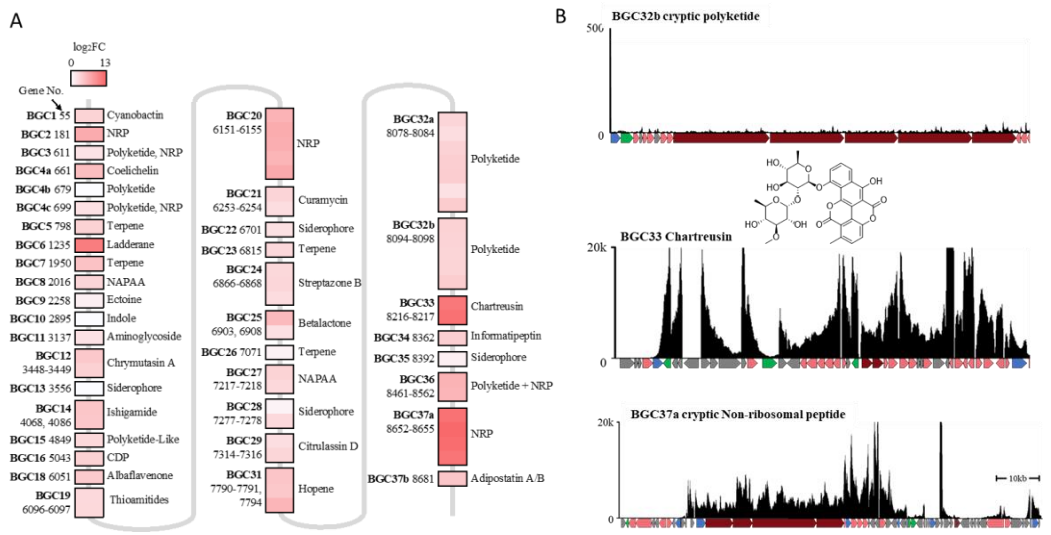


Figure 4. Change in BGC expression profile of *S. sungeiensis* SD3 between 12hrs and 96hrs. (A) Core biosynthetic genes of each BGC in SD3 are displayed with their protein encoding gene (PEG) number as annotated by RAST server. The log2FC values of these biosynthetic genes comparing 96hrs to 12hrs as analyzed and normalized by DESeq2 are used to generate a heat map with the higher intensity indicating a higher fold change value. (b) BGC32b, BGC33 and BGC37a read count profiles from 96hrs are mapped onto the *S. sungeiensis* SD3 genome using HISAT2 v 2.2.1 and visualized using Integrative genomics viewer (IGV) v 2.16.2 (Kim et al., 2019; Robinson et al., 2011). The biosynthetic genes of the BGCs as annotated by antiSMASH 7.0 are shown using arrows in the schematic (dark red for core biosynthetic genes, pink for supporting biosynthetic genes, green for regulatory genes, blue for transport related genes, and grey for others).

The main metabolic pathways that support the production of the aromatic polyketide chartreusin were further investigated to understand why the fermentation titer (~2.7 mg/L) for chartreusin was high. Glucose and maltose served as the primary carbon sources in the GYM liquid media used for chartreusin production

and RNA isolation. Analysis of the glycolysis pathway revealed downregulation of the majority of genes associated with the glycolysis pathway and the citric acid cycle (log2FC < 1, P < 0.05, DESeq2), indicating a slowdown of metabolism and growth at 96hr. In contrast, the majority of genes involved in the conversion of pyruvate to acetate (11 out of 19 genes) and the genes that convert acetate to acetyl-CoA and malonyl-CoA were upregulated. Moreover, 34 out of 67 genes from the beta-oxidation pathway produces acetyl-CoA also showed upregulation at 96 hr. In addition to acetyl-CoA and malonyl-CoA, monosaccharides that include D-fucose and D-digitalose are also required as the precursors for chartreusin biosynthesis (Xu et al., 2005). The majority of associated genes involved in D-glucose-1P to dTDP-L-rhamnose production were upregulated at 96 hours (Figure S4). The observations suggest that the high fermentation titer of chartreusin is likely to the synergized result of the high-level expression of chartreusin biosynthetic genes and the upregulation of the primary metabolic genes for acetyl-CoA and monosaccharide production.

2.3.5 Assessment of the strength of orthogonal promoters in *S. sungeiensis* SD3

For heterologous production of secondary metabolites, constitutively active promoters are often required to activate gene expression. To identify orthogonal promoters that are functional in *S. sungeiensis* SD3, we assessed the strength of several widely used promoters that were shown to function in multiple actinomycete strains. The chosen promoters ermEp*, sf14p, KasOp*, gapdhp(EL),

gapdhp(KR) were placed upstream of the engineered green fluorescent protein (EGFP) reporter gene in the pIJ8630 vector (Shao et al., 2013; Sun et al., 1999; Wang et al., 2013). The promoter-containing plasmids were conjugated into *S. sungeiensis* SD3 and the EGFP fluorescence intensity of the culture broth was used as a measure of the promoter strength.

By measuring the fluorescence intensity of the cells, we found that KasOp*, gapdhp(EL), gapdhp(KR) induced the highest EGFP expression, whereas SF14p induced moderate EGFP expression (Figure 5). The ermEp* promoter is one of the most commonly promoters used for driving gene expression in actinomycetes (Myronovskyi and Luzhetskyy, 2016). Surprisingly, ermEp* did not seem to induce significant gene expression in *S. sungeiensis* SD3, in contrast to the relatively greater strength reported for ermEp* in *S. coelicolor*, *S. avermitilis* NRRL8165 and other *Streptomyces* chassis (Liu et al., 2016; Luo et al., 2015; Wang et al., 2013).

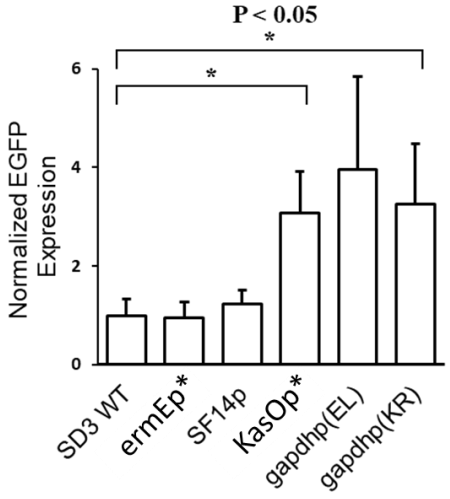


Figure 5. Comparison of the strength of constitutively active promoters in *S. sungeiensis* SD3. The normalized EGFP fluorescence intensities are shown for the different promoters

in *S. sungeiensis* SD3. The values are means with standard deviation (SD) from three independent experiments.

2.3.6 *S. sungeiensis* SD3 is amenable to CRISPR/Cas9-based genome editing

Strain optimization is often required to improve the performance of microbial heterologous hosts. Aside from the UV light or chemical induced random mutagenesis employed for the improvement of industrial strains, gene-targeted genome editing has become a common practice in strain improvement. For example, the removal of competing biosynthetic pathways by genetic approaches in *S. lividans*, S. *avermitilis*, and *S. albidoflavus* has resulted in cleaner metabolite background and improved fermentation yield (Ahmed et al., 2020; Hwang et al., 2023; Myronovskyi et al., 2018). Moreover, CRISPR/Cas9-assisted methods have emerged as efficient genome editing tools for actinomycetes in recent years (Tao et al., 2018).

We tested two Cas9-expressing vectors, including the pSG5-based pCRISPR-Cas9, and the pIJ101-based pQS-idgS (Tong et al., 2015; Wang et al., 2020), for gene deletion and insertion in *S. sungeiensis* SD3. The conjugation of pCRISPR-Cas9 into *S. sungeiensis* SD3 was not successful after repeated attempts, whereas the conjugation of pQS-idgS into *S. sungeiensis* SD3 yielded exconjugants. The observations suggest that the temperature sensitive pSG5 replicon in pCRISPR-Cas9 is incompatible with *S. sungeiensis* SD3. A similar observation was reported earlier for *Saccharopolyspora erythraea*, and the authors proposed that the pSG5 replicon can cause unpredicted gene recombination and instability (Mo et al., 2019).

Wang et al. also observed similar issues, which led to their use of the pIJ101 replicon for the construction of the pQS-idgS (Wang et al., 2020).

We continued to validate the utility of pQS-idgS for gene deletion and promoter insertion in *S. sungeiensis* SD3. We chose to target the first *pks* gene (ctg1_8216) of the silent cryptic BGC32b by deleting the first 1,323bp fragment of the *pks* gene (see Figure 6A). We also constructed the pQS-idgS-based plasmid to insert the KasOp* promoter in front of ctg1_8216 to test whether the constitutively active KasOp* can activate the expression of the BGC to produce the cryptic product (Figure 6B). We further constructed the plasmid to target the beta-ketoacyl synthase genes *chaA* and *chaB* within the chartreusin BGC to disrupt the production of chartreusin while inserting a PhiC31 attB attachment site and a *sfp* gene. Disrupting *chaA* and *chaB* would abolish chartreusin production, which will not only eliminate the chartreusin-associated HPLC peaks to facilitate the detection of metabolites produced by heterologous biosynthetic pathways but also free up malonyl-CoA and NADPH to boost fermentation titer. The additional attB site is expected to allow the integration of an additional copy of exogenous BGC in the chromosome; whereas the insertion of the *sfp* gene that codes for the promiscuous Sfp 4'- Phosphopantetheinyl transferase may lead to the activation of silent BGCs (Figure 6C) (Zhang et al., 2017). Upon the conjugation of the pQS-idgS-based plasmids that harbor the corresponding sgRNA and homology arms into *S. sungeiensis* SD3, exconjugants were readily observed. After the curing of the plasmid following the procedure described by Zhang and co-workers (Wang et al., 2020), we confirmed the successful deletion of the *pks* gene fragment and insertion

of the KasOp* promoter by PCR and DNA sequencing (Figure 6D). The deletion of *chaA* and *chaB* genes was successful as well, which was further confirmed by the observation that the mutant strain no longer produced chartreusin (Figure 6E).

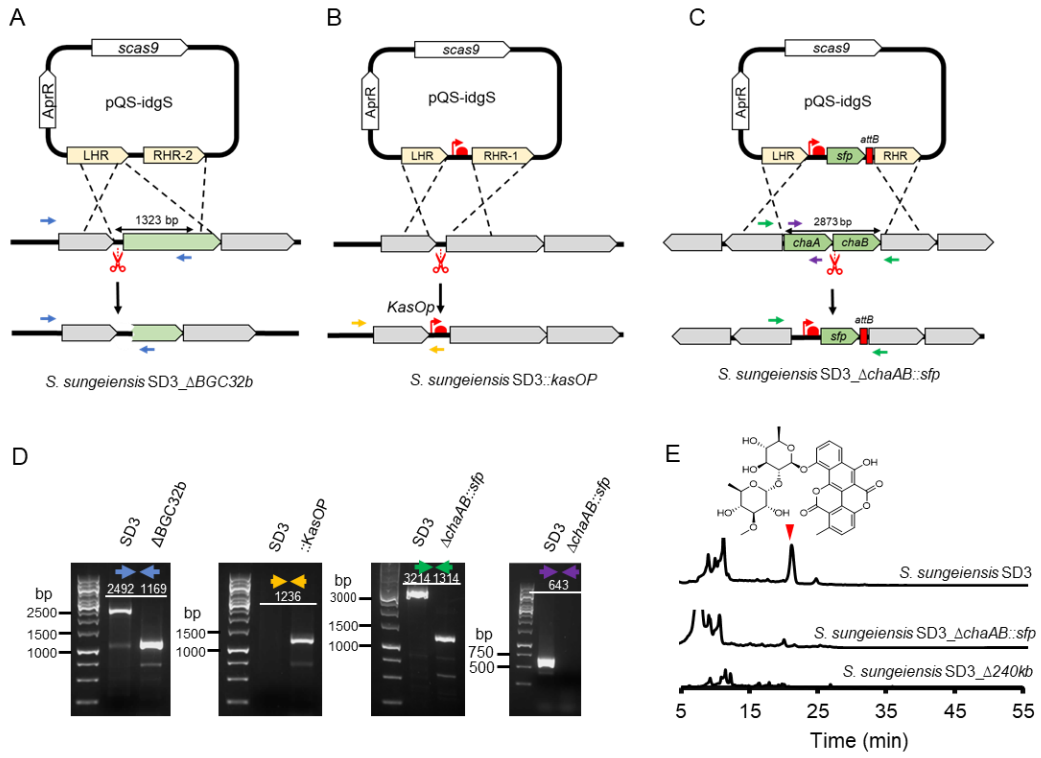


Figure 6. Gene deletion and insertion via CRISPR-Cas9 enabled genome editing in *S. sungeiensis*. (A-C) Experimental designs to insert the constitutive promoter KasOp* in front of the first T1PKS module of BGC32b, delete the first AT domain of BGC32b, and delete the *chaA* and *chaB* biosynthetic genes of BGC33 by using CRISPR-Cas9. (D) PCR results to confirm the promoter insertion or AT domain deletion in BGC32b or loss of *chaA* gene in the *S. sungeiensis* SD3_Δ*chaAB*::*sfp* strain. The primer regions are denoted by the coloured arrows. (E) HPLC analysis confirmed the mutant strains have lost the ability to produce chartreusin. The chartreusin peak is denoted by the red triangle.

While the intended gene deletion and promoter or gene insertions were successful, our whole genome sequencing of one of the mutant strains unearthed an unexpectedly extensive deletion for one of the *chaAB* deletion clones. The deletion resulted in the loss of a 240.6 kb DNA fragment, stretching 206,425 bp upstream of *chaA* and 30,843 bp downstream of *chaB*. As a result, the entirety of BGC33 (chartreusin BGC) was eradicated while half of the upstream BGC32a/b was also lost. It is worth noting that while the large deletion was unexpected, the experiment proved efficacious in disrupting chartreusin production and resulting in a notable genome reduction of over 200 Kb. The genome reduction did not elicit any noticeable variances in growth and sporulation rate or morphology.

Taken together, the outcomes from the gene deletion and promoter insertion experiments underscored the compatibility of *S. sungeiensis* SD3 with CRISPR-Cas9-based genome editing techniques. The unexpected large deletion is not an uncommon phenomenon as it has been observed by other research labs (Hoff et al., 2018). The observations reinforce the view that judicious selection of the PAM and Cas9-cutting sites is needed to avert untoward instances of homologous recombination events in *S. sungeiensis* SD3 and other *Streptomyces* strains.

2.3.7 Heterologous production of secondary metabolites using *S. sungeiensis* SD3

To establish *S. sungeiensis* SD3 as a useful microbial chassis for expressing heterologous BGCs from closely related phylogenetic strains, we tested the BGCs from the *S. tasikensis* P46 strain. While sharing a close phylogenetic relationship

with *S. sungeiensis* SD3 (see Figure 2B), *S. tasikensis* P46 has a slower growth rate, typically requiring more than 14 days to reach maturity on agar plates, in contrast to the rapid five-day growth cycle of *S. sungeiensis* SD3 (Ma et al., 2022). Based on our experience, the production of secondary metabolites, including tasikamides, furaquinocins, and pentalenolactones, are inconsistent in *S. tasikensis* P46. *S. tasikensis* P46 is also difficult to manipulate genetically and not amenable to CRISPR/Cas9 editing due to the toxicity of Cas9 for the strain. Hence, if the expression of the BGCs of *S. tasikensis* P46 in *S. sungeiensis* SD3 is successful, it will not only enable the production of the secondary metabolites in a less erratic manner but also offer the opportunity to activate the novel cryptic BGCs of *S. tasikensis* P46 by heterologous expression.

We obtained a genomic DNA library of *S. tasikensis* P46 constructed using the P1-derived artificial chromosome (PAC) vector pESAC13A from BioS&T (Canada, https://www.biost.com/pesac_13-libraries). From the genomic library, six clones that contain the BGCs predicted to produce polyketides, non-ribosomal peptides, terpenoids, and ribosomal synthesized and post-translationally modified peptides (RiPPs) were obtained. The six BGC-harbouring pESAC13A plasmids were subsequently integrated into the chromosome of *S. sungeiensis* SD3 via *E. coli-Streptomyces* conjugation. The transformed *S. sungeiensis* SD3 strains were cultivated in various liquid and solid fermentation cultures for the production of secondary metabolites. Among the six transformed strains, *S. sungeiensis* SD3::P46_BGC23 and *S. sungeiensis* SD3::P46_BGC28 were found to produce

compounds that were not produced by the *S. sungeiensis* SD3 control strain transformed with empty pESAC13A (Figure 7A).

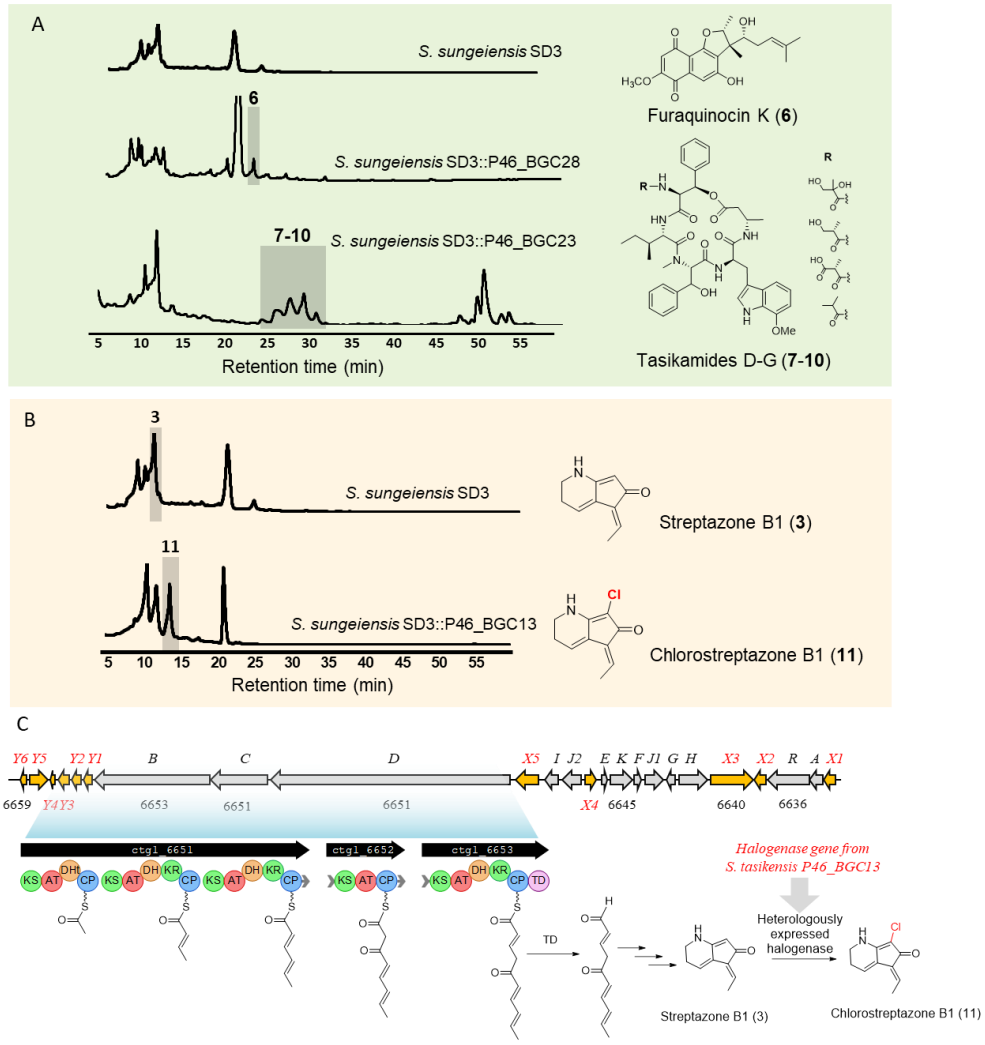


Figure 7. Heterologous production of secondary metabolites using the *S. sungeiensis* SD3 host. (A) HPLC chromatogram and the compounds produced by the heterologously expressed BGC28 and BGC23 from *S. tasikensis* P46. (B) HPLC chromatogram showing the production of streptazone B1 (**3**) and chlorostreptazone B1 (**11**). (C) Heterologously expressed halogenase converts streptazone B1 (**3**) produced by an endogenous biosynthetic pathway to chlorostreptazone B1 (**11**) in *S. sungeiensis* SD3.

P46_BGC28 encodes a pathway predicted to produce the polyketide-terpene hybrid products known as furaquinocins. This pathway involves a prenyltransferase that introduces an isoprene unit and a Type III PKS that generates the 1,3,6,8-tetrahydroxynaphthalene (THN) skeleton through the condensation of five units of malonyl-CoA (Kumano et al., 2010). Analysis of the HPLC chromatogram from the *S. sungeiensis* SD3::BGC28 strain revealed the presence of a new compound (*m/z* 372.1573) in the culture broth when cultured in a GYMose liquid medium. Subsequent isolation and structural elucidation confirmed that it is a novel furaquinocin and herein named as furaquinocin M (**6**) (Figures 7A, S3, and S10).

P46_BGC23 codes for an NPRS pathway that produces the cyclic peptides tasikamides. HPLC analysis of the fermentation products of the *S. sungeiensis* SD3::BGC23 strain revealed the presence of a group of metabolites that exhibit similar UV-Vis spectra, molecular masses ($[M + H]^+$, *m/z* 809.38, 841.41, 855.38, and 857.40), and retention times as tasikamides D-G (**7**-**10**) (Ma et al., 2022). Tasikamides D-G represent a unique family of non-ribosomal peptides characterized by the presence of methylated and hydroxylated alpha and beta amino acid building blocks (Ma et al., 2022). The production of compounds **7**-**10**, in conjunction with the presence of several NRPS-encoding BGCs in the *S. sungeiensis* SD3 genome, suggests *S. sungeiensis* SD3 could be exploited as a robust chassis to produce non-ribosomal peptides. It is worth noting that the production of tasikamides in *S. sungeiensis* SD3 only required five days in GYMose liquid medium (versus 15 days for *S. tasikensis* P46 on solid medium), and that *S. sungeiensis* SD3 produces **7**-**10** in higher fermentation titer (2 to 4-fold by

estimation) than *S. tasikensis* P46, highlighting its advantages for the production of tasikamides and potentially other NRPS peptides.

P46_BGC13 encodes a marinopyrrole pathway that contains six putative halogenase genes. Although the *S. sungeiensis* SD3 host transformed with P46_BGC13 did not produce the anticipated marinopyrroles, the strain produced a new compound consistently when compared to the control strain when cultured in an oatmeal liquid medium (Figure 7B). We isolated and characterized the compound to establish its identify as a hitherto unreported compound, chlorostreptazone B1 (**11**, $[M + H]^+$, *m/z* 196.08). The structure of **11** was unambiguously established by 1D and 2D NMR spectroscopic studies (Figures S1 and S9). Unlike *S. sungeiensis* SD3, the *S. sungeiensis* SD3::P46_BGC13 strain produced negligible amount of streptazone B1. We reasoned that one of the six halogenases encoded by P46_BGC13 was expressed functionally in *S. sungeiensis* SD3, resulting in the cross-chlorination of streptazone B1 produced by one of the endogenous pathways encoded by BGC13 (Figure 7C). This unexpected finding suggests that some of the biosynthetic enzymes from the BGC13 were expressed in *S. sungeiensis* SD3, and that the halogenase encoded by one of the halogenase encoded by P46_BGC13 can chlorinate the non-native substrate streptazone B1 duo to relaxed substrate selectivity.

2.4 Discussion

We have identified *S. sungeiensis* SD3 as a promising microbial chassis for the heterologous production of secondary metabolites. This strain possesses several advantageous characteristics that make it well-suited for expressing heterologous biosynthetic pathways. *S. sungeiensis* SD3 occupies a unique position within a distinct branch on the phylogenetic tree, distinguishing it from previously developed *Streptomyces* chassis strains. This phylogenetic classification designates *S. sungeiensis* SD3 as the preferred chassis for expressing biosynthetic gene clusters (BGCs) originating from *Streptomyces* strains within the same phylogenetic clade. The use of closely related chassis could enhance the likelihood of successful expression and production of secondary metabolites, owing to the shared cellular environment and metabolic capabilities with the native producer. Furthermore, *S. sungeiensis* SD3 harbours primary metabolic networks that support the synthesis of polyketides and non-ribosomal peptides, two classes of natural products renowned for their diverse biological activities and promising therapeutic applications. Based on the strain's inherent metabolic capabilities, *S. sungeiensis* SD3 can be exploited as a versatile chassis for the heterologous production of the polyketides and non-ribosomal peptides. *S. sungeiensis* SD3 also demonstrates remarkable genetic tractability, rendering it readily amenable to genetic manipulation for strain improvement through rational approaches.

The successful production of the polyketide-terpenoid compound **6** and NRPS peptides **7-10** demonstrates the versatility of *S. sungeiensis* SD3 as a microbial chassis for expressing BGCs from closely related strains. The expedited

fermentation period and heightened fermentation titer further highlight the advantages associated with employing *S. sungeiensis* SD3 as a heterologous expression host. It is noteworthy that, among the six tested BGCs, four failed to produce in both the native producer (P46) and the host (SD3), indicating that silent BGCs in the native producer are likely to remain as inactive in a heterologous host. Consequently, the activation of these silent BGCs through genetic or non-genetic means remains imperative for the heterologous production of secondary metabolites. Unlike many undomesticated wild-type strains, the outstanding genetic tractability of *S. sungeiensis* SD3 will enable the genetic engineering of silent BGCs within the host, facilitating their activation. The availability of genome-editing tools further renders *S. sungeiensis* SD3 amenable to metabolite engineering, boosting precursor and cofactor supplies to enhance fermentation titers.

The accessibility of its comprehensive genome sequence and transcriptomic profile paves the way for advanced strain enhancement through genome editing. We envision that optimizing *S. sungeiensis* SD3 via genome editing will result in specialized mutant strains tailored for the efficient production of diverse classes of secondary metabolites. Notably, the removal of the chartreusin BGC and the adjacent 200kb region has produced the *S. sungeiensis* SD3_Δ240kb strain that is characterized by genome reduction and a cleaner HPLC background. Concurrently, the introduction of the *sfp* gene yielded the *S. sungeiensis* SD3_*ΔchaAB::sfp* strain that may facilitate the activation of silent BGCs reliant on post-translational modification through phosphopantetheinylation (Zhang et al., 2017). These two

mutant strains will serve as promising starting points for the development of customized chassis tailored for the production of polyketide natural products.

In conclusion, our findings underscore the potential of *S. sungeiensis* SD3 as a promising microbial chassis for heterologous expression of secondary biosynthetic pathways. The capacity to express diverse biosynthetic pathways, coupled with its genetic tractability, positions *S. sungeiensis* SD3 as a valuable and complementary chassis within the *Streptomyces* chassis repertoire for uncovering novel natural products and producing commercially valuable compounds. Ongoing optimization endeavours, encompassing both genetic and non-genetic strategies, aim to yield specialized mutant strains tailored for the production of distinct natural product classes. This optimization process also seeks to streamline metabolic networks, ultimately achieving higher fermentation titers essential for industrial applications.

Chapter 3 Characterization and engineering of *Streptomyces albus* MD102 as a heterologous host

3.1 Introduction

Actinomycetes are highly prolific microbial producers of secondary metabolites synthesized by the enzymes encoded by biosynthetic gene clusters (BGCs) (Lee et al., 2020). The abundance of BGCs and the widespread distribution of actinomycete strains in the environments and human digestive tracts suggest the underexplored potential of discovering secondary metabolites with novel structure and bioactivity (Lacey and Rutledge, 2022; Nepal and Wang, 2019). However, many cryptic BGCs remain dormant under typical laboratory culture conditions due to repressed gene expression and other inhibitory factors (Liu et al., 2021). Activating silent BGCs to disclose their cryptic secondary metabolites is often challenging, with many reported successes owing to serendipitous discovery resulted from screening elicitors or fermentation conditions (Bode et al., 2002; Craney et al., 2012; Xia et al., 2020)

One approach that is being explored to access the products of silent cryptic BGCs involves the expression of silent cryptic BGCs in a genetically manipulable and fast-growing heterologous chassis (Ahmed et al., 2020). Several actinomycetes strains from the *Streptomyces* genus, including *S. coelicolor*, *S. lividans*, *S. avermitilis*, *S. albus*, *S. chattanoogensis*, *S. griseofuscus*, *S. atratus*, and *S. venezuelae* have been rendered to genome engineering and chassis optimization (Ahmed et al., 2020; Gomez-Escribano and Bibb, 2011; Hwang et al., 2021; Liu et al., 2016; Myronovskyi et al., 2018; Whitford et al., 2021; Yang et al., 2022).

Among the *Streptomyces* chassis, *S. albidoflavus* J1074 has emerged as a robust and versatile chassis employed by many academic researchers. *S. albidoflavus* J1074 (formerly known as *S. albus* J1074) is a derivative of the *S. albus* strain G that lacks the SalI restriction endonuclease gene. *S. albidoflavus* J1074 can be used to express BGCs from both *Streptomyces* and non-*Streptomyces* strains such as Micromonospora, as demonstrated by the production of thiocoraline, tetarimycin A, Clareposxcins, landepoxins, and other secondary metabolites (Kang and Kim, 2021; Myronovskyi and Luzhetskyy, 2019). Myronovskyi et al created the *S. albidoflavus* J1074 Del14 mutant strain by deleting 15 BGCs from the genome of *S. albidoflavus* J1074 (Myronovskyi et al., 2018). The Del14 strain was further modified by introducing additional phage phiC31 attB sites to improve the production yield of heterologously produced secondary metabolites. Studies demonstrated that the engineered strains not only exhibited higher fermentation titre, but also produced secondary metabolites coded by silent BGCs (e.g., pyridnopyrone, salicylic acid, fralnimycin bhimamycin A and aloesaponarin II).

Here we report the characterization and engineering of *Streptomyces albidoflavus* MD102 as a suitable microbial chassis for the expression of BGCs. *S. albidoflavus* MD102 was a recently isolated strain that is closely related to the established *S. albidoflavus* J1074. *S. albidoflavus* MD102 displays rapid growth and efficient sporulation on standard solid and liquid culture media. Its susceptibility to commonly used antibiotics in research laboratories and natural competence allow easy transformation through *E. coli-Streptomyces* conjugation. Using CRISPR-Cas9 tools for genome editing, we successfully eliminated

competing biosynthetic gene clusters (BGCs) and introduced genes to improve the strain for heterologous expression of BGCs. *S. albidoflavus* MD102 stands as a valuable supplement to the repertoire of *Streptomyces* chassis, offering potential for utilization in heterologous expression systems and contributing to progress in synthetic biology and biotechnological research.

3.2 Materials and Methods

Isolation of actinobacterial strains

1m depth soil sediments were previously obtained from Sungei Buloh Wetland Reserve, Singapore (1°26'41.172"N, 103° 43'36.12"E) via a sterile stainless-steel soil sampler. The samples were desiccated using microwave treatment, and sterile water used in various dilutions to obtain single colonies on various solid agar media.

Actinobacterial culture conditions

Routine culturing of Streptomyces strains was done on mannitol soy (MS) solid media (20 g/L BactoTM agar, 20 g/L mannitol, and 20 g/L soya flour), tryptic soy broth (TSB) (30 g/L TSB powder) (Merck Millipore, Germany) liquid media and GYMose (20 g/L glucose, 10 g/L maltose, 5 g/L yeast extract and 2 g/L CaCO3) liquid media. Selective antibiotics were used at 50 µg/mL apramycin, 25 µg/mL chloramphenicol, 50 µg/mL kanamycin, and 25 µg/mL nalidixic acid. Culturing of *Streptomyces* strains was done at 30c, and for liquid culture shaking was carried out at 180 rpm.

Actinobacterial Genomic DNA extraction

The extraction of Streptomyces genomic DNA was done using the Hexadecyl trimethyl-ammonium bromide (CTAB) buffer (2 g CTAB, 10 mL 1 M Tris (pH 8), 4 mL 0.5 M ethylenediamine tetra-acetic acid disodium salt (pH 8), 28 mL 5 M NaCl, 1 g polyvinyl pyrrolidone (MW 40 kDa) and 58 mL water) method.

Phylogenetic and Comparative Genomics analysis

Next-generation whole genome sequencing via Illumina 2x150bp paired-end configuration was used to obtain the genome sequence of *S. albidoflavus* MD102. The establishment of the phylogenetic relationships and closest related strains for *S. albidoflavus* MD102 was carried out using the TYGS server (https://tygs.dsmz.de) and and visualised by the online software iTOL (https://itol.embl.de/). The OrthoANIu value was obtained from the ANI calculator from EZBioCloud. RAST (Rapid Annotation using Subsystem Technology) was used to predict the open reading frames and carry out genome annotation. OrthoVenn3 was used to compare orthologous protein clusters. The antiSMASH database was used to identify biosynthetic gene clusters. Unique aromatic degradation genes were blasted against common *Streptomyces* genomes in the NCBI including *Streptomyces coelicolor* A3(2) (taxid:100226), *Streptomyces avermitilis* MA-4680 (taxid:227882), and *Streptomyces albidoflavus* (taxid:1886). These genomes were used as queries for ortholog identification, and results were filtered by 75% sequence identity, 50% coverage, and 1e-10 expected value.

Fermentation and metabolite profiling

Streptomyces strains were first cultured on MS agar. Spores from the MS agar plates were inoculated into 10ml liquid TSB media and cultured at 30c 180rpm for 4-5 days. Subsequently, this starter culture was used to subculture 50ml of liquid media at 2% and grown for 1 week. Ethyl acetate was used for metabolite extraction

of the culture broth while methanol was used for the biomass. HPLC analysis was performed on Agilent1200 HPLC-DAD system equipped with an ODS column (Pursuit XRs: 250 mm × 4.6 mm, 5 μm) using a linear gradient of CH3CN in H2O with 0.1% formic acid (0–5 min, 10%–20% CH3CN; 5–35 min, 20%–70% CH3CN; 35–50 min, 70%–90% CH3CN; 50–60 min, 90%–100% CH3CN; 60–70 min, 100% CH3CN) at a flow rate of 1.0 mL/min.

Molecular biology techniques and molecular cloning

PCR reactions were carried out using Q5 High Fidelity Polymerase (NEB) in accordance with the manufacturer's instructions. The purification of plasmids (Qiagen Spin Miniprep) and gel extractions (Zymo Research Zymoclean Gel DNA Recovery Kit) were done according to the manufacturer's instructions. The purified DNA was eluted in Milli-Q water (Merck Millipore) or provided elution buffer. DNA digestion was carried out using various NEB restriction enzymes at their recommended optimum temperatures overnight. pCRISPR-Cas9 derivatives containing sgRNA and homology arms were constructed as described by Tong et al.

CRISPR/Cas9-assisted genome modification

For conjugation of plasmids including the pCRISPR-Cas9 vector into *Streptomyces*, *E. coli* ET12567/pUZ8002 strains containing the vector of interest were grown at 37c with 50 μg/μL apramycin, 25 μg/μL chloramphenicol and 50 μg/μL kanamycin. Successful exconjugants were screened on MS plate containing

50 µg/mL apramycin and 25 µg/mL nalidixic acid. Subsequent 5 µg/mL thiostrepton was used for the activation of the tipA inducible promoter on pCRISPR-Cas9. Curing of pCRISPR-Cas9 plasmid was done in TSB liquid culture at 37c for one week.

Measurement of promoter strength

Streptomyces spores were first inoculated into 50ml TSB liquid media starter cultures for two days, before subculturing into 50ml TSB liquid media for another three days with the addition of 30 5-mm glass beads. 10ml of culture was extracted and washed and resuspended with 1x PBS. 100ul of cells from 10ml of 1x PBS cell suspension was used for measurement of OD600 and normalization of cell concentration to 0.5 OD600. Fluorescence measurements were carried out using a Tecan Infinite® 200 PRO Plate Reader. For EGFP, the excitation and emission wavelengths used were 485 nm and 515 nm respectively and readings were done in triplicates for each sample.

RNA isolation and sequencing

S. albidoflavus MD102 was grown in TSB media for 72hrs in baffled flasks. Cells were collected and homogenized using a pestle. 1mg/ml working concentration of lysozyme was used for incubation at 37 ºC for half an hour. Cells were freeze-thawed using liquid nitrogen for 5-10 cycles. TRIzol Reagent was added according to the manufacturer's instructions (Invitrogen). Finally, total RNA was extracted and purified using the Purelink RNA Mini kit according to

manufacturer's instructions (Thermo Fisher Scientific). DNAse-I (Thermo Fisher Scientific) was added in accordance with the Purelink RNA Mini kit protocol.

Flaviolin production in *S. albidoflavus* MD102

pIJ8630-fur1 was conjugated into MD102SL01 as described above, and showed pink colouration when cultured on mannitol soy (MS) solid media. MD102SL01-fur1 exconjugants were cultured in liquid media (GYMose and TSB) in baffled flasks for 5 days before extraction of metabolites using ethyl acetate and methanol.

3.3 Results

3.3.1 Isolation and taxonomic classification of *Streptomyces albidoflavus* MD102

We recently isolated a collection of actinomycete strains from local soil and aquatic environments in Singapore. Among them, strain MD102, initially derived from sediment obtained from the Sungei Buloh Wetland Reserve, exhibited remarkable growth and sporulation, surpassing several commonly used *Streptomyces* chassis, including *Streptomyces* chassis, *S. coelicolor* M1152 and *S. lividans* TK24 (Figure 8). Initial 16S rRNA sequencing analysis suggested its close relationship with *Streptomyces violascens* ISP 5183, with a 100% similarity at 93.8% coverage according to the EzBioCloud database. However, recognizing the limitations of small subunit rRNA gene sequences for *Streptomyces spp*. phylogenetic classifications, we adopted a whole-genome approach to attain species-level resolution of MD102's phylogenetic relationship.

Utilizing a PacBio and Illumina hybrid approach, whole genome sequencing produced a total length of 7,047,131 bp and a GC content of 73.3%. Phylogenetic relationships were established by constructing a Genome BLAST Distance Phylogeny (GBDP) tree using the Type Strain Genome Software (TYGS). MD102 clustered within a monophylogenetic clade alongside *S. albidoflavus* J1074, *S. coelicolor* DSM 40233, *S. sampsonii* NBRC 13083, *S. limosus* NBRC 12790, *S. albidoflavus* DSM 40455, and *S. albidoflavus* NRRL B-1271. Notably, *S. albidoflavus* J1074 exhibited the highest GBDP score with MD102, registering d_0, d_4, and d_6 scores of 92.5, 90.1, and 94.4, respectively, with a 0.0% difference in GC

content. Further confirming this relationship, the OrthoANIu value of 98.81% obtained from the EzBioCloud database in the MD102 and *S. albidoflavus* J1074 comparison surpassed the ANI threshold value of 95%, leading to the classification and naming of the strain as *S. albidoflavus* MD102.

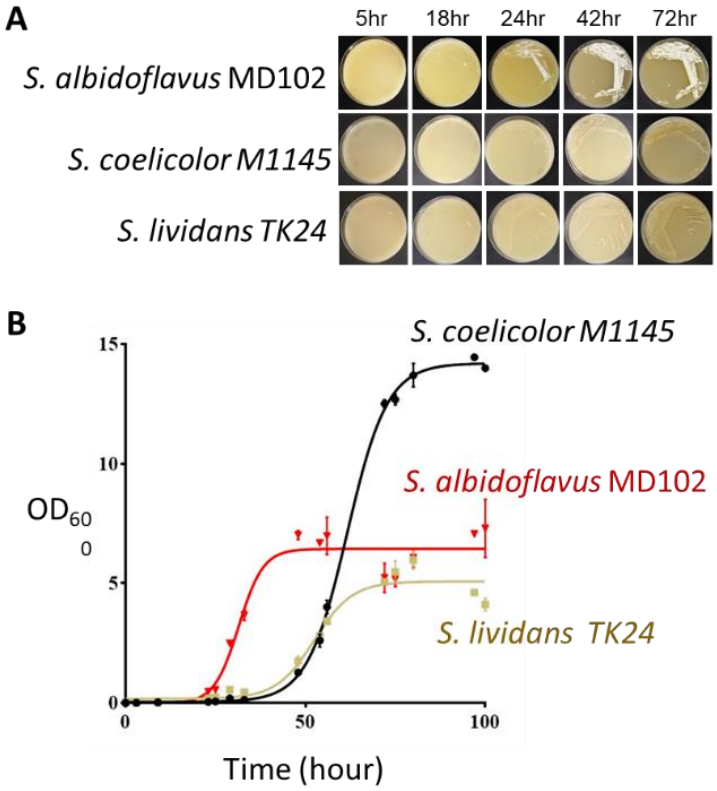


Figure 8. Comparison of the growth of *S. albidoflavus* MD102 to two commonly used *Streptomyces* hosts. A. solid medium. B. CGM liquid medium (30C).

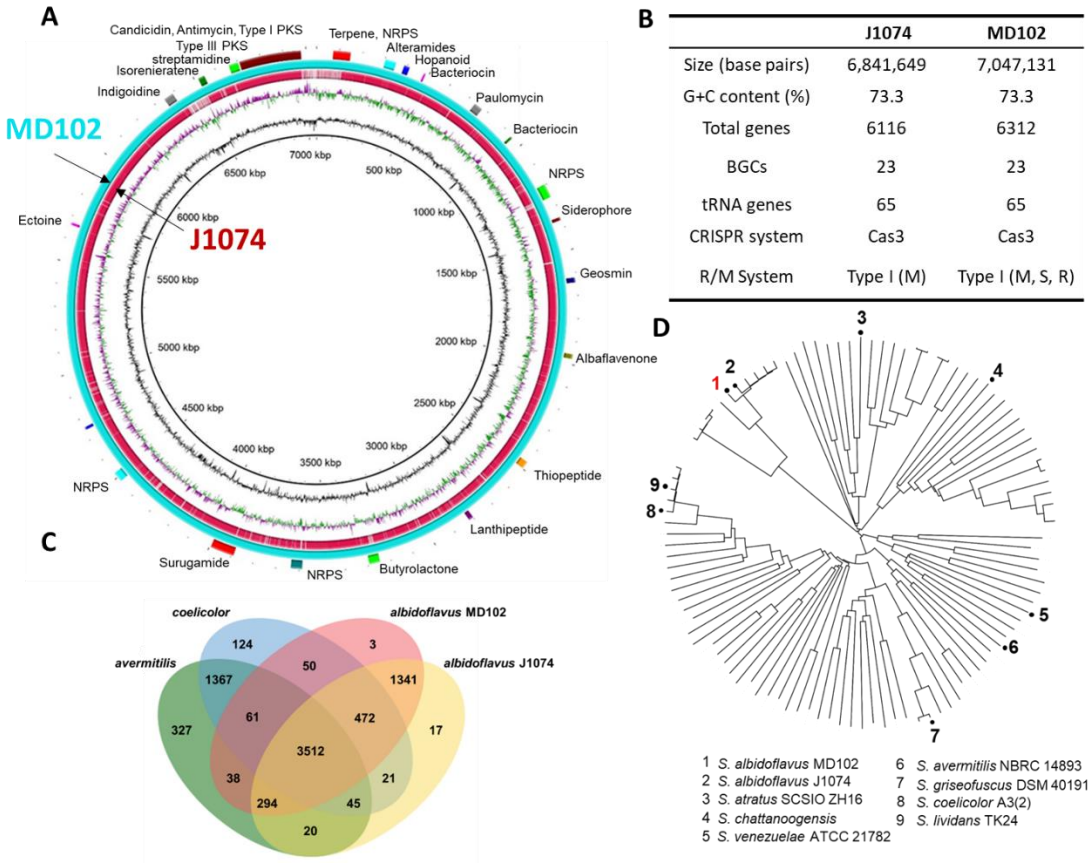
3.3.2 Genomic features of *S. albidoflavus* MD102

S. albidoflavus J1074 is one of the most used heterologous chassis known for its fast growth and natural competence. Along with *S. venezuelae*, *S. albidoflavus*

J1074 contains seven rRNA operons (16S-23S-5S) which might explain the fast growth rate of these two commonly used heterologous hosts. Its small genome of 6,841,649 bp also has one of the highest GC content (73.3%) among *Streptomycetes*. Compared to *S. albidoflavus* J1074, *S. albidoflavus* MD102 has a slightly larger genome (7,047,131 bp) and almost the same GC content. Genome annotation of *S. albidoflavus* MD102 using RAST identified 65 tRNA genes, and seven rRNA operons, the same as *S. albidoflavus* J1074 (Figure 9b, Table S10. Like *S. albidoflavus* J1074, *S. albidoflavus* MD102 lacks the SalI restriction modification system that facilitates genetic transformation. However, it possesses the restriction subunit R of the Type I restriction modification system, which could interfere with its ability to accept exogenous DNA.

RAST annotation led to the identification of 6,312 protein coding ORFs, slightly more than the 6,116 found in *S. albidoflavus* J1074. OrthoVenn3 was utilized to compare orthologous proteins of *S. albidoflavus* MD102 with *S. albidoflavus* J1074, *S. coelicolor* and *S. avermitilis*. A total of 5771 orthologous proteins were identified in *S. albidoflavus* MD102 by OrthoVenn3, with 1,341 shared with *S. albidoflavus* J1074, whilst only 50 with *S. coelicolor* and 38 with *S. avermitilis*, which agrees with their phylogenetic relationship (Figure 9C). In comparison with the other three model strains, *S. albidoflavus* MD102 has three unique and strain-specific uncharacterized protein clusters. Many of the unique genes are found in the terminal regions of the chromosome (refer to the blue and red circles in Figure 9A). Among the genes found in *S. albidoflavus* MD102 but not in *S. albidoflavus* J1074 are several genes involved in aromatic degradation

including biphenyl dioxygenases and other aromatic dioxygenases. This might be attributed to the prevalence of polycyclic aromatic hydrocarbons (PAHs) in the mangrove and coastal sediments from which *S. albidoflavus* MD102 was isolated, as a result of Singapore's status as one of the busiest ports in the world with concomitant heavy shipping and petroleum refinery activities (Basheer et al., 2003). Of these 9 genes present in *S. albidoflavus* MD102 but not in *S. albidoflavus* J1074, 6 of them appear to form a cluster with the same orientation, with PEG6088, PEG6089 and PEG6090, and PEG6092 and PEG6093 possessing ATGA unidirectional gene overlap regions (Table S9) (Wright et al., 2022). This putative cluster was not found in other *albidoflavus* or "*albus*" strains, with the exception of *Streptomyces albus* SM254, a strain isolated from copper-rich subsurface of a mine (Badalamenti et al., 2016).



	J1074	MD102
Size (base pairs)	6,841,649	7,047,131
G+C content (%)	73.3	73.3
Total genes	6116	6312
BGCs	23	23
tRNA genes	65	65
CRISPR system	Cas3	Cas3
R/M System	Type I (M)	Type I (M, S, R)

Figure 9. Phylogenetic and genomic characterization of *S. albidoflavus* MD102. (A) Chromosomal genome map of *S. albidoflavus* MD102 with comparisons against *S. albidoflavus* J1074. Tracks from outermost circle to the center: (i) Biosynthetic gene clusters of *S. albidoflavus* MD102 predicted by antiSMASH 7.0.0. (ii) Predicted coding DNA sequence (CDS) of *S. albidoflavus* MD102 in blue. (iii) Predicted coding DNA sequence (CDS) of *S. albidoflavus* J1074 in red. (vi) GC skew of *S. albidoflavus* MD102, the inward facing portion in green indicating higher than average GC skew, and the outward portion in purple indicating below average GC skew. (v) GC content of *S. albidoflavus* MD102, the outward facing portion indicating higher than genome-wide average GC content, whilst the inward portion indicating lower than average GC content. (vi) Marker for for genome size. The genome figure was constructed using DNAPlotter v 18.1.0. (B)

Genome features of *S. albidoflavus* J1074 and *S. albidoflavus* MD102. (C) Venn diagram of unique and shared orthologous gene clusters amongst *S. albidoflavus* MD102, *S. albidoflavus* J1074 (CP004370.1), *S. avermitilis* NBRC 14893 (NC_003155), and *S. coelicolor* A3(2) (CP042324). (D) Phylogenetic relationship between *S. albidoflavus* MD102 and commonly used *Streptomyces* chassis.

3.3.3 The biosynthetic capability of *S. albidoflavus* MD102

We mined the biosynthetic gene clusters (BGCs) of *S. albidoflavus* MD102 using antiSMASH 7.0 to evaluate the biosynthetic capacity of *S. albidoflavus* MD102. The AntiSMASH analysis identified 23 Biosynthetic gene clusters (BGCs), the same number as *S. albidoflavus* J1074 (Table S8). However, a close inspection of the BGCs revealed that two of the BGCs are different between the two strains. In *S. albidoflavus* MD102, BGC1 encodes terpene synthase and NRPS-like biosynthetic enzymes and BGC13 encodes polyketide synthase (PKS). In *S. albidoflavus* J1074, BGC1 codes for a T1PKs-NRPS hybrid and BGC12 codes for a Ribosomally synthesized and post-translationally modified peptides (RiPP) pathway.

We cultured *S. albidoflavus* MD102 in various culture media to identify constitutively active BGCs, with most compounds observed being produced in R5A liquid medium. HPLC and LC/MS-based metabolite profiling identified the production of several metabolites that include the polyketide-derived secondary metabolites alteramides and the polyketide-peptide hybrid antimycin (Figure 10B). The production of the antifungal polyene compound candicidin is in agreement with

the observation that *S. albidoflavus* MD102 displayed strong antifungal activity against *Candida albicans*, with a large zone of inhibition on antifungal assay plate overlay (date not shown). The results from the AntiSMASH mining analysis underscore the potential of *S. albidoflavus* MD102 to synthesize a wide array of secondary metabolites, indicating that its genetic makeup encompasses a robust metabolic network conducive to the heterologous production of secondary metabolites. Considering that only a small fraction of the predicted secondary metabolites has been identified by the metabolite profiling, many of the endogenous BGCs are likely to be silent under laboratory fermentation conditions. The discrepancy between the large number of secondary metabolites suggested by AntiSMASH analysis and the small number of metabolites detected underscores the highly regulated production of secondary metabolites in *S. albidoflavus* MD102 as observed for other *Streptomyces*.

3.3.4 Strain improvement by genome editing

To enhance *S. albidoflavus* MD102 as a heterologous host, we employed the CRISPR/Cas9 tool for rational genome editing. In comparison to the RedET double crossover methodology utilized for BGC deletion in *S. albidoflavus* J1074, the CRISPR/Cas9 method proves more efficient and avoids leaving scars in the chromosome. Our focus was on the targeted deletion of constitutively active BGCs, including alteramide, paulomycin, BGC16 (NRPS), candicidin, and antimycins. These deletions aimed to augment the availability of precursors (e.g., amino acids, acetyl-CoA, malonyl-CoA) and cofactors (e.g., NADPH) while simplifying the

HPLC profile. This streamlined profile would facilitate the detection of compounds and simplify the compound-isolation process. Moreover, we designed the knock-out CRISPR/Cas9 plasmids to concurrently introduce exogenous genes, to enhance heterologous expression and BGC activation.

The inserted genes encompass the global regulator bldA gene sourced from *S. coelicolor*, encoding tRNAleu (UAA). This rare codon UAA is known to constrain the expression of biosynthetic enzymes in *Streptomyces*. The integration of bldA, coupled with a constitutively active promoter, is anticipated to facilitate the translation of biosynthetic enzymes. Another incorporated gene is *gpps* from *S. tasikensis* P46, responsible for encoding a geranyl diphosphate synthase that synthesizes geranyl pyrophosphate (GPP) crucial for the biosynthesis of diverse terpenes and terpenoids. The introduction of *gpps* aims to augment the cellular supply of GPP, enhancing the production of terpene-based natural products. Additionally, we integrated an additional attachment site, ΦBT1-attB, into the BGC23a region. This novel attachment site differs from the ΦC31-attB site present in *S. albidoflavus* MD102, enabling the integration of biosynthetic genes into two distinct sites.

For the genome-editing experiments, the temperature-sensitive pSG5-based pCRISPR-Cas9 plasmids containing the relevant homologous arms and sgRNA sequences were introduced into *S. albidoflavus* MD102 (see Figure 10A). The culmination of the genome editing process, marked by the deletion of the antimycin and other constitutively expressed BGCs, led to a substantially simplified HPLC profile for the resulting mutant strain *S. albidoflavus* MD102SL01. Illumina-based

whole genome sequencing was conducted to check the integrity of the MD102SL01 genome after multiple rounds of CRISPR-Cas9 facilitated genome editing. Genome sequencing affirmed the successful deletion of paulomycin, BGC16-NRPS, candicidin, and antimycin BGCs, as well as the insertion of bldA, gpps, and ΦBT1-attB, thereby validating the efficacy of the CRISPR-Cas9 facilitated genome editing efforts. However, an unforeseen deletion of the 5' and 3' chromosome end regions was detected. At the 5' end, a loss of 154,300 bp occurred, and at the 3' end, a loss of 292,451 bp transpired, summing up to a total loss of 446,751 bp. This unexpected deletion resulted in the additional loss of BGC1 and BGC23 clusters. This substantial deletion event was likely an outcome of the BGC23b (antimycin BGC) deletion during the fifth round of CRISPR-Cas9 editing. Together, our results showed the compatibility of *S. albidoflavus* MD102 with CRISPR-Cas9-based genome editing techniques. The occurrence of this unexpected chromosomal deletion is not an uncommon phenomenon, as documented in our previous studies and those conducted by others (Hoff et al., 2018). These findings underscore the importance of carefully choosing the PAM and Cas9-cutting sites to mitigate the risk of inadvertent homologous recombination events in *S. albidoflavus* MD102 and other *Streptomyces* strains.

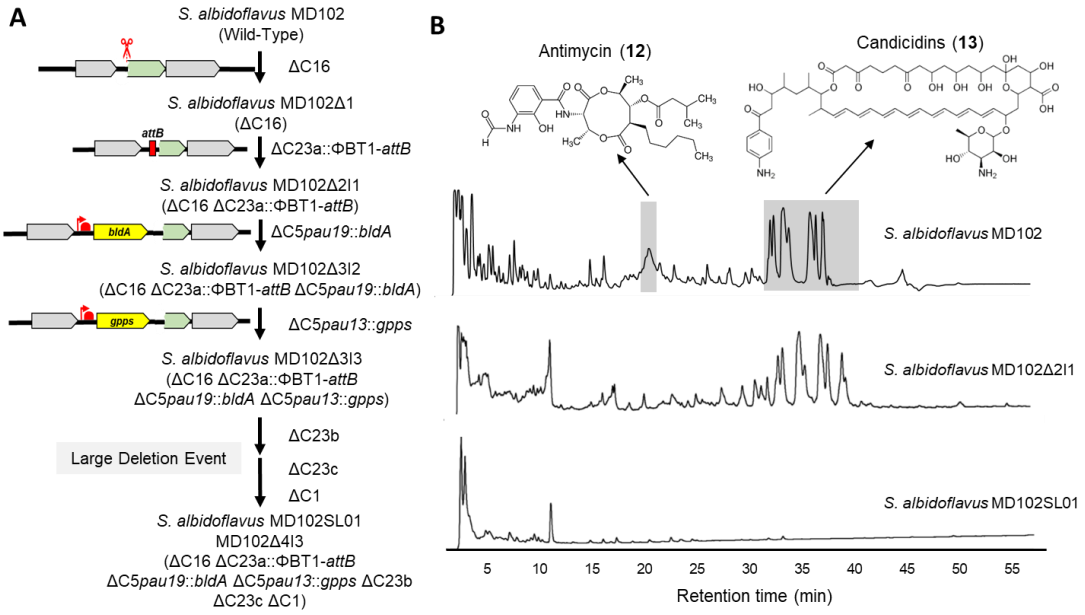


Figure 10. Production of secondary metabolites by *S. albidoflavus* MD102 wild-type and mutant strains. (A) Pedigree of *S. albidoflavus* MD102 mutant strains. Following several rounds of CRISPR-Cas9 facilitated genome editing and an unexpected large deletion event, four endogenous secondary metabolite gene clusters were deleted, and *bldA*, *gpps*, and ΦBT1-attB were introduced into the genome. (B) HPLC analysis confirmed the loss of antimycin and candicidins production in *S. albidoflavus* MD102 mutant strains

3.3.5 Assessment of genetic tools for heterologous expression in *S. albidoflavus* MD102

We assessed the natural competence of *S. albidoflavus* MD102 using commonly employed vectors by academic researchers, including conjugative/integrative vectors derived from pIJ101 and pSET152, such as pIJ12551, pSET, and pIJ8630. Through *E. coli-Streptomyces* conjugation, these vectors were efficiently introduced into *S. albidoflavus* MD102, indicating

successful conjugation and integration into the chromosome. Notably, vectors relying on the ΦC31 integrase gene for chromosomal integration, such as those derived from pIJ101 and pSET152, demonstrated functionality, suggesting that the attB site in the pirin-like gene of *S. albidoflavus* MD102 is active. For the heterologous expression of large biosynthetic gene clusters (BGCs), transformation by plasmids containing substantial DNA fragments (30-150Kb) is crucial. We successfully introduced cosmid-based pIJ10702 vector and BAC-based pSMART-BAC-S into *S. albidoflavus* MD102 via *E. coli-Streptomyces* conjugation. However, attempts with the PAC-based pESAC13A vector were unsuccessful, as *E. coli-Streptomyces* conjugation did not yield any exconjugants.

Having established the transformation protocols for *S. albidoflavus* MD102, we proceeded to evaluate the strength of various widely used constitutively active promoters. These promoters were cloned upstream of the EGFP gene in the reporter vector pIJ8630 and introduced into *S. albidoflavus* MD102. By quantifying EGFP fluorescence intensity in the culture broth and normalizing against the ermEp* promoter, we observed that KasOP*, gapdhp(EL), and gapdh(KR) promoters induced the highest EGFP expression. In contrast, ermEp* and SF14p induced moderate EGFP expression (Figure 11A). A similar expression pattern was noted in *S. albidoflavus* J1074, reaffirming the comparable cellular environment and genetic background of the two strains.

Due to the cost and time-intensive nature of cloning large biosynthetic gene clusters (BGCs), an extensive evaluation of *S. albidoflavus* MD102's ability to express diverse biosynthetic pathways from other actinomycetes has not been

undertaken in this study. However, given the high similarity between *S. albidoflavus* MD102 and the well-characterized *S. albidoflavus* J1074, we anticipate that *S. albidoflavus* MD102 will exhibit behaviour similar to J1074 as a heterologous chassis. Instead of focusing on diverse BGCs, our investigation aimed to determine whether *S. albidoflavus* MD102 can effectively express biosynthetic genes under the control of orthogonal or non-native promoters—a critical aspect for the expression of refactored BGCs and engineered pathways.

To assess this, we initially cloned the fur1 gene from the furaquinocin BGC of *S. tasikensis* P46 and integrated it into the MD102 chromosome using the conjugative/integrative vector pIJ8630, coupled with the robust promoter gapdhp(KR) (Figure 11B). The fur1 gene encodes a type III polyketide synthase (PKS) responsible for catalyzing the production of 1,3,6,8-tetrahydroxynaphthalene (THN) from malonyl-CoA molecules, leading to the formation of the red-colored pigment flaviolin. Culturing *S. albidoflavus* MD102SL01::fur1 in TSB liquid media resulted in the successful production of the anticipated red-colored pigment, flaviolin, validated against a standard (Figure 11B).

Furthermore, we undertook the cloning and refactoring of a cryptic BGC (from *S. tasikensis* P46) predicted to produce a diterpenoid natural product (Table S11). A refactored BGC was created with the three genes of the cryptic BGC equipped with two orthogonal promoters and terminators (Figure 11C). Following the design of the native BGC, the terpene synthase (ts) and the cytochrome P450 (P450) genes were kept in the same operon without adding an extra promoter for the P450 gene. While successful integration of the refactored BGC into the

MD102SL01 chromosome was achieved via *E. coli-Streptomyces* conjugation, the expected diterpenoid product was not detected in the fermentation culture across various media. RT-PCR analysis revealed expression of the geranylgeranyl pyrophosphate synthetase (ggpps) and the ts genes, but the mRNA of the cytochrome P450 was not detected, even with its own promoter (Figure 11D). This suggests that the anticipated diterpenoid was not observed likely due to the lack of cytochrome P450 expression. Together, our observations indicate that the genetic vectors and orthogonal promoters are compatible with *S. albidoflavus* MD102. However, predicting the expression levels of genes proves challenging, underscoring the necessity for experimental validation to confirm successful gene expression in this microbial chassis.

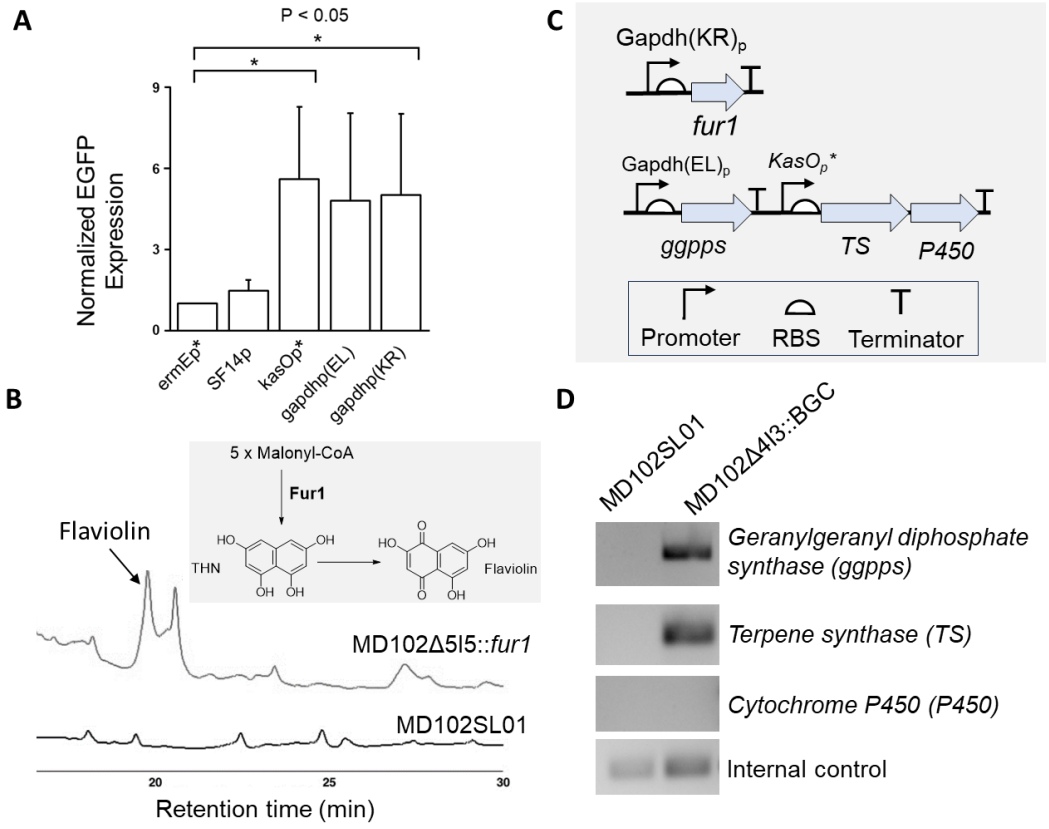


Figure 3. Heterologous production of secondary metabolites using *S. albidoflavus* MD102. (A) Promoter strength comparison in *S. albidoflavus* MD102 using constitutively active promoters and pIJ8630 EGFP reporter vector. EGFP Fluorescence intensities were normalised against ermEp, and the values are means with standard deviation (SD) from three independent experiments. (B) HPLC chromatogram showing the production of flaviolin in *S. albidoflavus* MD102 mutant. (C) Schematic showing the two exogenous biosynthetic gene circuits integrated into *S. albidoflavus* MD102. (D) Transcriptional analysis of the cryptic terpene BGC genes using RT-PCR. *hrdB* housekeeping gene is used as internal control.

3.4 Discussion

In this chapter, we present experimental data to support *S. albidoflavus* MD102 as a promising microbial chassis for the heterologous production of secondary metabolites. Sharing key characteristics with its close relative, *S. albidoflavus* J1074, this strain exhibits several advantages that make it well-suited for expressing heterologous biosynthetic pathways. *S. albidoflavus* MD102 possesses primary metabolic networks that are expected to support the production of polyketides, non-ribosomal peptides, and other natural products. Despite their close phylogenetic relationship, *S. albidoflavus* MD102 contains subsets of metabolic genes and enzymes that are not present in *S. albidoflavus* J1074. Given their distinct environmental origins, this divergence is likely the result of adaptative evolution.

The presence of genes predicted to be involved in biphenyl metabolism could indicate the contaminated nature of the environment from which *S. albidoflavus* MD102 was isolated. Polychlorinated biphenyls are among the most representative of man-made persistent organic contaminants, due to their relative volatility and chemical stability which facilitates their diffusion into the environment, whereby they bioaccumulate and remain as recalcitrant toxics (Garrido-Sanz et al., 2018). Hence, there remains a need for microorganisms that are able to facilitate the bioremediation of biphenyls and other persistent aromatic compounds (Zhu et al., 2023). Bioaugmentation is one such effective approach, by adding these microorganisms to the contaminated site for enhanced bioremediation, with the

requirement of using native strains so as to not alter the natural ecosystem (Popa et al., 2014).

The accessibility of the comprehensive genome sequence of *S. albidoflavus* MD102 enables strain enhancement through rational genome engineering. The strain displays exceptional genetic tractability, allowing for easy genetic manipulation to enhance its traits through rational approaches. The targeted removal of competing and interfering biosynthetic gene clusters (BGCs) and a 200kb region has resulted in mutant strains with reduced genome and cleaner chromatographic background. Additionally, the incorporation of an additional *attb* site, along with *bldA* and *gpps* genes, has yielded strains that will aid the expression and activation of silent BGCs, as well as enhancing fermentation titers. These engineered mutant strains represent promising starting points for the development of customized chassis tailored for the production of polyketides, terpenes, and various secondary metabolites.

While we have not conducted an exhaustive examination of the expression of large biosynthetic gene clusters (BGCs) from both *Streptomyces* and non-*Streptomyces* strains, our initial observations suggest that *S. albidoflavus* MD102 could serve as a useful microbial chassis for heterologous expression of secondary biosynthetic pathways. The distinctive genomic features of *S. albidoflavus* MD102, combined with its favourable traits, position it as a promising candidate for contributing to the field of heterologous expression systems and exploring the potential of diverse secondary metabolite production.

Chapter 4 Summary and future work

Recent developments in synthetic biology tools and genome sequencing technologies have paved the way for exploring the extensive reservoir of cryptic BGCs found in actinomycete species (Lee et al., 2019). However, addressing strain-specific challenges related to genetic tractability and suboptimal performance under standard laboratory conditions would be tedious and non-scalable (Ahmed et al., 2020). An approach to access the secondary metabolites produced by the cryptic BGCs involves the use of microbial chassis for heterologous expression of exogenous BGCs. Contrary to the notion of finding a universal "super-chassis" for producing *Streptomyces* natural products, recent studies point towards a more viable strategy of using a panel of phylogenetic diverse hosts (Hwang et al., 2021; Ke and Yoshikuni, 2020).

My Ph.D. studies focused on the characterisation and evaluation of two *Streptomyces* strains that were recently isolated by our lab (Table S12). *S. sungeiensis* SD3 was selected for its genetic tractability with BGC containing pESAC13A vectors, which allowed it to heterologously express large BGCs. The other isolated strain *S. albidoflavus* MD102 was chosen based on its growth rate, which was one of the fastest in our collection. Subsequent analysis on its genome sequence revealed it to be an *S. albidoflavus* strain, which explains its fast growth rate. For both strains, future iterative improvements can centre around increasing successful expression and production titres. To effectively utilise the native complex life cycle of *Streptomyces*, native promoters that are already linked to late growth phase of the *Streptomyces* life cycle could be identified for use in lieu of

strong constitutive promoters (Khalid et al., 2017). Additionally, global regulators that natively shift the *Streptomyces* towards secondary metabolite production like *AfsR* and the above mentioned *BldA* could be integrated into the genome (Peng et al., 2018). This would enable a metabolic approach that is more compatible with the metabolic network regulators that are already existing in the native strain. Additionally, increasing the number of attachment sites is a quick and easy method to modulate integrated BGC copy number, thereby potentially increasing metabolic production titre (Myronovskyi et al., 2018). Finally, BGC specific helper genes could be integrated depending on the chosen specialisation of BGC and natural product type. Overexpression of positive regulators governing specific BGCs integrated into the strain like luxR and SARP could be utilised to activate silent BGCs. Based on efforts to optimise metabolic flux, genes supporting production of metabolic precursors could be carried out once limiting metabolites are identified (Ko et al., 2020).

Additionally, certain issues arose which merits further study. Unfortunately, *S. albidoflavus* MD102 could not be transformed with pESAC13A, which limited the heterologous expression to smaller gene clusters using non-PAC/BAC integrative vectors. Current literature suggests either CRISPR or restriction modification systems as possible contributors to a lack of genetic tractability (Johnston et al., 2019; Yang et al., 2023). However, neither approach was successful in achieving genetic tractability for pESAC13A in *S. albidoflavus* MD102, which could implicate other defence mechanisms involved.

Another issue encountered during our work with these newly isolated strains pertained to the compatibility of the CRISPR-cas9 vectors. The initial unsuccessful attempt to transform *S. sungeiensis* SD3 using the pCRISPR vector prompted speculation about the potential contribution of reported CRISPR-cas9 toxicity (Alberti and Corre, 2019). Fortunately, a fortuitous turn of events occurred when a different CRISPR vector, pQS-idgS, featuring a different origin of replication was successfully transformed and subsequently employed to carry out genome modifications in *S. sungeiensis* SD3. This underscores the importance of considering both vector and genetic part compatibility. This is further exemplified by the differences in reported promoter strengths in model *Streptomyces* strains and with our two hosts. Despite the overarching goal of synthetic biology to establish standardized tools and parts, akin to what has been done in other engineering disciplines, biological components are seldom completely orthogonal, and instead are highly context specific in their functionality (Blount et al., 2012; Q. Liu et al., 2018).

Hence, going forward, the future of heterologous hosts for *Streptomyces* natural product expression could leverage laboratory automation to conduct extensive screens of biosynthetic gene clusters across a panel of hosts (J. Moore et al., 2023; Tong and Deng, 2020). With the advances of DNA sequencing and the increased use of machine learning in natural product prediction, putative bioactive compounds of interest and their concomitant gene sequences could be identified (Capecchi and Reymond, 2021; Wong et al., 2023; Zhang et al., 2021). Following which, the decline of DNA synthesis costs coming on new DNA synthesis

technologies including enzymatic approaches would enable the generation of sequences to be exogenously expressed in the hosts and can be optimised or refactored if required (Hoose et al., 2023). The combination of tools and parts for each host can be previously screened to facilitate this, and culturing on different media would reveal combinations that display successful and high production of the compound of interest (Hwang et al., 2021). To enable this approach, the diversity and quantity of the panel of hosts must first be increased. In parallel, the process of developing more hosts would also provide insights into which factors concerning genetic tools are pertinent to be incorporated into the screening process. Upon establishing this comprehensive framework, a diverse panel of hosts covering various branches of the phylogenetic tree will enhance our ability to tap into the extensive and diverse reservoir of natural products. This integration of automation, DNA sequencing, machine learning, and synthetic biology tools thereby positions us to unlock the full potential of *Streptomyces* to produce valuable bioactive compounds.

References

Abdelmohsen, U.R., Grkovic, T., Balasubramanian, S., Kamel, M.S., Quinn, R.J., Hentschel, U., 2015. Elicitation of secondary metabolism in actinomycetes. Biotechnology Advances 33, 798–811. https://doi.org/10.1016/j.biotechadv.2015.06.003

Ahmed, Y., 2019. Engineering of Streptomyces albus J1074 and Streptomyces lividans TK24 for natural products production. undefined.

Ahmed, Y., Rebets, Y., Estévez, M.R., Zapp, J., Myronovskyi, M., Luzhetskyy, A., 2020. Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters. Microb Cell Fact 19, 5. https://doi.org/10.1186/s12934-020-1277-8

Alam, K., Mazumder, A., Sikdar, S., Zhao, Y.-M., Hao, J., Song, C., Wang, Y., Sarkar, R., Islam, S., Zhang, Y., Li, A., 2022. Streptomyces: The biofactory of secondary metabolites. Frontiers in Microbiology 13.

Alberti, F., Corre, C., 2019. Editing streptomycete genomes in the CRISPR/Cas9 age. Natural Product Reports 36, 1237–1248. https://doi.org/10.1039/C8NP00081F

Atanasov, A.G., Zotchev, S.B., Dirsch, V.M., Supuran, C.T., 2021. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 20, 200–216. https://doi.org/10.1038/s41573-020-00114-z

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: Rapid Annotations using Subsystems Technology. BMC Genomics 9, 75. https://doi.org/10.1186/1471-2164-9-75

Badalamenti, J.P., Erickson, J.D., Salomon, C.E., 2016. Complete Genome Sequence of Streptomyces albus SM254, a Potent Antagonist of Bat White-Nose Syndrome Pathogen Pseudogymnoascus destructans. Genome Announc 4, e00290-16. https://doi.org/10.1128/genomeA.00290-16

Baltz, R.H., 2019. Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. Journal of Industrial Microbiology and Biotechnology 46, 281–299. https://doi.org/10.1007/s10295-018-2115-4

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

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

Baltz, R.H., 2008. Renaissance in antibacterial discovery from actinomycetes. Current Opinion in Pharmacology, Anti-infectives/New technologies 8, 557–563. https://doi.org/10.1016/j.coph.2008.04.008

Basheer, C., Obbard, J.P., Lee, H.K., 2003. Persistent Organic Pollutants in Singapore's Coastal Marine Environment: Part II, Sediments. Water, Air, & Soil Pollution 149, 315–325. https://doi.org/10.1023/A:1025673517831

Bérdy, J., 2012. Thoughts and facts about antibiotics: Where we are now and where we are heading. J Antibiot 65, 385–395. https://doi.org/10.1038/ja.2012.27
Bilyk, B., Luzhetskyy, A., 2014. Unusual site-specific DNA integration into the highly active pseudo-attB of the Streptomyces albus J1074 genome. Appl Microbiol Biotechnol 98, 5095–5104. https://doi.org/10.1007/s00253-014-5605-y
Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Medema, M.H., Weber, T., 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51, W46–W50. https://doi.org/10.1093/nar/gkad344
Blount, B.A., Weenink, T., Vasylechko, S., Ellis, T., 2012. Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. PLOS ONE 7, e33279. https://doi.org/10.1371/journal.pone.0033279
Bode, H.B., Bethe, B., Höfs, R., Zeeck, A., 2002. Big effects from small changes: possible ways to explore nature's chemical diversity. Chembiochem 3, 619–627. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9
Bu, Q.-T., Li, Y.-P., Xie, H., Li, J.-F., Lv, Z.-Y., Su, Y.-T., Li, Y.-Q., 2021. Rational engineering strategies for achieving high-yield, high-quality and high-stability of natural product production in actinomycetes. Metabolic Engineering 67, 198–215. https://doi.org/10.1016/j.ymben.2021.06.003
Bu, Q.-T., Yu, P., Wang, J., Li, Z.-Y., Chen, X.-A., Mao, X.-M., Li, Y.-Q., 2019. Rational construction of genome-reduced and high-efficient industrial Streptomyces chassis based on multiple comparative genomic approaches. Microb Cell Fact 18, 16. https://doi.org/10.1186/s12934-019-1055-7
Capecchi, A., Reymond, J.-L., 2021. Classifying natural products from plants, fungi or bacteria using the COCONUT database and machine learning. Journal of Cheminformatics 13, 82. https://doi.org/10.1186/s13321-021-00559-3
Carver, T., Thomson, N., Bleasby, A., Berriman, M., Parkhill, J., 2009. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25, 119–120. https://doi.org/10.1093/bioinformatics/btn578
Challis, G.L., 2008. Genome Mining for Novel Natural Product Discovery. J. Med. Chem. 51, 2618–2628. https://doi.org/10.1021/jm700948z
Chater, K.F., Biró, S., Lee, K.J., Palmer, T., Schrempf, H., 2010. The complex extracellular biology of Streptomyces. FEMS Microbiology Reviews 34, 171–198. https://doi.org/10.1111/j.1574-6976.2009.00206.x
Clardy, J., Fischbach, M.A., Walsh, C.T., 2006. New antibiotics from bacterial natural products. Nat Biotechnol 24, 1541–1550. https://doi.org/10.1038/nbt1266
Craney, A., Ozimok, C., Pimentel-Elardo, S.M., Capretta, A., Nodwell, J.R., 2012. Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. Chem Biol 19, 1020–1027. https://doi.org/10.1016/j.chembiol.2012.06.013
Culp, E.J., Yim, G., Waglechner, N., Wang, W., Pawlowski, A.C., Wright, G.D., 2019. Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. Nat Biotechnol 37, 1149–1154. https://doi.org/10.1038/s41587-019-0241-9

da Rosa, R., Schenkel, E.P., Campos Bernardes, L.S., 2020. Semisynthetic and newly designed derivatives based on natural chemical scaffolds: moving beyond natural products to fight Trypanosoma cruzi. Phytochem Rev 19, 105–122. https://doi.org/10.1007/s11101-020-09659-8

Davison, E.K., Brimble, M.A., 2019. Natural product derived privileged scaffolds in drug discovery. Current Opinion in Chemical Biology, Synthetic Biology • Synthetic Biomolecules 52, 1–8. https://doi.org/10.1016/j.cbpa.2018.12.007

Dias, D.A., Urban, S., Roessner, U., 2012. A Historical Overview of Natural Products in Drug Discovery. Metabolites 2, 303–336. https://doi.org/10.3390/metabo2020303

Donald, L., Pipite, A., Subramani, R., Owen, J., Keyzers, R.A., Taufa, T., 2022. Streptomyces: Still the Biggest Producer of New Natural Secondary Metabolites, a Current Perspective. Microbiology Research 13, 418–465. https://doi.org/10.3390/microbiolres13030031

Essarioui, A., Kistler, H.C., Kinkel, L.L., 2016. Nutrient use preferences among soil Streptomyces suggest greater resource competition in monoculture than polyculture plant communities. Plant Soil 409, 329–343. https://doi.org/10.1007/s11104-016-2968-0

Gao, Y., Zhao, Y., He, X., Deng, Z., Jiang, M., 2021. Challenges of functional expression of complex polyketide biosynthetic gene clusters. Current Opinion in Biotechnology, Chemical Biotechnology ● Pharmaceutical Biotechnology 69, 103–111. https://doi.org/10.1016/j.copbio.2020.12.007

García-Gutiérrez, C., Aparicio, T., Torres-Sánchez, L., Martínez-García, E., de Lorenzo, V., Villar, C.J., Lombó, F., 2020. Multifunctional SEVA shuttle vectors for actinomycetes and Gram-negative bacteria. Microbiologyopen 9, e1024. https://doi.org/10.1002/mbo3.1024

Garrido-Sanz, D., Manzano, J., Martín, M., Redondo-Nieto, M., Rivilla, R., 2018. Metagenomic Analysis of a Biphenyl-Degrading Soil Bacterial Consortium Reveals the Metabolic Roles of Specific Populations. Frontiers in Microbiology 9.

Gomez-Escribano, J.P., Alt, S., Bibb, M.J., 2016. Next Generation Sequencing of Actinobacteria for the Discovery of Novel Natural Products. Marine Drugs 14, 78. https://doi.org/10.3390/md14040078

Gomez-Escribano, J.P., Bibb, M.J., 2014. Heterologous expression of natural product biosynthetic gene clusters in Streptomyces coelicolor: from genome mining to manipulation of biosynthetic pathways. Journal of Industrial Microbiology and Biotechnology 41, 425–431. https://doi.org/10.1007/s10295-013-1348-5

Gomez-Escribano, J.P., Bibb, M.J., 2011. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters: *Streptomyces* host for heterologous expression of gene clusters. Microbial Biotechnology 4, 207–215. https://doi.org/10.1111/j.1751-7915.2010.00219.x

Gren, T., Whitford, C.M., Mohite, O.S., Jørgensen, T.S., Kontou, E.E., Nielsen, J.B., Lee, S.Y., Weber, T., 2021. Characterization and engineering of Streptomyces griseofuscus DSM 40191 as a potential host for heterologous expression of biosynthetic gene clusters. Sci Rep 11, 18301. https://doi.org/10.1038/s41598-021-97571-2

Hamed, M.B., Anné, J., Karamanou, S., Economou, A., 2018. Streptomyces protein secretion and its application in biotechnology. FEMS Microbiology Letters 365, fny250. https://doi.org/10.1093/femsle/fny250

Heng, E., Tan, L.L., Tay, D.W.P., Lim, Y.H., Yang, L.-K., Seow, D.C.S., Leong, C.Y., Ng, V., Ng, S.B., Kanagasundaram, Y., Wong, F.T., Koduru, L., 2023. Cost-effective hybrid long-short read assembly delineates alternative GC-rich Streptomyces hosts for natural product discovery. Synthetic and Systems Biotechnology 8, 253–261. https://doi.org/10.1016/j.synbio.2023.03.001

Hitch, T.C.A., Riedel, T., Oren, A., Overmann, J., Lawley, T.D., Clavel, T., 2021. Automated analysis of genomic sequences facilitates high-throughput and comprehensive description of bacteria. ISME COMMUN. 1, 1–16. https://doi.org/10.1038/s43705-021-00017-z

Hoff, G., Bertrand, C., Piotrowski, E., Thibessard, A., Leblond, P., 2018. Genome plasticity is governed by double strand break DNA repair in Streptomyces. Sci Rep 8, 5272. https://doi.org/10.1038/s41598-018-23622-w

Hoff, H., Drautz, H., Fiedler, H.P., Zähner, H., Schultz, J.E., Keller-Schierlein, W., Philipps, S., Ritzau, M., Zeeck, A., 1992. Metabolic products of microorganisms. 261. Obscurolides, a novel class of phosphodiesterase inhibitors from streptomyces. I. Production, isolation, structural elucidation and biological activity of obscurolides A1 to A4. J Antibiot (Tokyo) 45, 1096–1107. https://doi.org/10.7164/antibiotics.45.1096

Hoose, A., Vellacott, R., Storch, M., Freemont, P.S., Ryadnov, M.G., 2023. DNA synthesis technologies to close the gene writing gap. Nat Rev Chem 7, 144–161. https://doi.org/10.1038/s41570-022-00456-9

Hopwood, D.A., 1999. Forty years of genetics with Streptomyces: from in vivo through in vitro to in silico. Microbiology 145, 2183–2202. https://doi.org/10.1099/00221287-145-9-2183

Hopwood, D.A., Kieser, T., Wright, H.M., Bibb, M.J., 1983. Plasmids, Recombination and Chromosome Mapping in Streptomyces lividans 66. Microbiology 129, 2257–2269. https://doi.org/10.1099/00221287-129-7-2257

Hutchings, M.I., Truman, A.W., Wilkinson, B., 2019. Antibiotics: past, present and future. Current Opinion in Microbiology, Antimicrobials 51, 72–80. https://doi.org/10.1016/j.mib.2019.10.008

Hwang, S., Joung, C., Kim, W., Palsson, B., Cho, B.-K., 2023. Recent advances in non-model bacterial chassis construction. Current Opinion in Systems Biology 100471. https://doi.org/10.1016/j.coisb.2023.100471

Hwang, S., Lee, Y., Kim, J.H., Kim, G., Kim, H., Kim, W., Cho, S., Palsson, B.O., Cho, B.-K., 2021. Streptomyces as Microbial Chassis for Heterologous Protein Expression. Front. Bioeng. Biotechnol. 9, 804295. https://doi.org/10.3389/fbioe.2021.804295

J. Moore, S., Lai, H.-E., Li, J., S. Freemont, P., 2023. Streptomyces cell-free systems for natural product discovery and engineering. Natural Product Reports 40, 228–236. https://doi.org/10.1039/D2NP00057A

Johnston, C.D., Cotton, S.L., Rittling, S.R., Starr, J.R., Borisy, G.G., Dewhirst, F.E., Lemon, K.P., 2019. Systematic evasion of the restriction-modification barrier in bacteria. Proceedings of the National Academy of Sciences 116, 11454–11459. https://doi.org/10.1073/pnas.1820256116

Kang, H.-S., Kim, E.-S., 2021. Recent advances in heterologous expression of natural product biosynthetic gene clusters in Streptomyces hosts. Current Opinion in

Biotechnology, Chemical Biotechnology ● Pharmaceutical Biotechnology 69, 118–127. https://doi.org/10.1016/j.copbio.2020.12.016
Katz, L., Baltz, R.H., 2016. Natural product discovery: past, present, and future. Journal of Industrial Microbiology and Biotechnology 43, 155–176. https://doi.org/10.1007/s10295-015-1723-5
Ke, J., Yoshikuni, Y., 2020. Multi-chassis engineering for heterologous production of microbial natural products. Current Opinion in Biotechnology, Energy Biotechnologyy ● Environmental Biotechnology 62, 88–97. https://doi.org/10.1016/j.copbio.2019.09.005
Khalid, A., Takagi, H., Panthee, S., Muroi, M., Chappell, J., Osada, H., Takahashi, S., 2017. Development of a Terpenoid-Production Platform in Streptomyces reveromyceticus SN-593. ACS Synth. Biol. 6, 2339–2349. https://doi.org/10.1021/acssynbio.7b00249
Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L., 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37, 907–915. https://doi.org/10.1038/s41587-019-0201-4
Kim, E.J., Yang, I., Yoon, Y.J., 2015. Developing Streptomyces venezuelae as a cell factory for the production of small molecules used in drug discovery. Arch. Pharm. Res. 38, 1606–1616. https://doi.org/10.1007/s12272-015-0638-z
Kim, J.H., Komatsu, M., Shin-ya, K., Omura, S., Ikeda, H., 2018. Distribution and functional analysis of the phosphopantetheinyl transferase superfamily in Actinomycetales microorganisms. Proceedings of the National Academy of Sciences 115, 6828–6833. https://doi.org/10.1073/pnas.1800715115
Ko, Y.-S., Woong Kim, J., An Lee, J., Han, T., Bae Kim, G., Eum Park, J., Yup Lee, S., 2020. Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. Chemical Society Reviews 49, 4615–4636. https://doi.org/10.1039/D0CS00155D
Komatsu, M., Komatsu, K., Koiwai, H., Yamada, Y., Kozone, I., Izumikawa, M., Hashimoto, J., Takagi, M., Omura, S., Shin-ya, K., Cane, D.E., Ikeda, H., 2013. Engineered Streptomyces avermitilis Host for Heterologous Expression of Biosynthetic Gene Cluster for Secondary Metabolites. ACS Synth. Biol. 2, 384–396. https://doi.org/10.1021/sb3001003
Komatsu, M., Uchiyama, T., Ōmura, S., Cane, D.E., Ikeda, H., 2010. Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism. Proceedings of the National Academy of Sciences 107, 2646–2651. https://doi.org/10.1073/pnas.0914833107
Krause, J., Handayani, I., Blin, K., Kulik, A., Mast, Y., 2020. Disclosing the Potential of the SARP-Type Regulator PapR2 for the Activation of Antibiotic Gene Clusters in Streptomycetes. Frontiers in Microbiology 11.
Kudo, K., Hashimoto, T., Hashimoto, J., Kozone, I., Kagaya, N., Ueoka, R., Nishimura, T., Komatsu, M., Suenaga, H., Ikeda, H., Shin-ya, K., 2020. In vitro Cas9-assisted editing of modular polyketide synthase genes to produce desired natural product derivatives. Nat Commun 11, 4022. https://doi.org/10.1038/s41467-020-17769-2
Kumano, T., Tomita, T., Nishiyama, M., Kuzuyama, T., 2010. Functional Characterization of the Promiscuous Prenyltransferase Responsible for Furaquinocin Biosynthesis: IDENTIFICATION OF A PHYSIOLOGICAL POLYKETIDE SUBSTRATE AND ITS

PRENYLATED REACTION PRODUCTS*. Journal of Biological Chemistry 285, 39663–39671. https://doi.org/10.1074/jbc.M110.153957
Kumar, S., Abedin, M.M., Singh, A.K., Das, S., 2020. Role of Phenolic Compounds in Plant-Defensive Mechanisms, in: Plant Phenolics in Sustainable Agriculture. Springer, Singapore, pp. 517–532. https://doi.org/10.1007/978-981-15-4890-1_22
Kurumbang, N.P., Liou, K., Sohng, J.K., 2011. Biosynthesis of Ribostamycin Derivatives by Reconstitution and Heterologous Expression of Required Gene Sets. Appl Biochem Biotechnol 163, 373–382. https://doi.org/10.1007/s12010-010-9045-6
Lacey, H.J., Rutledge, P.J., 2022. Recently Discovered Secondary Metabolites from Streptomyces Species. Molecules 27, 887. https://doi.org/10.3390/molecules27030887
LeBlanc, N., Charles, T.C., 2022. Bacterial genome reductions: Tools, applications, and challenges. Frontiers in Genome Editing 4.
Lee, N., Hwang, S., Kim, J., Cho, S., Palsson, B., Cho, B.-K., 2020. Mini review: Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in Streptomyces. Computational and Structural Biotechnology Journal 18, 1548–1556. https://doi.org/10.1016/j.csbj.2020.06.024
Lee, N., Hwang, S., Kim, W., Lee, Y., Kim, J.H., Cho, S., Kim, H.U., Yoon, Y.J., Oh, M.-K., Palsson, B.O., Cho, B.-K., 2021. Systems and synthetic biology to elucidate secondary metabolite biosynthetic gene clusters encoded in Streptomyces genomes. Nat. Prod. Rep. 38, 1330–1361. https://doi.org/10.1039/D0NP00071J
Lee, N., Hwang, S., Lee, Y., Cho, S., Palsson, B., Cho, B.-K., 2019. Synthetic Biology Tools for Novel Secondary Metabolite Discovery in Streptomyces 29, 667–686. https://doi.org/10.4014/jmb.1904.04015
Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research 49, W293–W296. https://doi.org/10.1093/nar/gkab301
Li, F., Wang, Y., Li, D., Chen, Y., Dou, Q.P., 2019. Are we seeing a resurgence in the use of natural products for new drug discovery? Expert Opinion on Drug Discovery 14, 417–420. https://doi.org/10.1080/17460441.2019.1582639
Li, S., Li, Z., Pang, S., Xiang, W., Wang, W., 2021. Coordinating precursor supply for pharmaceutical polyketide production in Streptomyces. Current Opinion in Biotechnology, Chemical Biotechnology ● Pharmaceutical Biotechnology 69, 26–34. https://doi.org/10.1016/j.copbio.2020.11.006
Li, Y., Rebuffat, S., 2020. The manifold roles of microbial ribosomal peptide–based natural products in physiology and ecology. Journal of Biological Chemistry 295, 34–54. https://doi.org/10.1074/jbc.REV119.006545
Liao, C.-H., Yao, L., Xu, Y., Liu, W.-B., Zhou, Y., Ye, B.-C., 2015. Nitrogen regulator GlnR controls uptake and utilization of non-phosphotransferase-system carbon sources in actinomycetes. Proceedings of the National Academy of Sciences 112, 15630–15635. https://doi.org/10.1073/pnas.1508465112
Liu, J., Wang, X., Dai, G., Zhang, Y., Bian, X., 2022. Microbial chassis engineering drives heterologous production of complex secondary metabolites. Biotechnology Advances 59, 107966. https://doi.org/10.1016/j.biotechadv.2022.107966
Liu, Q., Schumacher, J., Wan, X., Lou, C., Wang, B., 2018. Orthogonality and Burdens of Heterologous AND Gate Gene Circuits in E. coli. ACS Synth. Biol. 7, 553–564. https://doi.org/10.1021/acssynbio.7b00328

Liu, Q., Xiao, L., Zhou, Y., Deng, K., Tan, G., Han, Y., Liu, X., Deng, Z., Liu, T., 2016. Development of Streptomyces sp. FR-008 as an emerging chassis. Synthetic and Systems Biotechnology 1, 207–214. https://doi.org/10.1016/j.synbio.2016.07.002
Liu, R., Deng, Z., Liu, T., 2018. Streptomyces species: Ideal chassis for natural product discovery and overproduction. Metabolic Engineering, Metabolic Engineering Host Organism Special Issue 50, 74–84. https://doi.org/10.1016/j.ymben.2018.05.015
Liu, Z., Zhao, Y., Huang, C., Luo, Y., 2021. Recent Advances in Silent Gene Cluster Activation in Streptomyces. Frontiers in Bioengineering and Biotechnology 9.
Low, Z.J., Pang, L.M., Ding, Y., Cheang, Q.W., Le Mai Hoang, K., Thi Tran, H., Li, J., Liu, X.-W., Kanagasundaram, Y., Yang, L., Liang, Z.-X., 2018. Identification of a biosynthetic gene cluster for the polyene macrolactam sceliphrolactam in a Streptomyces strain isolated from mangrove sediment. Sci Rep 8, 1594. https://doi.org/10.1038/s41598-018-20018-8
Luo, H., Vong, C.T., Chen, H., Gao, Y., Lyu, P., Qiu, L., Zhao, M., Liu, Q., Cheng, Z., Zou, J., Yao, P., Gao, C., Wei, J., Ung, C.O.L., Wang, S., Zhong, Z., Wang, Y., 2019. Naturally occurring anti-cancer compounds: shining from Chinese herbal medicine. Chin Med 14, 48. https://doi.org/10.1186/s13020-019-0270-9
Luo, Y., Zhang, L., Barton, K.W., Zhao, H., 2015. Systematic Identification of a Panel of Strong Constitutive Promoters from Streptomyces albus. ACS Synth. Biol. 4, 1001–1010. https://doi.org/10.1021/acssynbio.5b00016
Ma, G.-L., Candra, H., Pang, L.M., Xiong, J., Ding, Y., Tran, H.T., Low, Z.J., Ye, H., Liu, M., Zheng, J., Fang, M., Cao, B., Liang, Z.-X., 2022. Biosynthesis of Tasikamides via Pathway Coupling and Diazonium-Mediated Hydrazone Formation. J. Am. Chem. Soc. 144, 1622–1633. https://doi.org/10.1021/jacs.1c10369
Manteca, Á., Yagüe, P., 2018. Streptomyces Differentiation in Liquid Cultures as a Trigger of Secondary Metabolism. Antibiotics 7, 41. https://doi.org/10.3390/antibiotics7020041
Matsutani, M., Yakushi, T., 2018. Pyrroloquinoline quinone-dependent dehydrogenases of acetic acid bacteria. Appl Microbiol Biotechnol 102, 9531–9540. https://doi.org/10.1007/s00253-018-9360-3
Meier-Kolthoff, J.P., Göker, M., 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10, 2182. https://doi.org/10.1038/s41467-019-10210-3
Mo, J., Wang, S., Zhang, W., Li, C., Deng, Z., Zhang, L., Qu, X., 2019. Efficient editing DNA regions with high sequence identity in actinomycetal genomes by a CRISPR-Cas9 system. Synthetic and Systems Biotechnology 4, 86–91. https://doi.org/10.1016/j.synbio.2019.02.004
Myronovskyi, M., Luzhetskyy, A., 2019. Heterologous production of small molecules in the optimized Streptomyces hosts. Natural Product Reports 36, 1281–1294. https://doi.org/10.1039/C9NP00023B
Myronovskyi, M., Luzhetskyy, A., 2016. Native and engineered promoters in natural product discovery. Nat Prod Rep 33, 1006–1019. https://doi.org/10.1039/c6np00002a
Myronovskyi, M., Rosenkränzer, B., Nadmid, S., Pujic, P., Normand, P., Luzhetskyy, A., 2018. Generation of a cluster-free Streptomyces albus chassis strains for

improved heterologous expression of secondary metabolite clusters. Metabolic Engineering 49, 316–324. https://doi.org/10.1016/j.ymben.2018.09.004
Nepal, K.K., Wang, G., 2019. Streptomycetes: Surrogate hosts for the genetic manipulation of biosynthetic gene clusters and production of natural products. Biotechnology Advances 37, 1–20. https://doi.org/10.1016/j.biotechadv.2018.10.003
Newman, D.J., Cragg, G.M., 2020. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. J. Nat. Prod. 83, 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285
Ohno, S., Katsuyama, Y., Tajima, Y., Izumikawa, M., Takagi, M., Fujie, M., Satoh, N., Shin-Ya, K., Ohnishi, Y., 2015. Identification and Characterization of the Streptazone E Biosynthetic Gene Cluster in Streptomyces sp. MSC090213JE08. Chembiochem 16, 2385–2391. https://doi.org/10.1002/cbic.201500317
Otani, H., Udwary, D.W., Mouncey, N.J., 2022. Comparative and pangenomic analysis of the genus Streptomyces. Sci Rep 12, 18909. https://doi.org/10.1038/s41598-022-21731-1
Owen, J.G., Charlop-Powers, Z., Smith, A.G., Ternei, M.A., Calle, P.Y., Reddy, B.V.B., Montiel, D., Brady, S.F., 2015. Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. Proc Natl Acad Sci U S A 112, 4221–4226. https://doi.org/10.1073/pnas.1501124112
Pan, R., Bai, X., Chen, J., Zhang, H., Wang, H., 2019. Exploring Structural Diversity of Microbe Secondary Metabolites Using OSMAC Strategy: A Literature Review. Frontiers in Microbiology 10.
Park, J.W., Park, S.R., Nepal, K.K., Han, A.R., Ban, Y.H., Yoo, Y.J., Kim, E.J., Kim, E.M., Kim, D., Sohng, J.K., Yoon, Y.J., 2011. Discovery of parallel pathways of kanamycin biosynthesis allows antibiotic manipulation. Nat Chem Biol 7, 843–852. https://doi.org/10.1038/nchembio.671
Park, S.R., Yoon, J.A., Paik, J.H., Park, J.W., Jung, W.S., Ban, Y.-H., Kim, E.J., Yoo, Y.J., Han, A.R., Yoon, Y.J., 2009. Engineering of plant-specific phenylpropanoids biosynthesis in Streptomyces venezuelae. Journal of Biotechnology 141, 181–188. https://doi.org/10.1016/j.jbiotec.2009.03.013
Parra, J., Beaton, A., Seipke, R.F., Wilkinson, B., Hutchings, M.I., Duncan, K.R., 2023. Antibiotics from rare actinomycetes, beyond the genus Streptomyces. Current Opinion in Microbiology 76, 102385. https://doi.org/10.1016/j.mib.2023.102385
Payne, D.J., Gwynn, M.N., Holmes, D.J., Pompliano, D.L., 2007. Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat Rev Drug Discov 6, 29–40. https://doi.org/10.1038/nrd2201
Peng, Q., Gao, G., Lü, J., Long, Q., Chen, X., Zhang, F., Xu, M., Liu, K., Wang, Y., Deng, Z., Li, Z., Tao, M., 2018. Engineered Streptomyces lividans Strains for Optimal Identification and Expression of Cryptic Biosynthetic Gene Clusters. Frontiers in Microbiology 9.
Pham, J.V., Yilma, M.A., Feliz, A., Majid, M.T., Maffetone, N., Walker, J.R., Kim, E., Cho, H.J., Reynolds, J.M., Song, M.C., Park, S.R., Yoon, Y.J., 2019. A Review of the Microbial Production of Bioactive Natural Products and Biologics. Frontiers in Microbiology 10.

Pinedo-Rivilla, C., Aleu, J., Durán-Patrón, R., 2022. Cryptic Metabolites from Marine-Derived Microorganisms Using OSMAC and Epigenetic Approaches. Marine Drugs 20, 84. https://doi.org/10.3390/md20020084

Popa, C., Favier, L., Dinica, R., Semrany, S., Djelal, H., Amrane, A., Bahrim, G., 2014. Potential of newly isolated wild Streptomyces strains as agents for the biodegradation of a recalcitrant pharmaceutical, carbamazepine. Environmental Technology 35, 3082–3091. https://doi.org/10.1080/09593330.2014.931468

Puder, C., Krastel, P., Zeeck, A., 2000. Streptazones A, B1, B2, C, and D: New Piperidine Alkaloids from Streptomycetes. J. Nat. Prod. 63, 1258–1260. https://doi.org/10.1021/np0001373

Qiu, S., Yang, B., Li, Z., Li, S., Yan, H., Xin, Z., Liu, J., Zhao, X., Zhang, L., Xiang, W., Wang, W., 2024. Building a highly efficient Streptomyces super-chassis for secondary metabolite production by reprogramming naturally-evolved multifaceted shifts. Metabolic Engineering 81, 210–226. https://doi.org/10.1016/j.ymben.2023.12.007

Quinn, G.A., Banat, A.M., Abdelhameed, A.M., Banat, I.M., 2020. Streptomyces from traditional medicine: sources of new innovations in antibiotic discovery. J Med Microbiol 69, 1040–1048. https://doi.org/10.1099/jmm.0.001232

Rajwani, R., Ohlemacher, S.I., Zhao, G., Liu, H.-B., Bewley, C.A., 2021. Genome-Guided Discovery of Natural Products through Multiplexed Low-Coverage Whole-Genome Sequencing of Soil Actinomycetes on Oxford Nanopore Flongle. mSystems 6, e01020-21. https://doi.org/10.1128/mSystems.01020-21

Rateb, M.E., Houssen, W.E., Harrison, W.T.A., Deng, H., Okoro, C.K., Asenjo, J.A., Andrews, B.A., Bull, A.T., Goodfellow, M., Ebel, R., Jaspars, M., 2011. Diverse Metabolic Profiles of a Streptomyces Strain Isolated from a Hyper-arid Environment. J. Nat. Prod. 74, 1965–1971. https://doi.org/10.1021/np200470u

Rateb, M.E., Yang, D., Vodanovic-Jankovic, S., Yu, Z., Kron, M.A., Shen, B., 2015. Adipostatins A–D from Streptomyces sp. 4875 inhibiting Brugia malayi asparaginyl-tRNA synthetase and killing adult Brugia malayi parasites. J Antibiot 68, 540–542. https://doi.org/10.1038/ja.2015.22

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat Biotechnol 29, 24–26. https://doi.org/10.1038/nbt.1754

Romano, S., Jackson, S.A., Patry, S., Dobson, A.D.W., 2018. Extending the "One Strain Many Compounds" (OSMAC) Principle to Marine Microorganisms. Marine Drugs 16, 244. https://doi.org/10.3390/md16070244

Sánchez, S., Chávez, A., Forero, A., García-Huante, Y., Romero, A., Sánchez, M., Rocha, D., Sánchez, B., Ávalos, M., Guzmán-Trampe, S., Rodríguez-Sanoja, R., Langley, E., Ruiz, B., 2010. Carbon source regulation of antibiotic production. J Antibiot 63, 442–459. https://doi.org/10.1038/ja.2010.78

Schlatter, D.C., Kinkel, L.L., 2014. Global biogeography of Streptomyces antibiotic inhibition, resistance, and resource use. FEMS Microbiology Ecology 88, 386–397. https://doi.org/10.1111/1574-6941.12307

Schlimpert, S., Elliot, M.A., 2023. The Best of Both Worlds—Streptomyces coelicolor and Streptomyces venezuelae as Model Species for Studying Antibiotic Production and Bacterial Multicellular Development. Journal of Bacteriology 205, e00153-23. https://doi.org/10.1128/jb.00153-23

Schmidt, R., Ulanova, D., Wick, L.Y., Bode, H.B., Garbeva, P., 2019. Microbe-driven chemical ecology: past, present and future. ISME J 13, 2656–2663. https://doi.org/10.1038/s41396-019-0469-x

Scott Zarins-Tutt, J., Triscari Barberi, T., Gao, H., Mearns-Spragg, A., Zhang, L., J. Newman, D., Miriam Goss, R.J., 2016. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. Natural Product Reports 33, 54–72. https://doi.org/10.1039/C5NP00111K

Shao, Z., Rao, G., Li, C., Abil, Z., Luo, Y., Zhao, H., 2013. Refactoring the Silent Spectinabilin Gene Cluster Using a Plug-and-Play Scaffold. ACS Synth Biol 2, 662–669. https://doi.org/10.1021/sb400058n

Sherwood, E.J., Hesketh, A.R., Bibb, M.J., 2013. Cloning and Analysis of the Planosporicin Lantibiotic Biosynthetic Gene Cluster of Planomonospora alba. J Bacteriol 195, 2309–2321. https://doi.org/10.1128/JB.02291-12

Sorokina, M., Steinbeck, C., 2020. Review on natural products databases: where to find data in 2020. J Cheminform 12, 20. https://doi.org/10.1186/s13321-020-00424-9

Sosio, M., Giusino, F., Cappellano, C., Bossi, E., Puglia, A.M., Donadio, S., 2000. Artificial chromosomes for antibiotic-producing actinomycetes. Nat Biotechnol 18, 343–345. https://doi.org/10.1038/73810

Subramani, R., Sipkema, D., 2019. Marine Rare Actinomycetes: A Promising Source of Structurally Diverse and Unique Novel Natural Products. Marine Drugs 17, 249. https://doi.org/10.3390/md17050249

Sun, J., Kelemen, G.H., Fernández-Abalos, J.M., Bibb, M.J., 1999. Green fluorescent protein as a reporter for spatial and temporal gene expression in Streptomyces coelicolor A3(2)This paper is dedicated to the memory of Kathy Kendrick, whose devotion to understanding the biology of Streptomyces was unsurpassed. Microbiology 145, 2221–2227. https://doi.org/10.1099/00221287-145-9-2221

Swallah, M.S., Sun, H., Affoh, R., Fu, H., Yu, H., 2020. Antioxidant Potential Overviews of Secondary Metabolites (Polyphenols) in Fruits. International Journal of Food Science 2020, e9081686. https://doi.org/10.1155/2020/9081686

Świątek, M.A., Urem, M., Tenconi, E., Rigali, S., van Wezel, G.P., 2012. Engineering of N-acetylglucosamine metabolism for improved antibiotic production in Streptomyces coelicolor A3(2) and an unsuspected role of NagA in glucosamine metabolism. Bioengineered 3, 280–285. https://doi.org/10.4161/bioe.21371

Tao, W., Yang, A., Deng, Z., Sun, Y., 2018. CRISPR/Cas9-Based Editing of Streptomyces for Discovery, Characterization, and Production of Natural Products. Frontiers in Microbiology 9.

Tarasova, E.V., Luchnikova, N.A., Grishko, V.V., Ivshina, I.B., 2023. Actinomycetes as Producers of Biologically Active Terpenoids: Current Trends and Patents. Pharmaceuticals 16, 872. https://doi.org/10.3390/ph16060872

Tomm, H.A., Ucciferri, L., Ross, A.C., 2019. Advances in microbial culturing conditions to activate silent biosynthetic gene clusters for novel metabolite production. Journal of Industrial Microbiology and Biotechnology 46, 1381–1400. https://doi.org/10.1007/s10295-019-02198-y

Tong, Y., Charusanti, P., Zhang, L., Weber, T., Lee, S.Y., 2015. CRISPR-Cas9 Based Engineering of Actinomycetal Genomes. ACS Synth Biol 4, 1020–1029. https://doi.org/10.1021/acssynbio.5b00038

Tong, Y., Deng, Z., 2020. An aurora of natural products-based drug discovery is coming. Synthetic and Systems Biotechnology 5, 92–96. https://doi.org/10.1016/j.synbio.2020.05.003
Wang, Q., Xie, F., Tong, Y., Habisch, R., Yang, B., Zhang, L., Müller, R., Fu, C., 2020. Dual-function chromogenic screening-based CRISPR/Cas9 genome editing system for actinomycetes. Appl Microbiol Biotechnol 104, 225–239. https://doi.org/10.1007/s00253-019-10223-4
Wang, W., Li, X., Wang, J., Xiang, S., Feng, X., Yang, K., 2013. An Engineered Strong Promoter for Streptomycetes. Appl Environ Microbiol 79, 4484–4492. https://doi.org/10.1128/AEM.00985-13
Whitford, C.M., Cruz-Morales, P., Keasling, J.D., Weber, T., 2021. The Design-Build-Test-Learn cycle for metabolic engineering of Streptomycetes. Essays in Biochemistry 65, 261–275. https://doi.org/10.1042/EBC20200132
Wong, F., Zheng, E.J., Valeri, J.A., Donghia, N.M., Anahtar, M.N., Omori, S., Li, A., Cubillos-Ruiz, A., Krishnan, A., Jin, W., Manson, A.L., Friedrichs, J., Helbig, R., Hajian, B., Fiejtek, D.K., Wagner, F.F., Soutter, H.H., Earl, A.M., Stokes, J.M., Renner, L.D., Collins, J.J., 2023. Discovery of a structural class of antibiotics with explainable deep learning. Nature 1–9. https://doi.org/10.1038/s41586-023-06887-8
Wørmer, G.J., Poulsen, T.B., 2022. The [4.3.0] Piperidine Alkaloids: Architectures, Biology, Biosyntheses, and the Complete Details of the Asymmetric Syntheses of Streptazone A and Abikoviromycin. Synlett 33, 637–654. https://doi.org/10.1055/a-1688-0826
Wright, B.W., Molloy, M.P., Jaschke, P.R., 2022. Overlapping genes in natural and engineered genomes. Nat Rev Genet 23, 154–168. https://doi.org/10.1038/s41576-021-00417-w
Wu, C.-C., Liles, M., Kakirde, K., Ye, R., Wagner, M., Krerowicz, A., Staley, M., Jasinovica, S., Drinkwater, C., Godiska, R., Mead, D., 2012. Next-generation functional and structural soil metagenomics.
Xia, H., Li, X., Li, Z., Zhan, X., Mao, X., Li, Y., 2020. The Application of Regulatory Cascades in Streptomyces: Yield Enhancement and Metabolite Mining. Frontiers in Microbiology 11.
Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., Zhang, G., Gu, Y.Q., Coleman-Derr, D., Xia, Q., Wang, Y., 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res 47, W52–W58. https://doi.org/10.1093/nar/gkz333
Xu, M., Wang, Y., Zhao, Z., Gao, G., Huang, S.-X., Kang, Q., He, X., Lin, S., Pang, X., Deng, Z., Tao, M., 2016. Functional Genome Mining for Metabolites Encoded by Large Gene Clusters through Heterologous Expression of a Whole-Genome Bacterial Artificial Chromosome Library in Streptomyces spp. Applied and Environmental Microbiology 82, 5795–5805. https://doi.org/10.1128/AEM.01383-16
Xu, Z., Jakobi, K., Welzel, K., Hertweck, C., 2005. Biosynthesis of the Antitumor Agent Chartreusin Involves the Oxidative Rearrangement of an Anthracyclic Polyketide. Chemistry & Biology 12, 579–588. https://doi.org/10.1016/j.chembiol.2005.04.017
Xu, Z., Ji, L., Tang, W., Guo, L., Gao, C., Chen, X., Liu, J., Hu, G., Liu, L., 2022. Metabolic engineering of Streptomyces to enhance the synthesis of valuable natural

products. Engineering Microbiology 2, 100022. https://doi.org/10.1016/j.engmic.2022.100022
Yagüe, P., Lopez-Garcia, M.T., Rioseras, B., Sanchez, J., Manteca, A., 2012. New insights on the development of Streptomyces and their relationships with secondary metabolite production. Curr Trends Microbiol 8, 65–73.
Yagüe, P., Rodríguez-García, A., López-García, M.T., Martín, J.F., Rioseras, B., Sánchez, J., Manteca, A., 2013. Transcriptomic Analysis of Streptomyces coelicolor Differentiation in Solid Sporulating Cultures: First Compartmentalized and Second Multinucleated Mycelia Have Different and Distinctive Transcriptomes. PLOS ONE 8, e60665. https://doi.org/10.1371/journal.pone.0060665
Yanai, K., Murakami, T., Bibb, M., 2006. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of Streptomyces kanamyceticus. Proc Natl Acad Sci U S A 103, 9661–9666. https://doi.org/10.1073/pnas.0603251103
Yang, D., Zhu, X., Wu, X., Feng, Z., Huang, L., Shen, B., Xu, Z., 2011. Titer improvement of iso-migrastatin in selected heterologous Streptomyces hosts and related analysis of mRNA expression by quantitative RT–PCR. Appl Microbiol Biotechnol 89, 1709–1719. https://doi.org/10.1007/s00253-010-3025-1
Yang, M., Meng, F., Gu, W., Li, F., Tao, Y., Zhang, Z., Zhang, F., Yang, X., Li, J., Yu, J., 2020. Effects of Natural Products on Bacterial Communication and Network-Quorum Sensing. BioMed Research International 2020, e8638103. https://doi.org/10.1155/2020/8638103
Yang, P., Yang, J., Lin, T., Liu, Q., Yin, Y., Chen, D., Yang, S., 2023. Efficient Genome Editing in Most Staphylococcus aureus by Using the Restriction-Modification System Silent CRISPR-Cas9 Toolkit. ACS Synth. Biol. 12, 3340–3351. https://doi.org/10.1021/acssynbio.3c00339
Yang, Z., Liu, C., Wang, Y., Chen, Y., Li, Q., Zhang, Y., Chen, Q., Ju, J., Ma, J., 2022. MGCEP 1.0: A Genetic-Engineered Marine-Derived Chassis Cell for a Scaled Heterologous Expression Platform of Microbial Bioactive Metabolites. ACS Synth. Biol. 11, 3772–3784. https://doi.org/10.1021/acssynbio.2c00362
Yin, S., Li, Z., Wang, X., Wang, H., Jia, X., Ai, G., Bai, Z., Shi, M., Yuan, F., Liu, T., Wang, W., Yang, K., 2016. Heterologous expression of oxytetracycline biosynthetic gene cluster in Streptomyces venezuelae WVR2006 to improve production level and to alter fermentation process. Appl Microbiol Biotechnol 100, 10563–10572. https://doi.org/10.1007/s00253-016-7873-1
Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J., 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67, 1613–1617. https://doi.org/10.1099/ijsem.0.001755
Yue, X., Sheng, D., Zhuo, L., Li, Y.-Z., 2023. Genetic manipulation and tools in myxobacteria for the exploitation of secondary metabolism. Engineering Microbiology 3, 100075. https://doi.org/10.1016/j.engmic.2023.100075
Zabolotna, Y., Ertl, P., Horvath, D., Bonachera, F., Marcou, G., Varnek, A., 2021. NP Navigator: A New Look at the Natural Product Chemical Space. Molecular Informatics 40, 2100068. https://doi.org/10.1002/minf.202100068

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

Zhan, C., Lee, N., Lan, G., Dan, Q., Cowan, A., Wang, Z., Baidoo, E.E.K., Kakumanu, R., Luckie, B., Kuo, R.C., McCauley, J., Liu, Y., Valencia, L., Haushalter, R.W., Keasling, J.D., 2023. Improved polyketide production in C. glutamicum by preventing propionate-induced growth inhibition. Nat Metab 5, 1127–1140. https://doi.org/10.1038/s42255-023-00830-x

Zhang, B., Tian, W., Wang, S., Yan, X., Jia, X., Pierens, G.K., Chen, W., Ma, H., Deng, Z., Qu, X., 2017. Activation of Natural Products Biosynthetic Pathways via a Protein Modification Level Regulation. ACS Chem. Biol. 12, 1732–1736. https://doi.org/10.1021/acschembio.7b00225

Zhang, R., Li, X., Zhang, X., Qin, H., Xiao, W., 2021. Machine learning approaches for elucidating the biological effects of natural products. Nat. Prod. Rep. 38, 346–361. https://doi.org/10.1039/D0NP00043D

Zhao, Y., Li, G., Chen, Y., Lu, Y., 2020. Challenges and Advances in Genome Editing Technologies in Streptomyces. Biomolecules 10, 734. https://doi.org/10.3390/biom10050734

Zhu, D.-H., Nie, F.-H., Song, Q.-L., Wei, W., Zhang, M., Hu, Y., Lin, H.-Y., Kang, D.-J., Chen, Z.-B., Chen, J.-J., 2023. Isolation and genomic characterization of Klebsiella Lw3 with polychlorinated biphenyl degradability. Environmental Technology 44, 3656–3666. https://doi.org/10.1080/09593330.2022.2068381

Zong, G., Fu, J., Zhang, P., Zhang, W., Xu, Y., Cao, G., Zhang, R., 2022. Use of elicitors to enhance or activate the antibiotic production in streptomyces. Critical Reviews in Biotechnology 42, 1260–1283. https://doi.org/10.1080/07388551.2021.1987856

Appendix

Supplementary Tables

Table S1. Plasmids used in this study.

Plasmid	Description	Reference
pESAC13A	PAC-derived *E. coli-Streptomyces* shuttle vector	(Sosio et al., 2000)
pESAC13A-P46-BGC1	pESAC13A containing BGC1 from *Streptomyces tasikensis* P46	This work
pESAC13A-P46-BGC13	pESAC13A containing BGC13 from *Streptomyces tasikensis* P46	This work
pESAC13A-P46-BGC15	pESAC13A containing BGC15 from *Streptomyces tasikensis* P46	This work
pESAC13A-P46-BGC23	pESAC13A containing BGC23 from *Streptomyces tasikensis* P46	This work
pESAC13A-P46-BGC28	pESAC13A containing BGC28 from *Streptomyces tasikensis* P46	This work
pESAC13A-P46-BGC30	pESAC13A containing BGC30 from *Streptomyces tasikensis* P46	This work
pESAC13A-SD50-BGC7	pESAC13A containing BGC7 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC18	pESAC13A containing BGC18 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC23	pESAC13A containing BGC23 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC32	pESAC13A containing BGC32 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC38	pESAC13A containing BGC38 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC44	pESAC13A containing BGC44 from *Streptomyces sp.* SD50	This work
pESAC13A-SD50-BGC45	pESAC13A containing BGC45 from *Streptomyces sp.* SD50	This work
pIJ8630	EGFP reporter vector	(Sun et al., 1999)
pIJ8630-ermEp*	EGFP reporter vector under the control of ermEp*	This work
pIJ8630-SF14p	EGFP reporter vector under the control of SF14p	This work
pIJ8630-kasOp*	EGFP reporter vector under the control of kasOp*	This work
pIJ8630-gapdhp(EL)	EGFP reporter vector under the control of gapdhp(EL)	This work
pIJ8630-gapdhp(KR)	EGFP reporter vector under the control of gapdhp(KR)	This work

pIJ8630-fur1	Type III PKS *fur1* gene in pIJ8630 vector under the control of gapdhp(KR)	This work
pIJ12551	Constitutive expression vector under the control of ermEp*	(Sherwood et al., 2013)
pSEVA28c1	Replicative vector using pIJ101 origin	(García-Gutiérrez et al., 2020)
pIJ10702	Cosmid vector derived from SuperCos1	(Yanai et al., 2006)
pSMART-BAC-S	BAC-derived *E. coli-Streptomyces* shuttle vector	(Wu et al., 2012)
pCRISPR-Cas9	CRISPR-Cas9 vector for actinomycetes	(Tong et al., 2015)
pCRISPR-Cas9-BGC16	pCRISPR-Cas9 derivative for the knockout of BGC16	This work
pCRISPR-Cas9-BGC23a-ΦBT1-attB	pCRISPR-Cas9 derivative for the knockout of BGC23a and insertion of *ΦBT1-attB*	This work
pCRISPR-Cas9-BGC5P19-bldA	pCRISPR-Cas9 derivative for the knockout of BGC5's P19 gene and insertion of *bldA*	This work
pCRISPR-Cas9-BGC5P13-gpps	pCRISPR-Cas9 derivative for the knockout of BGC5's P13 gene and insertion of *gpps*	This work
pCRISPR-Cas9-BGC23b	pCRISPR-Cas9 derivative for the knockout of BGC23b	This work
pQS-idgS	CRISPR-Cas9 vector for actinomycetes	(Wang et al., 2020)
pQS-idgS-ΔBGC32b	pQS-idgS derivative for the knockout of BGC32b	This work
pQS-idgS-BGC32b::KasOp	pQS-idgS derivative for the insertion of KasOp* in front of the first gene of BGC32b	This work
pQS-idgS-ΔBGC33::ΔchaAB::sfp	pQS-idgS derivative for the knockout of BGC33 *ChaA* and *ChaB* and replacement with promiscuous pptase *sfp* gene and SD3 *phiC31 attB* site	This work

Table S2. Strains used in this study.

Strain	Description
E. coli TOP10	For cloning
E. coli ET12567/pUZ8002	For conjugative transfer of plasmids into *Streptomyces spp.*
Streptomyces sungeiensis SD3	Isolated from Sungei Buloh Wetland Reserve, Singapore (1°26'41.5"N, 103°43'19.3"E); wild type genotype
SD3_ΔBGC32b	SD3 strain containing the knockout of BGC32b
SD3::kasOP	SD3 strain containing the insertion of KasOp* in front of the first gene of BGC32b
SD3_ΔchaAB::sfp	SD3 strain containing the knockout of BGC33 *ChaA* and *ChaB* and replacement with promiscuous pptase *sfp* gene and SD3 *phiC31 attB* site
SD3_Δ240kb	SD3 strain containing a 240.6 kb deletion region

SD3::P46-BGC13	SD3 strain containing the integration of BGC13 from P46
SD3::P46-BGC23	SD3 strain containing the integration of BGC23 from P46
SD3::P46-BGC28	SD3 strain containing the integration of BGC28 from P46
S. coelicolor M1154	Streptomyces chassis strain
S. lividans TK24	Streptomyces chassis strain
S. albidoflavus MD102	Isolated from Sungei Buloh Wetland Reserve, Singapore (1°26'41.172"N, 103°43'36.12"E); wild type genotype
S. albidoflavus MD102Δ1	ΔC16
S. albidoflavus MD102Δ2I1	ΔC16 ΔC23a::ΦBT1-*attB*
S. albidoflavus MD102Δ3I2	ΔC16 ΔC23a::ΦBT1-*attB* ΔC5P19::*bldA*
S. albidoflavus MD102Δ3I3	ΔC16 ΔC23a::ΦBT1-*attB* ΔC5P19::*bldA* ΔC5P13::*gpps*
S. albidoflavus MD102SL01	ΔC16 ΔC23a::ΦBT1-*attB* ΔC5P19::*bldA* ΔC5P13::*gpps* ΔC23b ΔC23c ΔC1
S. albidoflavus MD102SL01::fur1	*S. albidoflavus* MD102SL01 with pIJ8630-fur1 for heterologous production of flaviolin

Table S3. Primers used in this study.

Primer name	Sequence (5'-3')	Function
BGC32b-KaSOP-Gibson-F	tgttcacattcgaaccgtctct	Amplify KasOp for Gibson assembly with BGC32b HR arms
BGC32b-KaSOP-Gibson-R	gcaaaacccctaggtcgtcg	
BGC32b-LHR-Gibson-F	ctcgtcgaaggcactagaggcctacgacctgttgctgcc	Amplify BGC32b LHR arms for Gibson assembly
BGC32b-LHR-Gibson-R	caaggagatggagcgcctg	
BGC32b-RHR-Gibson-F	gtgtccgatcaagctcgct	Amplify BGC32b RHR arms for Gibson assembly for promoter insertion
BGC32b-RHR-Gibson-R	ggtcgatccccgcatataggctgcgaagtgtgctgtgga	
BGC32b-RHR-del-Gibson-F	acaggcgctccatctccttgaccgtccacagcacacttc	Amplify BGC32b RHR arms for Gibson assembly for deletion
BGC32b-RHR-del-Gibson-R	ggtcgatccccgcatataggcatgaggccgaggtagatgc	
Fw C34 LHR pQS	ctcgtcgaaggcactagaggtcggtcgcacttttccaaca	Amplify BGC33 LHR for Gibson assembly
Rw C34 LHR	gacggttcgaatgtgaacagggagtcggcttggagatacg	
Rw sfpattB	taaactagtggtggagtacgcg	Amplify BGC33 LHR with sfp-attb for Gibson assembly
Fw C34 RHR	cgtactccaccactagtttaggaacaccccacaaccgaa	Amplify BGC33 RHR for Gibson assembly
Rw C34 RHR pQS	ggtcgatccccgcatataggcggtatcagttgtttcgcgg	
Fw Full sfpattb chk	taggtctgcaaggtcgtggt	Checking primer for BGC33 deleted/knock-in region size
Rw Full sfpattb chk	aactcgatgtcgcggatgtc	
Fw BGC33 ChaA chk	cgcccaaggaactgaggag	

Rw BGC33 ChaA chk	cccgtcatgtgataggcgtt	Checking primer for ChaA in BGC33
Fw C32b chk	ggcgtctacaccgacatca	Checking primer for BGC32b deleted region
Rw C32b chk	ctgcgaagtgtgctgtgga	
Fw C32b-KasOp chk	gtctacaccgacatcacggc	Checking primer for KasOp in BGC32b
Rw C32b-KasOp chk	tcagaccggaaagcacatcc	
pQS-32b-Fw	catggttattcggacggacgactgggttttagag	Amplify guide region for BGC32b for use in pQS-idgS
pQS-32b-Rw	ctagctctaaaacccagtcgtccgtccgaataac	
pQS-33-Fw	catggcctatcacatgacgggtctggttttagag	Amplify guide region for BGC33 for use in pQS-idgS
pQS-33-Rw	ctagctctaaaaccagacccgtcatgtgataggc	

Table S4. Conjugation efficiency of *S. sungeiensis* SD3 and *S. lividans* TK24. $5x10^8$ of ET12567(pR9604) donor cells containing pESAC13A-P46-BGC28 vector was used in intergeneric conjugation with 10^8 of *S. sungeiensis* SD3 and *S. lividans* TK24 recipient spores. Exconjugant frequency was calculated by dividing the number of exconjugant colonies by the number of the recipient spores used. Values represent the mean frequencies from three independent experiments.

Streptomyces recipient	Exconjugant frequency
S. sungeiensis SD3	$5.37x10^{-7}$
S. lividans TK24	$8.13x10^{-7}$

Table S5. BGCs predicted by AntiSMASH analysis.

Cluster No.	Locus	Most similar known cluster and related product	Percent similarity (%)	Type
1	31020-53566	Cyanobactin	-	Cyanobactin
2	135552-244275	NRPS	-	NRPS
3	644684-691027	Type I PKS, NRPS	-	Type I PKS, NRPS
4a	697584-813883	Coelichelin	100	NRPS
4b	697584-813883	Type I PKS	-	Type I PKS
4c	697584-813883	Type I PKS NRPS	-	Type I PKS, NRPS
5	893321-914067	Terpene	-	Terpene
6	1356647-1396129	Ladderane	-	Ladderane
7	2146523-2167249	Terpene	-	Terpene
8	2218605-2251407	ε-Poly-L-Lysine	-	NAPAA
9	2492957-2503361	Ectoine	100	Ectoine
10	3173805-3194941	Indole	-	Indole
11	3450682-3471923	Aminoglycoside	-	Aminoglycoside
12	3785347-3857226	Mayamycin	77	Type II PKS + Butyrolactone + Melanin
13	3936687-3947464	Desferrioxamin B/Desferrioxamine E	83	Siderophore
14	4484945-4563474	Ishigamide	100	Ladderane + NRPS

15	5337328-5378356	PKS-like	-	PKS-like
16	5542323-5563078	CDPS	-	CDPS
17	5988283-6008957	Redox-Cofactor	-	Redox-Cofactor
18	6627695-6644600	Albaflavenone	100	Terpene
19	6677129-6699044	Thioamitides	-	Thioamitides
20	6726302-6782167	NRPS	-	NRPS
21	6819367-6890024	Spore pigment	100	Type II PKS
22	7362472-7372916	Siderophore	-	Siderophore
23	7492907-7512870	Terpene	-	Terpene
24	7547773-7607441	Streptazone E	75	Type I PKS
25	7616496-7641969	Betalactone	-	Betalactone
26	7804677-7830034	γ-butyrolactone + Geosmin	100	Butyrolactone + Terpene
27	7980246-8015371	NAPAA	-	NAPAA
28	8053375-8065709	Siderophore	-	Siderophore
29	8091518-8114069	Citrulassin	100	Lassopeptide
30	8274405-8296487	Redox-Cofactor	-	Redox-Cofactor
31	8634943-8661657	Hopene	92	Terpene
32a	8936697-9059458	Type I PKS	-	Type I PKS
32b	9059587-9172263	Type I PKS		Type I PKS
33	9236269-9308760	Chartreusin	100	Type II PKS
34	9413045-9440425	Informatipeptin	100	RiPP
35	9456412-9469902	Siderophore	-	Siderophore
36	9519915-9569936	Type I PKS + NRPS	-	Type I PKS + NRPS
37a	9736358-9799254	Isocomplestatin	100	NRPS + PKS
37b	9799397-9835234	alkylresorcinol		Type III PKS

Table S6. ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR Data (δ ppm, in DMSO-d_6) for furaquinocin M (**6**)[a]

No.	δ_H (*J* in Hz)	δ_C	No.	δ_H (*J* in Hz)	δ_C
2	4.51 (q, 6.7)	90.8	8	6.05 (s)	111.6
2-CH$_3$	1.47 (d, 6.7)	15.0	9	--	182.6
3	--	51.7	9a	--	107.9
3-CH$_3$	1.35 (s)	20.7	9b	--	162.1
3a	--	126.3	10	3.67 (dd, 9.5, 2.4)	74.5
4	--	159.2	11	2.12 (m), 2.17 (m)	32.5
5	7.01 (s)	109.4	12	5.15 (m)	123.2
5a	--	133.1	13	--	131.8
6	--	179.9	14	1.62 (s)	26.1
7	--	159.2	15	1.49 (s)	18.2
7-OCH$_3$	3.77 (s)	56.6			

[a] Assignments were made by a combination of 1D and 2D NMR experiments.

Table S7. ^{1}H (400 MHz) and ^{13}C (100 MHz) NMR data (δ in ppm, in DMSO-d_6) for chlorostreptazone B1 (**11**)[a]

No.	δ_H (*J* in Hz)	δ_C
1	8.05 (1H, br s)	--
2	3.33 (2H, t, 7.0)	38.7
3	2.47 (2H, td, 7.0, 4.5)	23.3

4	6.36 (1H, t, 4.5)	118.1
4a	--	128.3
5	--	103.5
6	--	183.4
7	--	129.1
7a	--	157.3
8	6.29 (1H, q, 7.5)	126.3
9	2.21 (3H, d, 7.5)	13.4

[a] Assignments were made by a combination of 1D and 2D NMR experiments.

Table S8. The biosynthetic gene clusters in S. albidoflavus MD102 predicted by antiSMASH are shown, along with the subsequent genome modification efforts.

Cluster No.	Locus	Most similar known cluster	Type [a]	In J1074	In Post Modification Strain
1	72006-147709	Unknown Terpene-NRPS	Terpene, NRPS	No	Deleted
2	301189-348721	Alteramide	T1PKS, NRPS	Yes	Yes
3	385976-411017	Hopanoid	Terpene	Yes	Yes
4	477446-485282	Unknown Bacteriocin	RiPP-like	Yes	Yes
5	733851-772692	Paulomycin	PKS-like	Yes	Deleted::*bldA, gpp*
6	943245-953524	Unknown Bacteriocin	RiPP-like	Yes	Yes
7	1196271-1256511	Unknown NRPS	NRPS	Yes	Yes
8	1352620-1366013	Unknown Siderophore	NRPS	Yes	Yes
9	1621886-1642836	Geosmin	Terpene	Yes	Yes
10	1955508-1975564	Albaflavenone	Terpene	Yes	Yes
11	2465198-2497664	Thiopeptide	Thiopeptide	Yes	Yes
12	2799470-2820002	Lanthipeptide	Lanthipeptide	Yes	Yes
13	3247182-3294215	Butyrolactone	Type I PKS	No	Yes
14	3586813-3636066	Unknown NRPS	NRPS	Yes	Yes
15	3889944-3994457	Surugamide	NRPS	Yes	Yes
16	4478518-4521430	Unknown NRPS	NRPS	Yes	Deleted
17	4746805-4758625	Desferrioxamine B	NRPS	Yes	Yes
18	5659455-5669853	Ectoine	Ectoine	Yes	Yes
19	6344976-6388218	Indigoidine	NRPS	Yes	Yes
20	6516019-6544745	Isorenieratene	RiPP-like, Terpene	Yes	Yes
21	6,633,062-6,644,429	streptamidine	RiPP-like	Yes	Yes
22	6659980-6701077	Unknown PKS	Type III PKS	Yes	Yes
23a	6707464-6978266	Candicidin	Type I PKS	Yes	Deleted::ϕBT1 site
23b		Antimycin	Type I PKS-NRPS	Yes	Deleted

23c	Unknown PKS	Type I PKS	Yes	Deleted

[a] Biosynthetic Cluster Types: NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase

Table S9. Genes involved in aromatic degradation found in *S. albidoflavus* MD102 but not in other common *Streptomyces* strains.

System	Function	*S. albidoflavus* MD102 (PEG)	*S. albidoflavus* J1074	*S. coelicolor* A3(2)	*S. avermitilis* MA-4680
Aromatic dioxygenase mess	2,3-dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	589	X	X	✓
Biphenyl Degradation	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)	590	X	X	X
Aromatic dioxygenase mess	Cysteine dioxygenase (EC 1.13.11.20)	1211	✓	X	X
Aromatic dioxygenase mess	Cysteine dioxygenase (EC 1.13.11.20)	2071	✓	X	X
Aromatic dioxygenase mess	Quercetin 2,3-dioxygenase (EC 1.13.11.24)	2946	X	X	X
Biphenyl Degradation	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)	3330	✓	X	X
Biphenyl Degradation	2-keto-4-pentenoate hydratase (EC 4.2.1.80)	6088	X	X	X
Biphenyl Degradation	Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol	6089	X	X	X
Biphenyl Degradation	4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39)	6090	X	X	X
Biphenyl Degradation	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-)	6091	X	X	X
Aromatic compound metabolism	3-(3-hydroxyphenyl)propanoate hydroxylase (EC 1.14.13.127)	6092	X	X	X
Aromatic dioxygenase mess	3-carboxyethylcatechol 2,3-dioxygenase (EC 1.13.11.16)	6093	X	X	X

Table S10. Genome comparison between *S. albidoflavus* MD102 and *S. albidoflavus* J1074.

Feature	*S. albidoflavus* J1074	*S. albidoflavus* MD102	*S. albidoflavus* MD102SL01
Size (Mbp)	6,841,649	7,047,131	6,618,450
G+C content (bp)	73.3	73.3	73.4
Total genes	6116	6312	6037
BGCs	23	23	21
tRNA genes	65	65	66
CRISPR system	Cas3	Cas3	Cas3
Restriction System	Type I (M)	Type I (M, S, R)	Type I (M, S, R)

Table S11. Protein sequences of BGC6 terpene genes from *S. tasikensis* P46 predicted to produce unknown diterpene.

Function	Amino acid sequence
GGPPS	MTDAALDEFLDRKSLTAPSHQMLELVETLRGFLSSGGKRIRPVMCLCGWYAA GGKETPRPVVKAAASLELFHACALIHDDVMDNSDARRGRLTLHRLLAERHRR RDPPGNAERFGTNAAILLGDLALAWSDEMFHTAGLSPAQVRAALPVLDAMRS EVMFGQYLDLLATGRPTGDVEEALMASRFKTAKYTVERPLHIGAALAGSGPAI RDALTAYALPVGEAFQLRDDLLGVFGDSRQTGKPVDDDLREGKCTVLMALA VARADAAQLRVLRSLVGRSDLDAEEADAIRDVLVSTGARSVVDRMITSRCRR ALSVLDRAPFPLPATNALRRLAHSASVRTS
Terpene synthetase	MFDIAPQRTGTSSPARHDEKAEGFFLPELPRLLPVAYHPKAAQIEFRSNAWLRR YLSGCFAGEGELLKFLRERVGLYGPLIAPTADEQHALDLADFYHFVGVIDTMA ADHSGLGASHCGARDVFDRIIADFAAGIEPGMSDNSPFGPAARDLWLRISAGL TPHQVERFRTSMSSFLRGVASELPYQLNGSVPDYDTYMAVRRDSFGCDFILLL TEYSLAVDMTELAASPQFAKVHAHAMRQLILVNDVLSLRKELGDPMNAVRV LRRHNGLTLQQAVDAVCELAERHERAYIAARDAVRYGPFGAHTDVRTYLEGL DHLLAGSQEYEYLTPRYFGDGSVWDGSTSGWISLTAPIARFLPEAGPTSEQRRT KAVRIHTRRS
Cytochrome P450	MTTAHEARAESCPSTRGTAPGGLPLLGHALPLRRRPLEFLAALPAQGDLVEVR LGPRRMYLACHPDLVQQVLRDSRTFDKGGPMFDKVRLLTGNGLATSCWAEH RRQRRLVQPAFHQERMAGYAAVMDHEITSMLDSWQEGRVLDVHAAMQALT ARVIVRTLFSTRIEDWAVDEIQRCLPVITRGFFTRMVAPLGLVQKLPTRSNREF DRALERMRNVIDQTVSSHPTGTDHGDLLSGLLRAEDEETGERLAGHEIHDQV MTLLMGAIETTSNTLAWTYHLLGENPEAEARLHREIDSVLPARRPGFDDLPHL GYTQRVVTEALRLYPPTWLLTRSTTCEAELAGLRLAPGTTVALSFYALGHNPA LFCDPERFDPDRWLPERAKTVPRGAWNSFGGGSRKCIGDRFATIETTLVLAAV

	ASGWRLKPHPGPSIRPEPKASLSTGPLPMIPERRKTSGAGAPAARSPWTLTAPP PDDHC

Table S12. Comparison of recently published *Streptomyces* chassis strains.

Comparisons	***S. griseofuscus* DSM 40191**	***S. atratus* SCSIO ZH16**	***S. sungeiensis* SD3**	***S. albidoflavus* MD102**
Publication date	2021	2022	2024	-
Type strain	*S. griseofuscus*	*S. atratus*	Novel	*S. albidoflavus*
BGCs	35	26	37	23
Genome editing method	CRISPR-Cas9 Based editing	Double crossover	CRISPR-Cas9	CRISPR-Cas9
BGC deletion and Genome minimisation	(5) pSGR1FU1 plasmid with 4 BGCs and BGC33	(1) atratumycin BGC	(1) Chartreusin BGC33	(4) BGC1, BGC5, BGC16, BGC23
Insertion of helper genes and elements	-	-	(2) PhiC31 attB site, *sfp* pptase gene	(3) *bldA*, g*pps*, ΦBT1-attB
Heterologous Natural Product expression	(1) actinorhodin	(7) Marinacarbolines, A201A, Deostamides, Pendolmycin, Pre-himastatin, Actinopyone and Grincamycins	(3) Furaquinocins, Tasikamides, chlorostreptazone B1	(1) 1,3,6,8-tetrahydroxynaphthalene

Supplementary Figures

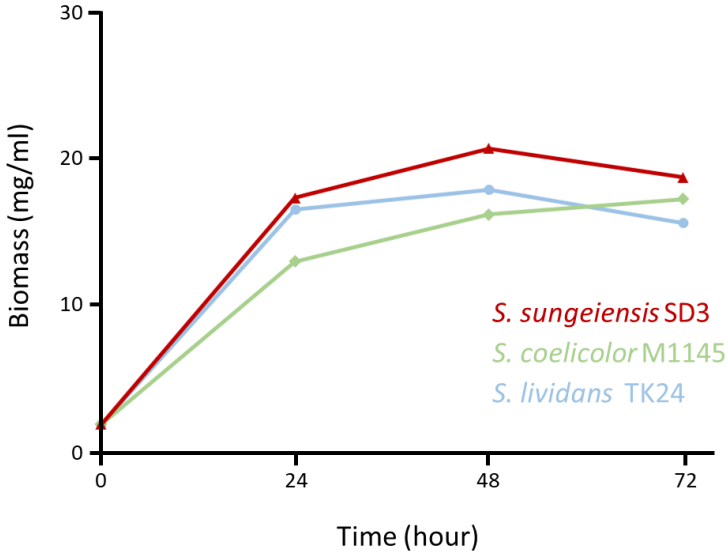


Figure S1. Growth profile of S. sungeiensis SD3 compared with other *Streptomyces* hosts.

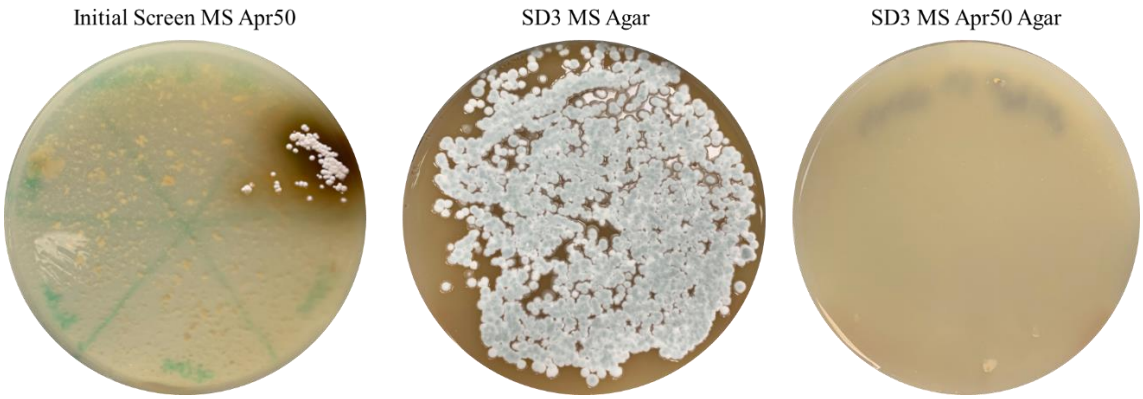


Figure S2. Antibiotic sensitivity test for *Streptomyces* strains. Strains were initially streaked on MS solid agar media containing various selective antibiotics and observed for a lack of growth. This antibiotic sensitivity was reconfirmed by adding liquid cultured biomass to antibiotic plates again. The initial screen plate in the figure shows the lack of

antibiotic sensitivity of the strain MD64, while the SD3 plates show the control and antibiotic selection plates used to reconfirm sensitivity of SD3 to Apr50.

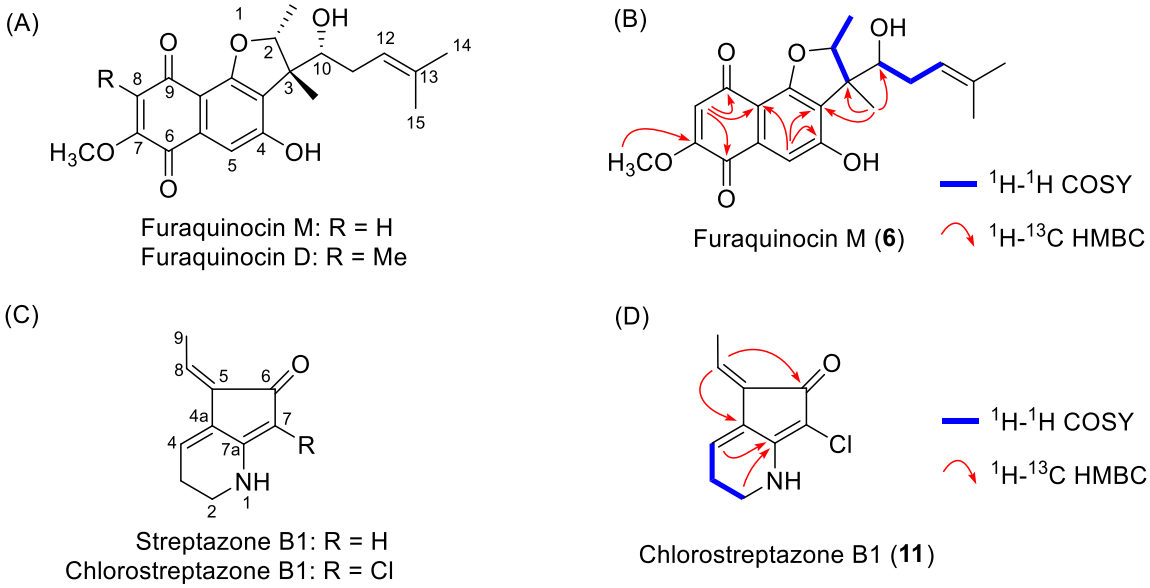


Figure S3. (A) Chemical structures for furaquinocins M (**6**) and D; (B) Key ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC correlations of furaquinocin M; (C) Chemical structure for streptazone B1 and chlorostreptazone B1 (**11**); (D) Key ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMBC correlations of chlorostreptazone B1 (**11**).

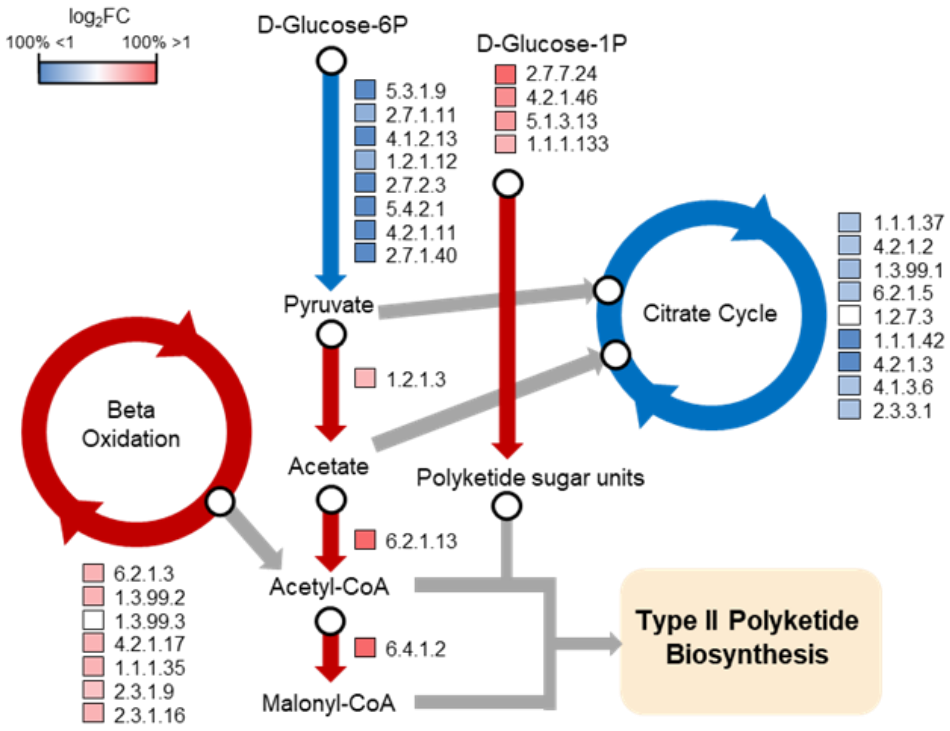


Figure S4. Primary metabolic pathways that underpin chartreusin biosynthesis. Major metabolic pathways involved in the biosynthesis of chartreusin are shown. The enzyme commission (EC) numbers of relevant metabolic reactions of each metabolic pathway are displayed, along with colored boxes displaying the percent of annotated genes in each EC with to <1 or >1 log2FC ($P < 0.05$, DESeq2). The arrows corresponding to each metabolic pathway are colored red or blue based on the pathway being majority up or down regulated.

Figure S5. ^{1}H and ^{13}C NMR spectra of known abscurolide A2 (**1**) (in DMSO-d_6).

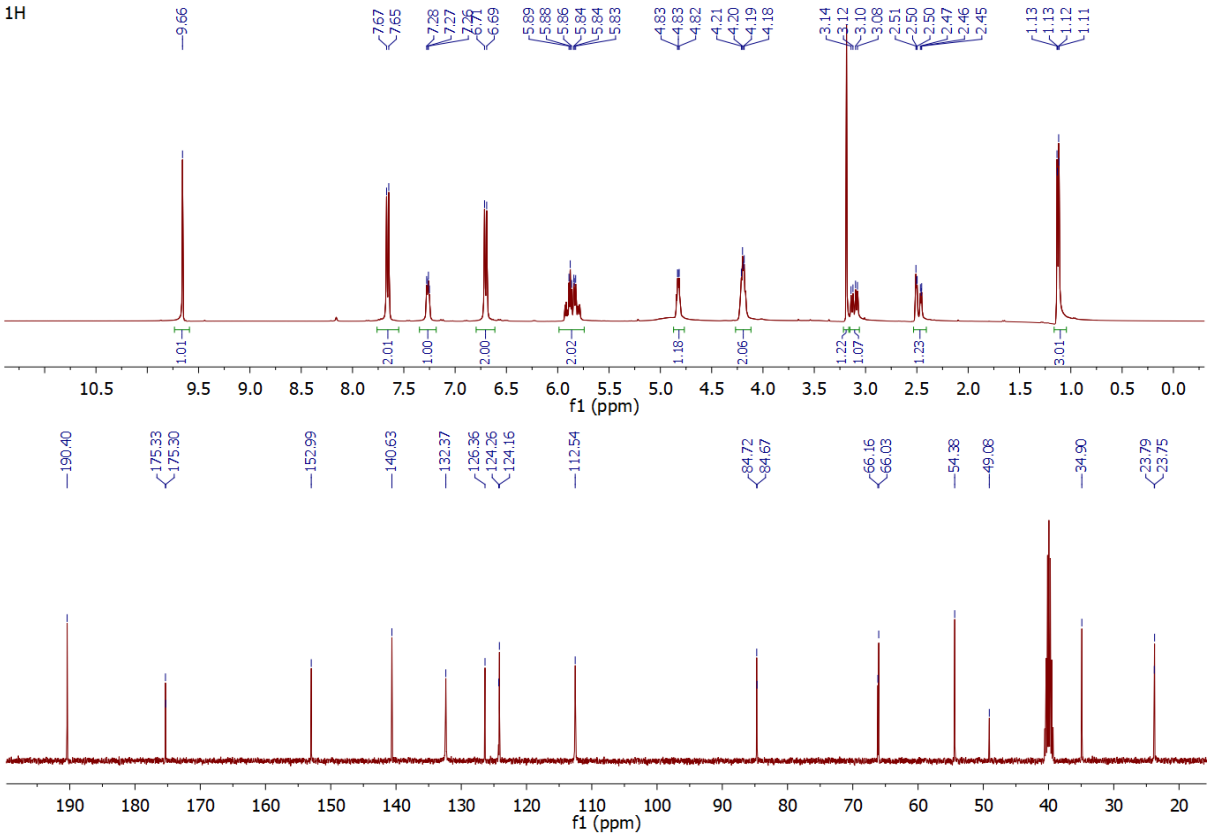


Figure S6. ^{1}H and ^{13}C NMR spectra of known chartreusin (**2**) (in DMSO-d_6).

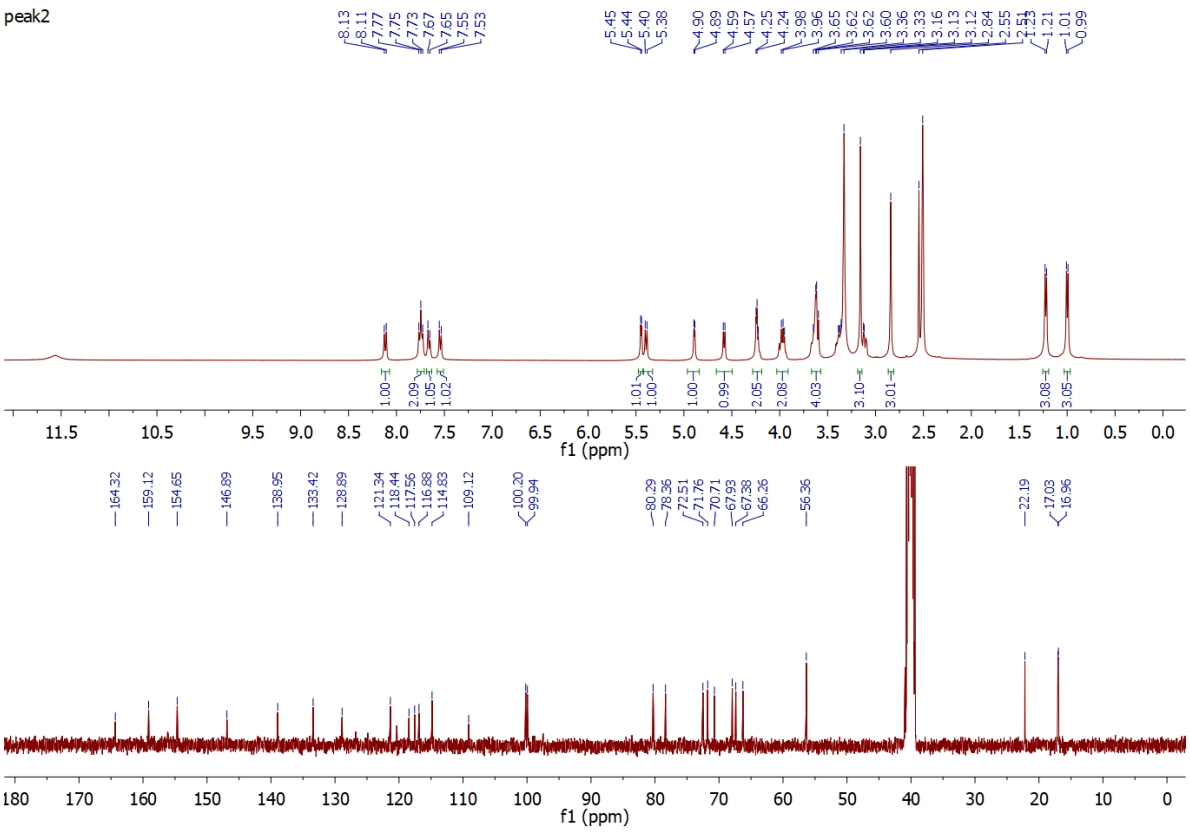


Figure S7. ^{1}H NMR spectrum of known streptazone B1 (**3**) (in MeOD-d_4).

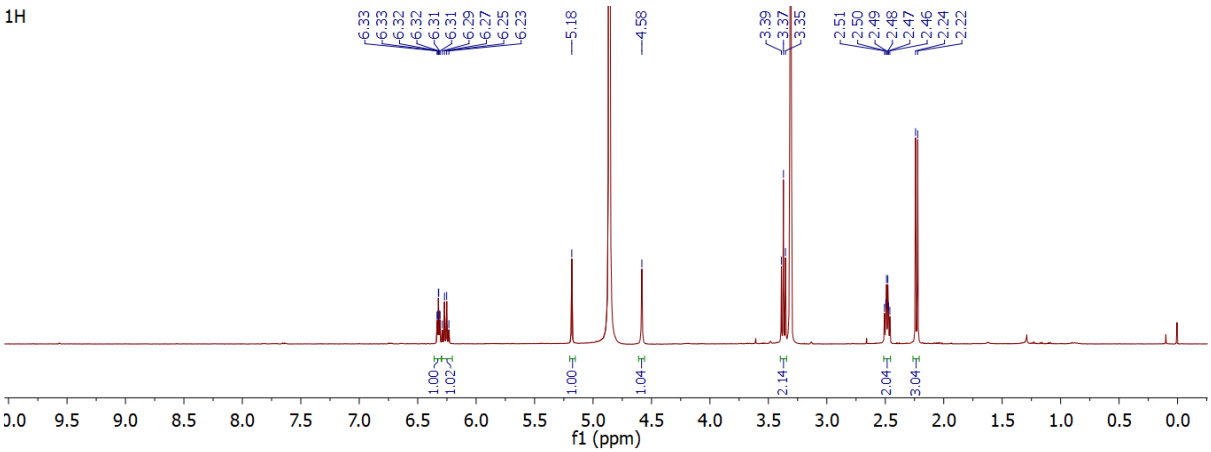


Figure S8. ^{1}H and ^{13}C NMR spectra of known adipostatin A (**4**) (in MeOD-d_4).

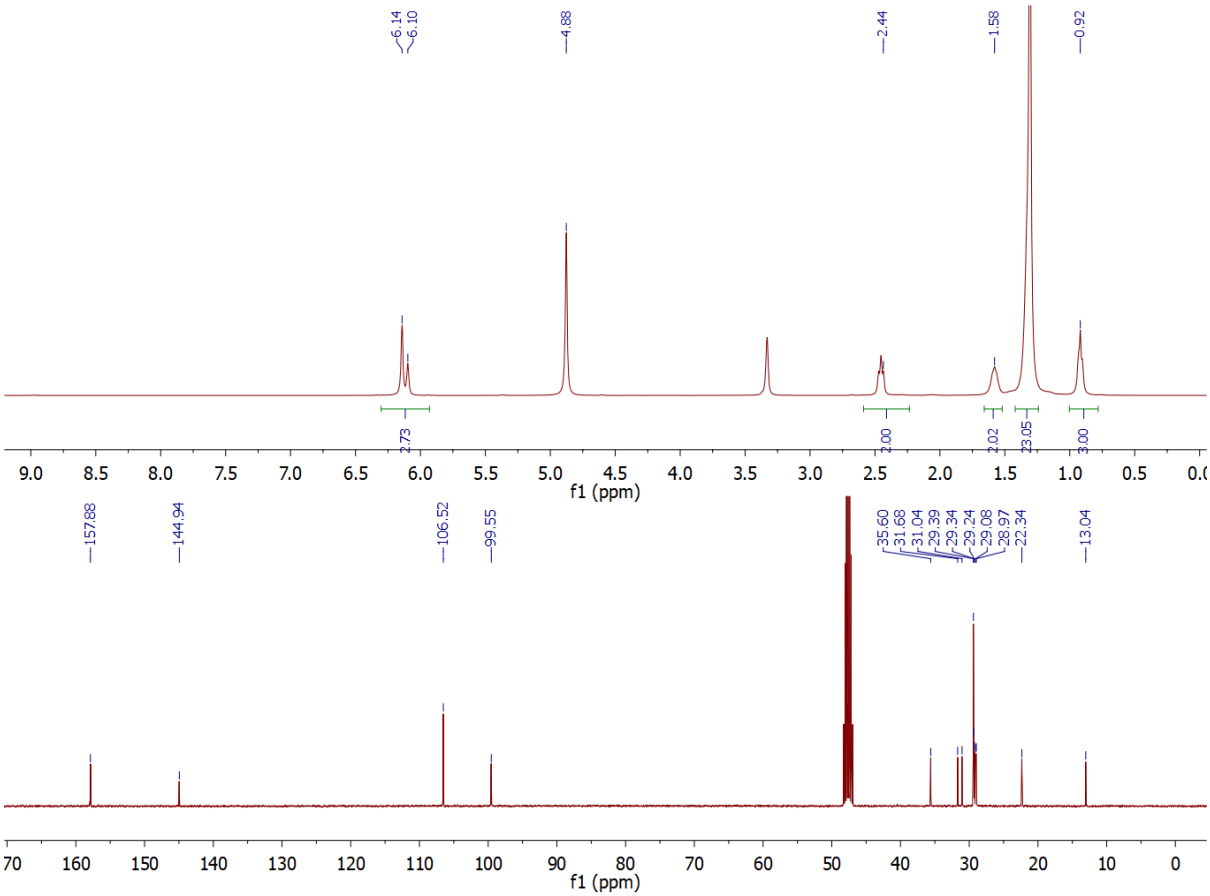


Figure S9. ^{1}H NMR spectrum of known adipostatin B (**5**) (in MeOD-d_4).

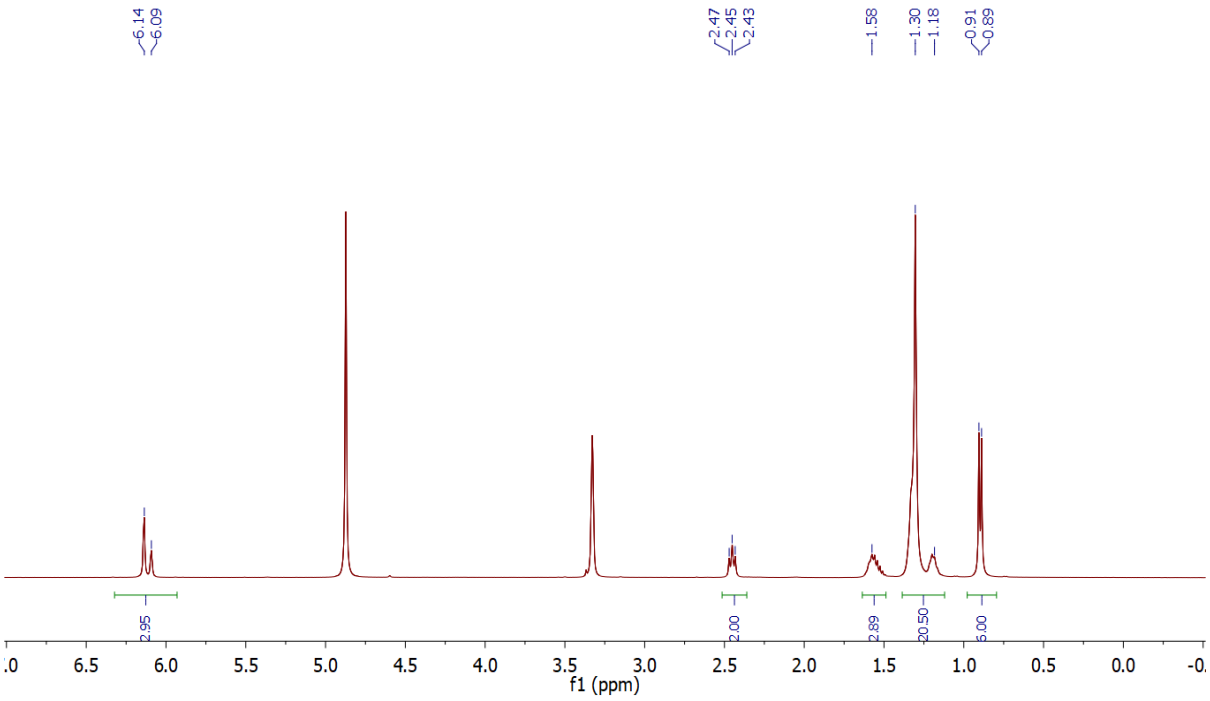


Figure S10-1. ^{1}H NMR spectrum of new furaquinocin M (**6**) (400 MHz, in DMSO-d_6).

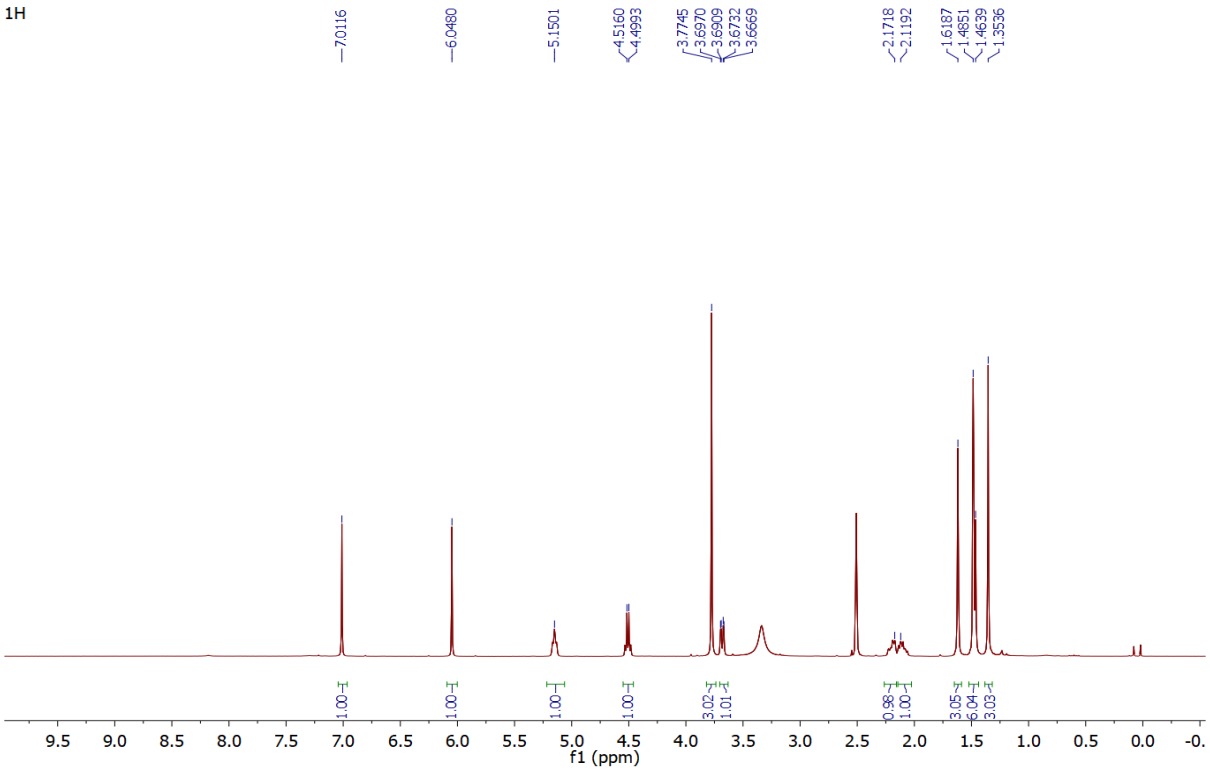


Figure S10-2. ^{13}C NMR spectrum of new furaquinocin M (**6**) (100 MHz, in DMSO-d_6)

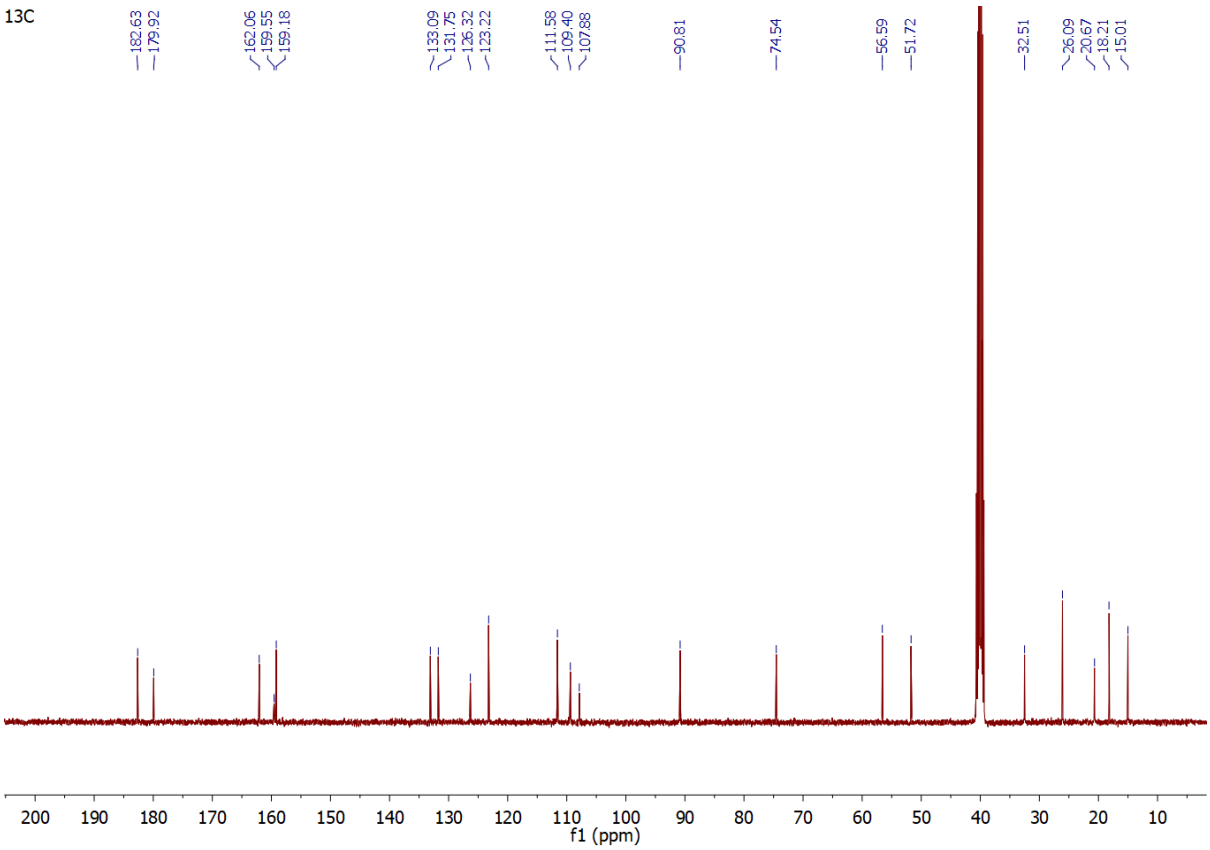


Figure S10-3. ^{1}H-^{1}H COSY NMR spectrum of new furaquinocin M (**6**) (400 MHz, in DMSO-d_6)

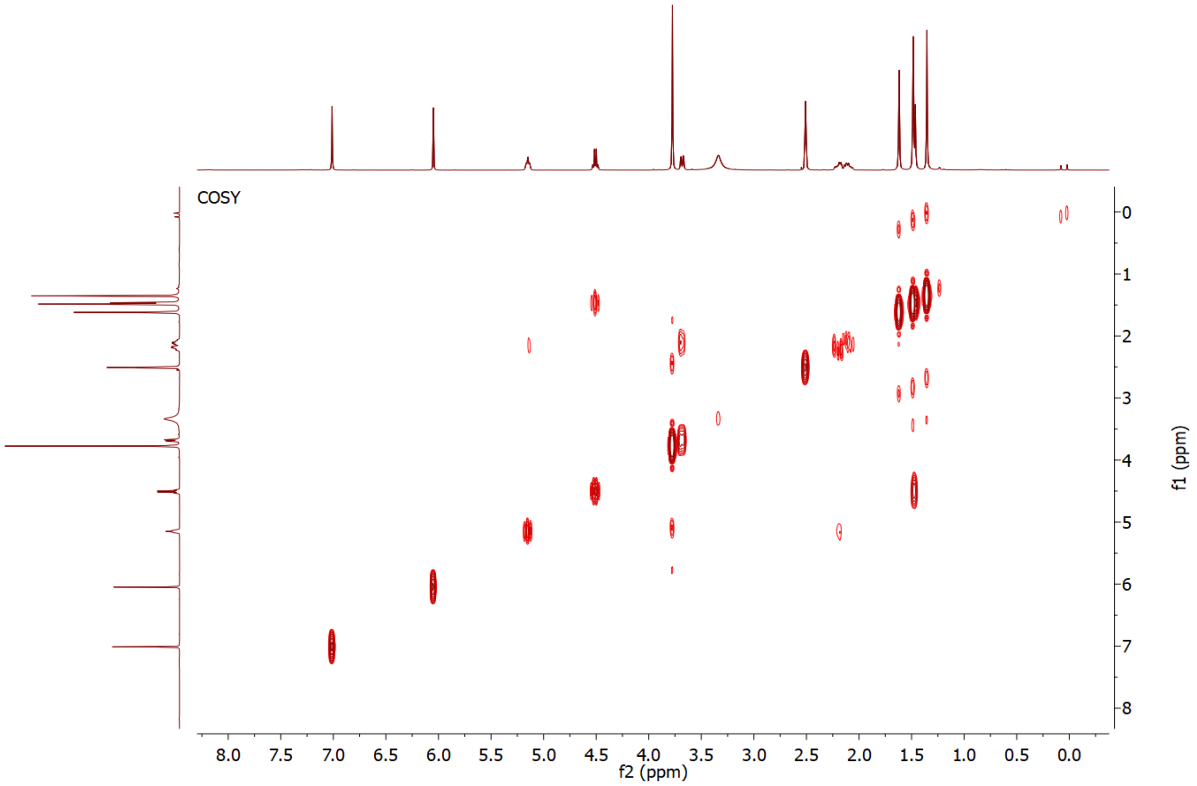


Figure S10-4. ^{1}H-^{13}C HSQC NMR spectrum of new furaquinocin M (**6**) (400 MHz, in DMSO-d_6)

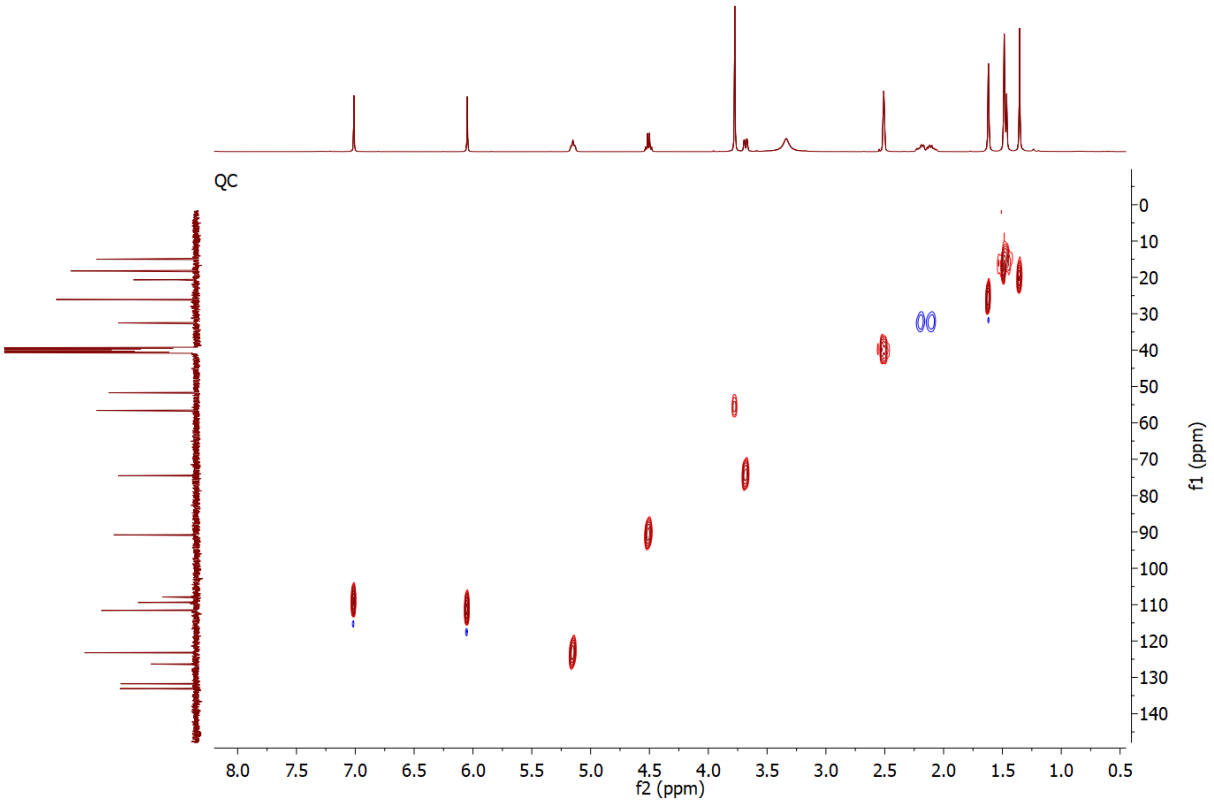


Figure S10-5. ^{1}H-^{13}C HMBC NMR spectrum of new furaquinocin M (**6**) (400 MHz, in DMSO-d_6)

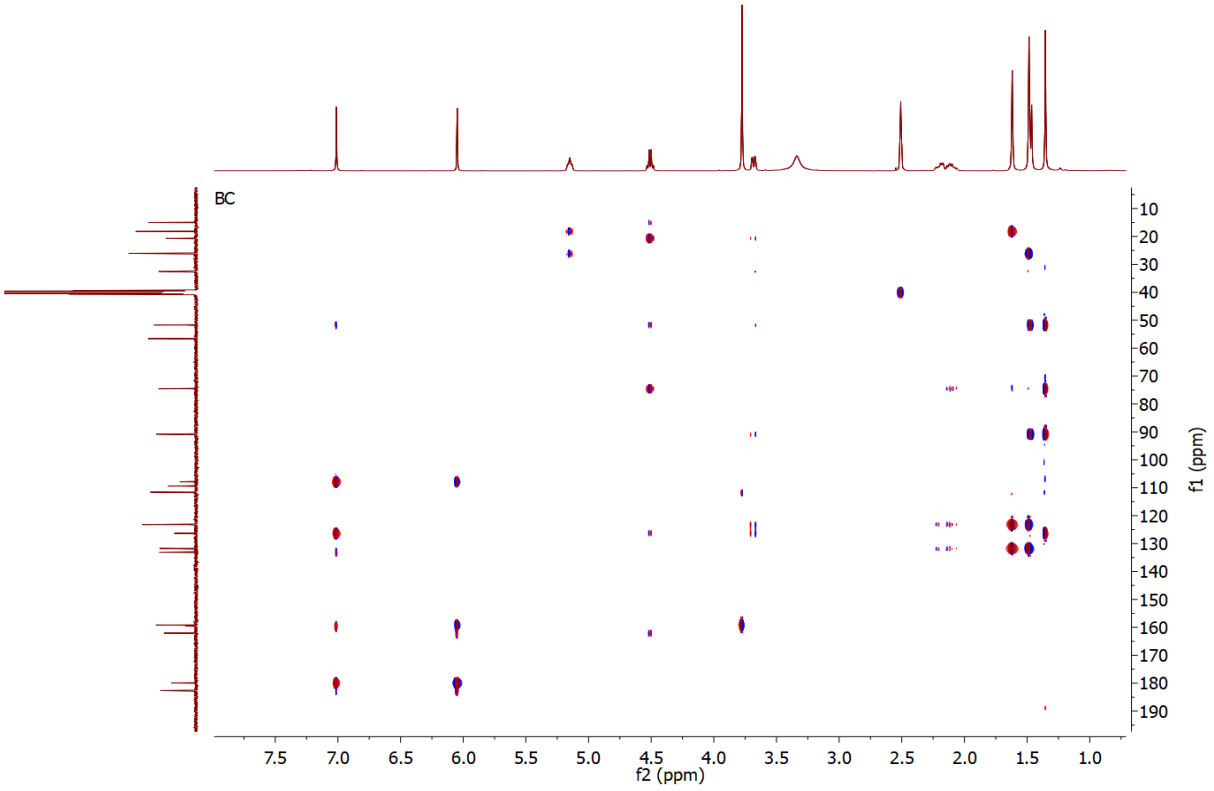


Figure S10-6. NOESY NMR spectrum of new furaquinocin M (**6**) (400 MHz, in DMSO-d_6)

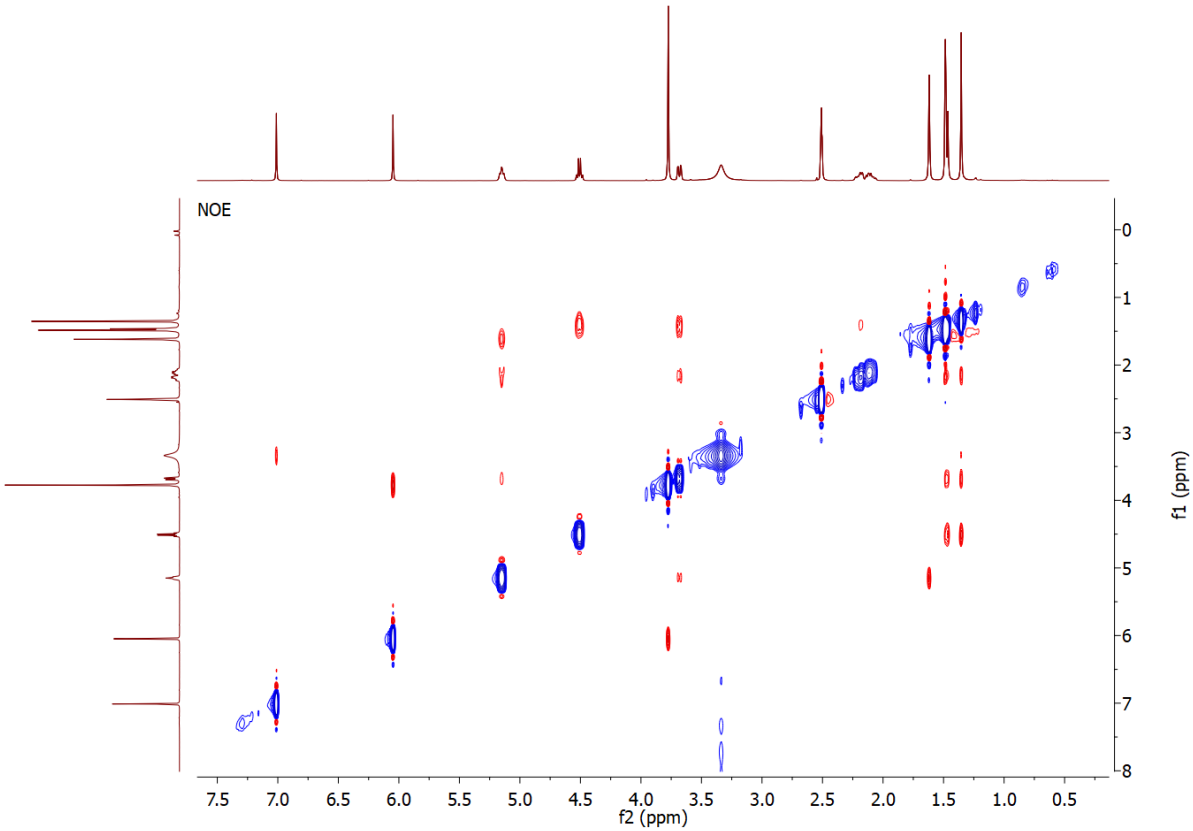


Figure S10-7. HRESIMS of furaquinocin M (**6**)

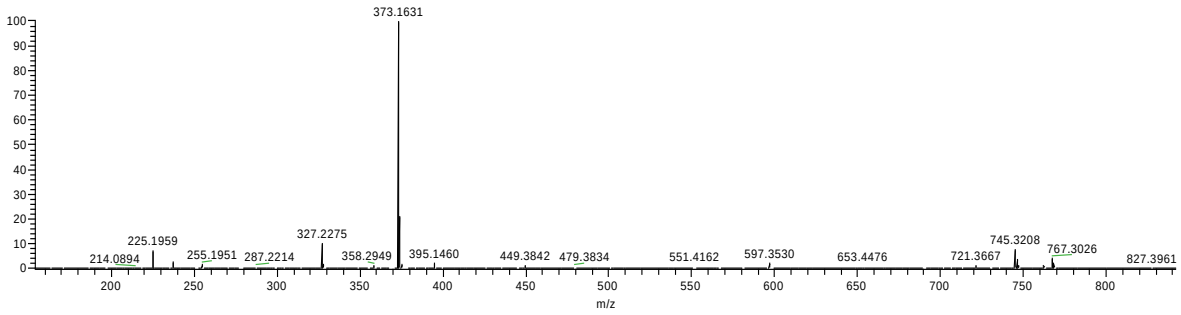


Figure S11-1. ^{1}H and ^{13}C NMR spectra of new chlorostreptazone B1 (**11**) (in DMSO-d_6).

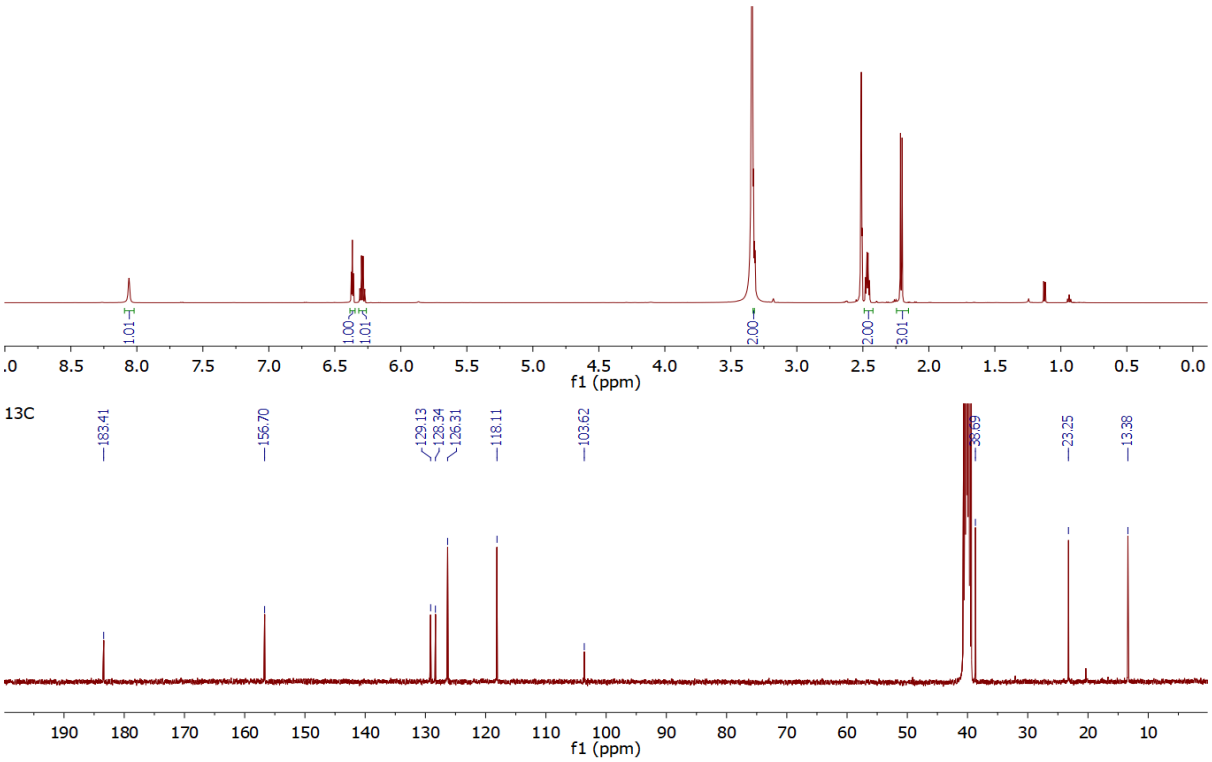


Figure S11-2. ^{1}H-^{1}H COSY NMR spectrum of new chlorostreptazone B1 (**11**) (in DMSO-d_6).

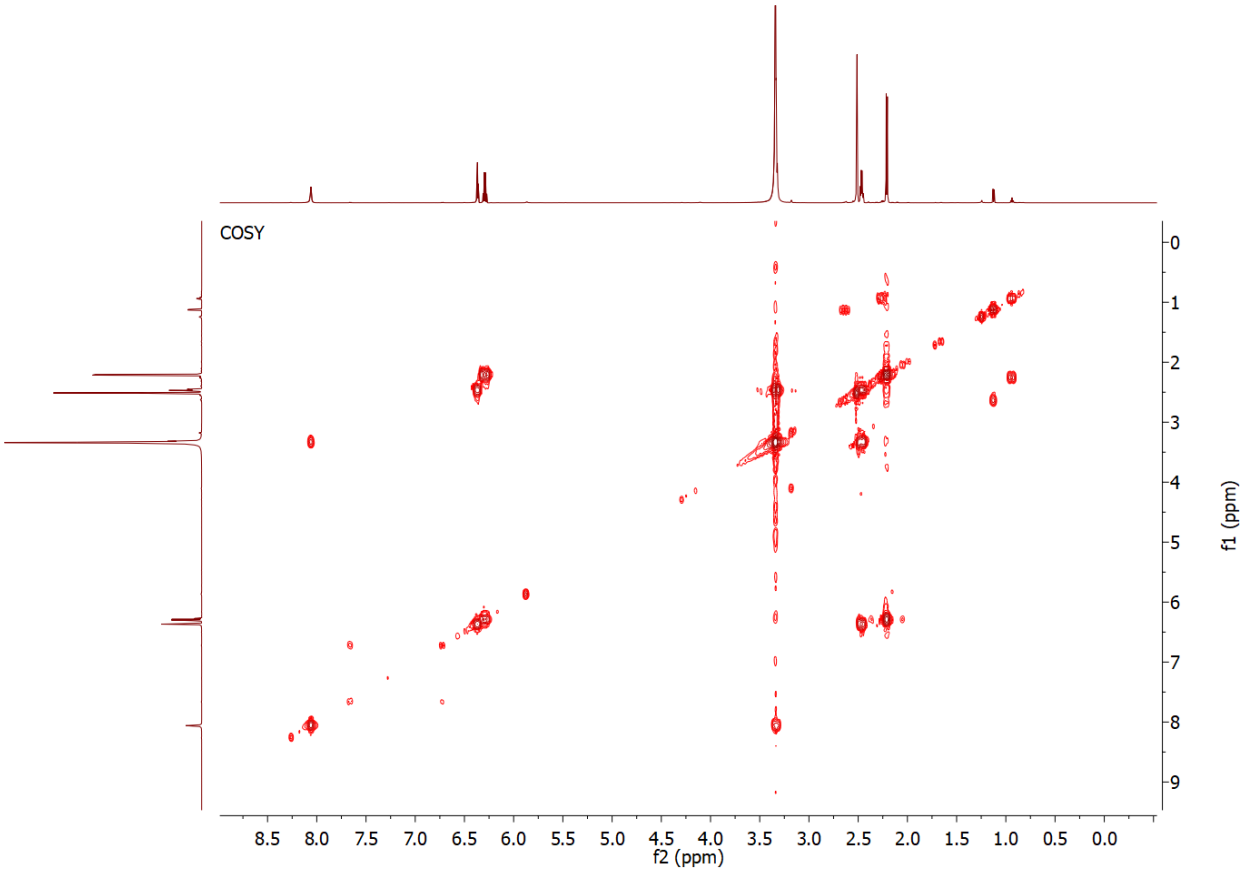


Figure S11-3. ^{1}H-^{13}C HSQC NMR spectrum of new chlorostreptazone B1 (**11**) (in DMSO-d_6).

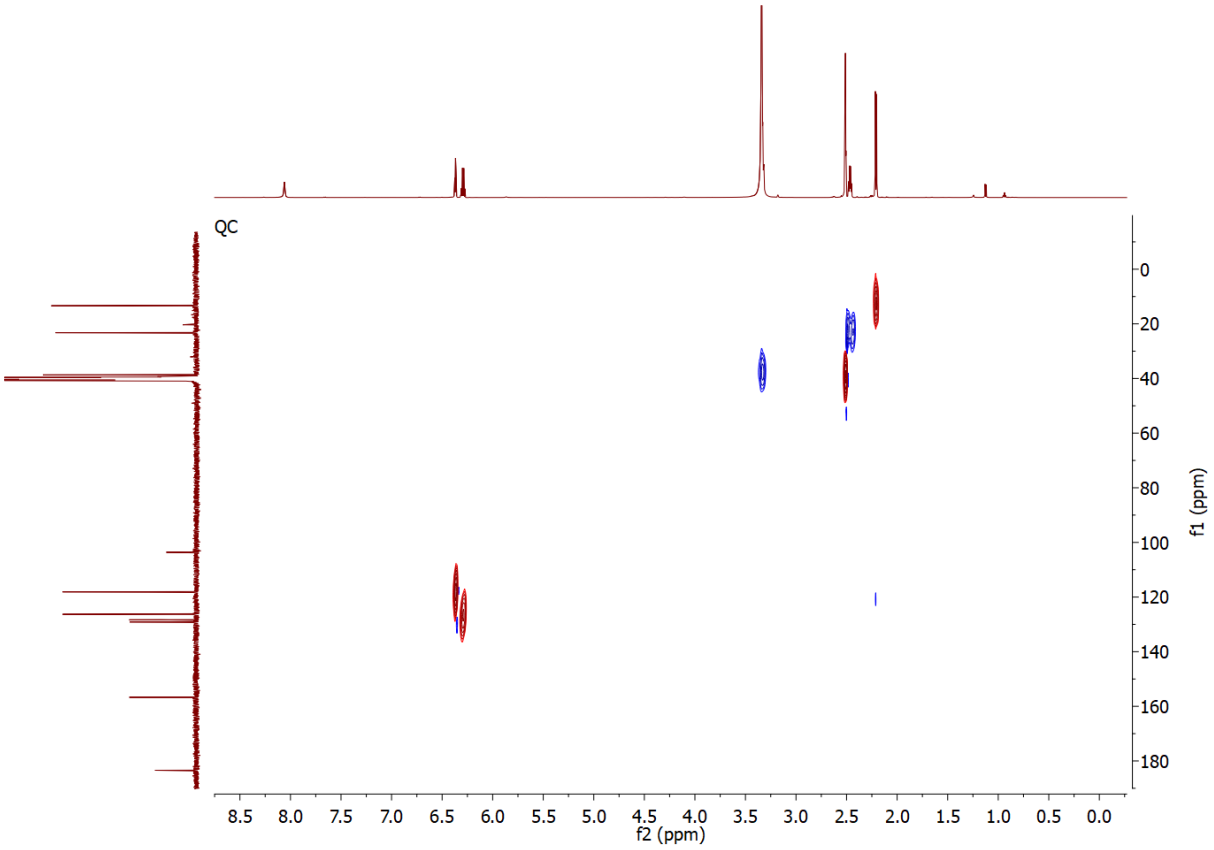


Figure S11-4. ^{1}H-^{13}C HMBC NMR spectrum of new chlorostreptazone B1 (**11**) (in DMSO-d_6).

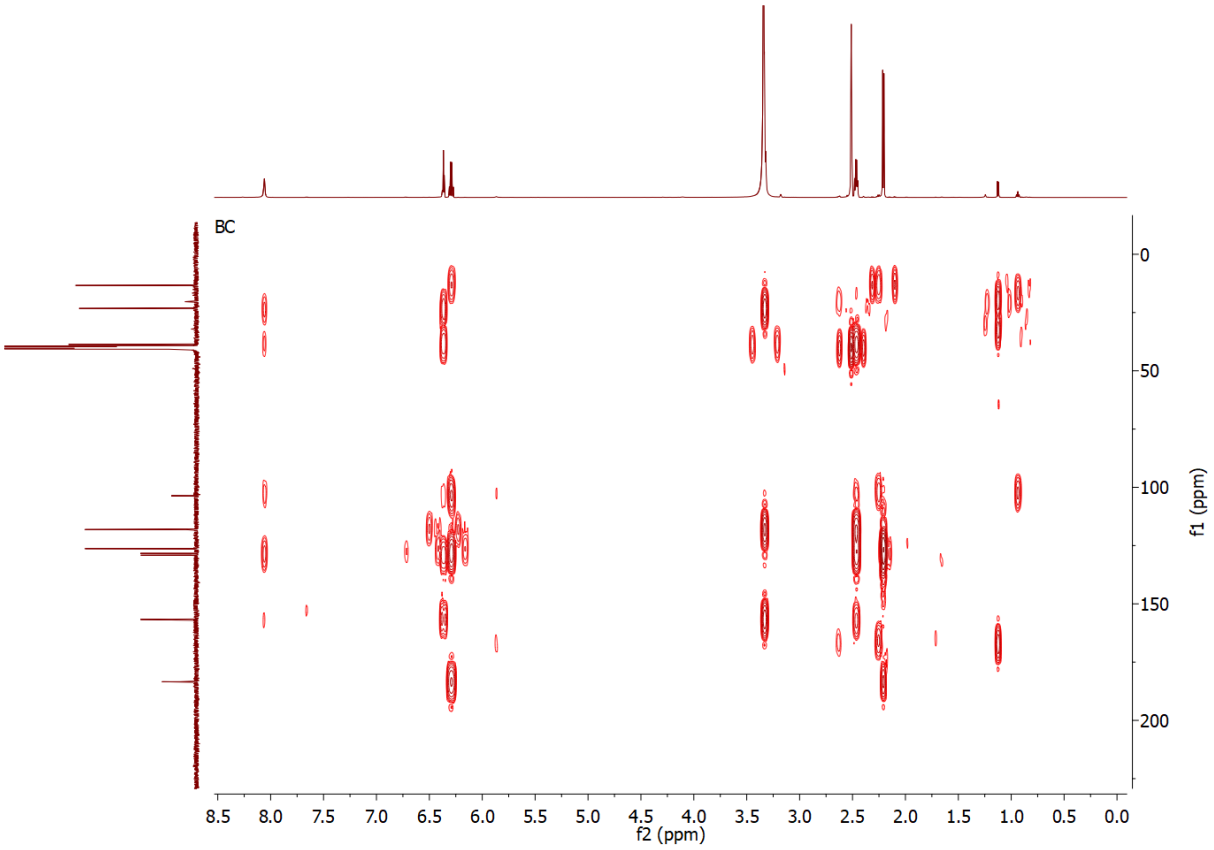
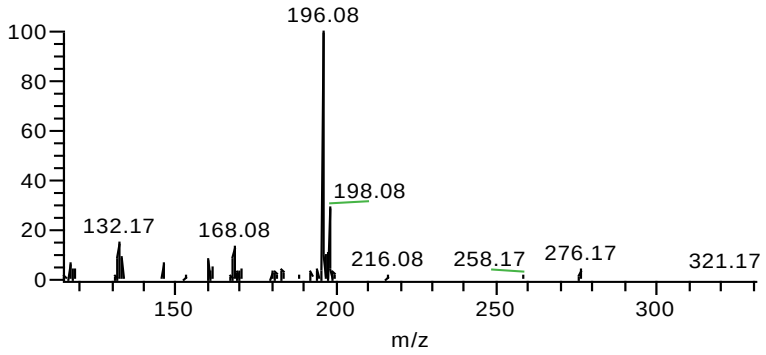


Figure S11-5. ESIMS spectrum of chlorostreptazone B1 (**11**).



Publication list

Peer-reviewed publications

1. Ma, G.-L.; Xin, L.; Liao, Y.; Chong, Z.-S.; Candra, H.; Pang, L. M.; Lee, S. Q. E.; Gakuubi, M. M.; Ng, S. B.; Liang, Z.-X. Characterization of the Biosynthetic Gene Cluster and Shunt Products Yields Insights into the Biosynthesis of Balmoralmycin. Appl Environ Microbiol 2022, 88 (23), e0120822.

2. Candra, H.; Ma, G.-L.; En, S. L. Q.; Liang, Z.-X. Enaminone Formation Drives the Coupling of Biosynthetic Pathways to Generate Cyclic Lipopeptides. ChemBioChem 2022, 23 (22), e202200457.

3. Liao, Y.; Wang, X.-J.; Ma, G.-L.; Candra, H.; Qiu En, S. L.; Khandelwal, S.; Liang, Z.-X. Biosynthesis of Octacosamicin A: Uncommon Starter/Extender Units and Product Releasing via Intermolecular Amidation. Chembiochem 2024, 25 (1), e202300590. https://doi.org/10.1002/cbic.202300590.

4. Lee, S.Q.E., Ma, G.-L., Candra, H., Khandelwal, S., Pang, L.M., Low, Z.J., Cheang, Q.W., Liang, Z.-X., 2024. *Streptomyces sungeiensis* SD3 as a microbial chassis for the heterologous production of secondary metabolites. ACS Synthetic Biology.

Submitted for publication

5. Characterization and optimization of *Streptomyces albidoflavus* MD102 as a heterologous expression chassis. S. Q. E. Lee, G. L. Ma, H. Candra, L. M. Pang, Z. J. Low, Q. W. Cheang, Z. X. Liang

Competing interests

The novel strain *Streptomyces sungeiensis* is currently part of a patent application process:

Liang Z.-X; Sean LEE Qiu En; MA Guang-Lei; Hartono CANDRA; Srashti KHANDELWAL; PANG Li Mei; LOW Zhen Jie; CHEANG Qing Wei, Applications Of Streptomyces Sungeiensis SD3 as a Microbial Heterologous Host. 2024 Singapore provisional patent application number 10202400722U.