MASTERARBEIT | MASTER'S THESIS

Titel | Title

Discovery of new natural products from Micromonospora sp. isolated from Leontopodium nivale rhizosphere via genome mining

verfasst von | submitted by

Una Bajrić BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Magistra pharmaciae (Mag. pharm.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 605
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Pharmazie
Betreut von | Supervisor:	Univ.-Prof. Sergey Zotchev PhD
Mitbetreut von | Co-Supervisor:	Dr. Jaime Felipe Guerrero Garzón BSc MSc

Acknowledgements

I would like to express my deepest gratitude to Univ.-Prof. Dr. Sergey B. Zotchev for giving me the opportunity to work on such a challenging and interesting project and for allowing me to be a part of the most welcoming research team – one that I deeply enjoyed working and spending time with. His support and guidance throughout the project immensely helped me overcome any difficulty I came across along the way.

I am profoundly thankful for the co-mentorship of Dr. Jaime Felipe Guerrero Garzón. His patience, understanding, and extraordinary ability to explain concepts both theoretically and practically, have helped me greatly throughout my Master's thesis research, as well as during the writing process.

I am deeply appreciative of the support I received from Dr. Olga Sekurova and Dr. Anna Vignolle, who were always there to listen and answer all my questions. I would also like to extend my heartfelt thanks to all the members of the pharmaceutical biotechnology research team, from whom I was able to learn so much about the field of science that has fascinated me since my first encounter with it at the University of Vienna. I will cherish all the knowledge I acquired and every laugh I shared with this wonderful team. I would also like to thank Dr. Martin Zehl for performing the analytical part of my project.

Without my beloved friends, this journey towards my Master's degree would have been a mission impossible. I am incredibly grateful to Melis, Isabella, Stefan, Irene and Christoph, who have been my rock in Vienna and made it feel like home, far away from Bosnia. I want to thank them for motivating me to never give up, for listening to me about the challenges I faced throughout my university experience, and for never failing to make any problem seem a little smaller.

I also want to thank my colleagues from the "Schwan" pharmacy for the kindness, shared knowledge, and the feeling of having a second family at work. Without my boss's understanding for all my university-related obligations, my co-workers' encouragement before every exam, and the pride they felt after every little goal I achieved, this whole experience would have been incomplete.

Last but not least, I want to thank my parents, my aunt, and my uncle, for the unconditional love, the sacrifices they made to help me reach my dreams, all the long talks that kept me going, and the unwavering faith they had in me when I was struggling at having it. I will never be able to find words strong enough to thank my mother for dedicating her life to my happiness, lifting me up every time things felt hard, spending sleepless nights helping me prepare for exams, and being present in every moment of my life in Vienna even when we were miles apart. This achievement is as much yours as it is mine, and I will never take for granted how lucky I am to have a family like you.

Zusammenfassung

Die Zunahme von Krankheitserregern, die gegen die derzeit auf dem Markt befindlichen antimikrobiellen Mittel resistent sind, und der Rückgang bei der Entdeckung neuer Wirkstoffe machen es deutlich, dass es dringend notwendig ist, neue Strategien zur Entwicklung von Medikamenten zu generieren. *Micromonospora* Arten sind dafür bekannt, dass sie verschiedene Sekundärmetaboliten mit unterschiedlichen Bioaktivitäten produzieren. Insbesondere Stämme, die in wenig erforschten Umgebungen leben, gelten als vielversprechende Quellen für neue Wirkstoffe. Die Analyse dieser einzigartigen Stämme durch Anwendung neuartiger Genom-Mining-Strategien, Klonierungstechniken und heterologer Expression in gut charakterisierten Wirten, könnte zur Entdeckung neuer bioaktiver Naturstoffe führen.

BGC 1.8, ein biosynthetischer Gencluster im Genom von *Micromonospora* sp. RLA083, isoliert aus der Rhizosphäre von *Leontopodium nivale* subsp. *alpinum*, stand im Mittelpunkt dieser Studie. Mit Hilfe des Bioinformatik-Tools antiSMASH wurde vorhergesagt, dass BGC 1.8 für die Biosynthese eines glykosylierten Terpens kodiert, einer Klasse von Naturstoffen, für die nur wenige Beispiele in Bakterien identifiziert wurden, die aber sehr interessante Aktivitäten aufweisen. Um das Klonen des Clusters und die heterologe Expression zu ermöglichen, wurde die genomische Bibliothek von *Micromonospora* sp. RLA083 erstellt. Der Cluster, der eine geringe Ähnlichkeit mit bereits beschriebenen BGCs für Sekundärmetaboliten aufweist, wurde in einem einzigen Klon der genomischen Bibliothek mit Hilfe eines gepoolten PCR-Ansatzes identifiziert.

Um die Produktion der von BGC 1.8 kodierten Verbindung zu gewährleisten, wurde ein breites Spektrum an genetischen Hilfsmitteln eingesetzt. BGC 1.8 wurde in ein pYES-Plasmid geklont, wobei die natürliche homologe Rekombinationsfähigkeit der Hefe genutzt wurde. Für die heterologe Expression des Clusters wurden drei verschiedene Wirtsstämme und drei verschiedene Fermentationsmedien verwendet, wobei die höchste Produktionsausbeute in *Streptomyces coelicolor* M1154 beobachtet wurde. Die LC-MS-Analyse der Fermentationsextrakte ergab die Produktion eines Terpen-Aglykons, was auf das Fehlen einer Glykosyltransferase-Aktivität hinweist und das Merkmale aufweist, die keiner bereits beschriebenen Verbindung zugeordnet werden können. Die mit den Extrakten durchgeführten Bioassays ergaben keine antimikrobielle Aktivität gegenüber den getesteten Stämmen. Es sind jedoch weitere Bioaktivitätstests, eine Optimierung der Produktionsausbeute und eine NMR-Strukturcharakterisierung erforderlich, um das Potenzial der hergestellten Verbindung als neuartiges Molekül für die Arzneimittelentwicklung weiter zu bewerten.

Abstract

The rise of pathogens resistant to currently marketed antimicrobials and the decline in novel compound discovery highlight the pressing need to establish new drug development strategies. *Micromonospora* species are known to produce diverse secondary metabolites with various bioactivities. In particular, strains inhabiting underexplored environments are seen as promising sources of novel compounds. The analysis of these unique strains by applying novel genome mining strategies, cloning techniques and heterologous expression in well-characterized hosts, may result in discoveries of novel bioactive natural products.

BGC 1.8, a biosynthetic gene cluster located in the genome of *Micromonospora* sp. RLA083, isolated from *Leontopodium nivale* subsp. *alpinum* rhizosphere, was the focus of this study. Using the bioinformatics tool antiSMASH, it was predicted that the BGC 1.8 encodes the biosynthesis of a glycosylated terpene, a class of natural compounds with few examples identified from bacteria but very interesting activities. To enable the cluster cloning, and heterologous expression, the *Micromonospora* sp. RLA083 genomic library was created. The cluster, which shows low similarity to already described secondary metabolite BGCs, was identified in a single genomic library clone using a pooled PCR approach.

To ensure the production of the compound encoded by BGC 1.8, a wide range of genetic tools was employed. The BGC 1.8 was cloned into a pYES plasmid using yeast's natural homologous recombination ability. For the heterologous expression of the cluster, three different host strains and three different fermentation media were used, with the highest production yield observed in *Streptomyces coelicolor* M1154. LC-MS analysis of the fermentation extracts revealed the production of a terpene aglycone, indicating the absence of glycosyltransferase activity, and exhibiting characteristics not associated with any already described compound. The bioassays performed with the extracts did not show any antimicrobial activity against the tested strains. However, other bioactivity tests, optimization of production yields, and NMR structural characterization are necessary to further evaluate the produced compound's potential as a novel molecule for drug development.

Table of Contents

List of Abbreviations

2XYT – Tryptone Yeast Extract medium
Act – Actinorhodin
antiSMASH – Antibiotic and Secondary Metabolite Analysis Shell
BAC – Bacterial Artificial Chromosome
BGC – Biosynthetic Gene Cluster
bp – Base pair
$CaCO_3$ – Calcium carbonate
CAZy – Carbohydrate-Active enZYmes Database
Cda – Calcium-dependent antibiotic
Clm – Chloramphenicol
Cpk – Coelimycin
DH – Dehydratase
DMAPP – Dimethylallyl diphosphate
EDTA – Ethylenediaminetetraacetic acid
gDNA – Genomic DNA
GT – Glycosyltransferase
HPLC – High-Performance Liquid Chromatography
IPP – Isopentenyl diphosphate
kb – Kilobase
Km – Kanamycin
KR – Ketoreductase
LA medium – Luria Agar medium
LB medium – Luria Broth medium
LC-MS – Liquid Chromatography-Mass Spectrometry
M – Molar
MEP – 2-C-methyl-D-erythritol 4-phosphate
$MgCl_2$ – Magnesium chloride
MIBiG – Minimum Information about a Biosynthetic Gene Cluster
mM – Millimolar
MS/MS – Tandem Mass Spectrometry
MW – Molecular weight
MVA – Mevalonate pathway
NaCl – Sodium chloride
NaOH – Sodium hydroxide
Nal – Nalidixic acid
NCBI – National Center for Biotechnology Information
NGS – Next-generation sequencing
NH_3 – Ammoniac
NMR – Nuclear Magnetic Resonance
OD – Optical density
OPP – Pyrophosphate group
PDB – Phage Dilution Buffer
PCR – Polymerase Chain Reaction
PEG – Polyethylene glycol
PKS – Type I Polyketide Synthase
Red – Undecylprodigiosin
rpm – Revolutions per minute
SDS – Sodium dodecyl sulfate
SFM – Soy Flour Mannitol
SS DNA – Single-stranded salmon sperm DNA
Taq – *Thermus aquaticus* polymerase
TAR – Transformation-associated recombination
TBE – Tris-Borate-EDTA buffer
TE – Thioesterase
TRIS – Tris(hydroxymethyl)aminomethane
V – Volt
YEPD/YPD – Yeast Extract Peptone Dextrose

1. Aims of the study

The main objective of this Master's thesis project was to examine the potential of *Micromonospora* sp. RLA083 BGC 1.8 for production of novel compound(s). The steps required to accomplish this aim should be: construction of the strain's genomic library, identification of the BGC within the library, followed by its cloning and heterologous expression in different hosts and under different conditions.

To be able to analyze the BGC of interest, it will be necessary to create the *Micromonospora* sp. RLA083 genomic library and screen it using pooled PCR to identify clone(s) harboring the cluster or its fragments. In the next step, a capture vector with the flanking regions complementary to the BGC's ends will be created. The capture vector and the BGC of interest will be assembled using the natural homologous recombination ability of *Saccharomyces cerevisiae*. Once assembled, the construct will be transformed for the heterologous expression in three different hosts. These hosts will be fermented, and methanol extracts of the cultures analyzed using HPLC and LC-MS to detect compound production and predict its structural characteristics, while bioassays should provide insight into the compound's bioactivity.

2. Introduction

2.1. Microorganisms as a source of novel bioactive secondary metabolites

2.1.1. Secondary metabolites

Secondary metabolites are compounds characterized by their low molecular weight and produced by microorganisms under conditions, such as lack of nutrients, that require activation of survival mechanisms. In contrast to primary metabolites, they are not crucial for microbial growth or propagation. However, they are very important for the adaptational behaviors of microorganisms exposed to different kinds of selection pressure. These adaptations often include competing against other microorganisms inhabiting the same environment, signaling as a way of chemical communication, obtaining nutrients, and developing defense systems. Different classes of secondary metabolites, typically synthesized during the stationary phase of growth, possess either antibiotic or other biological activities and have a broad range of chemical structures. (1,2)

2.1.2. Actinomycetes

Actinomycetes belong to the class of Gram-positive bacteria that inhabit almost all natural ecosystems, from terrestrial to aquatic ones. Their genomes have high G+C content, and code for genes governing biosynthesis of a wide range of secondary metabolites, of which many are in clinical use nowadays. They are members of the phylum Actinomycetota and have an intriguing lifecycle that is very similar to that of fungi. (3)

The life cycle of Actinomycetes starts with the germination of the dormant spores under favorable environmental and nutrient conditions, leading to the formation of hyphae that later develop substrate mycelium on solid nutrient media. When the availability of nutrients becomes limited, the bacteria produce vertically growing aerial hyphae, which overcome surface tension and extend into the air. The formation of prespore chains that mature into thick-walled spores occurs through a controlled cell division of aerial hyphae, increasing the general resistance to environmental stress. The spores tend to remain dormant until the conditions improve again, ensuring the survival of actinomycetes. (4) (Figure 1)

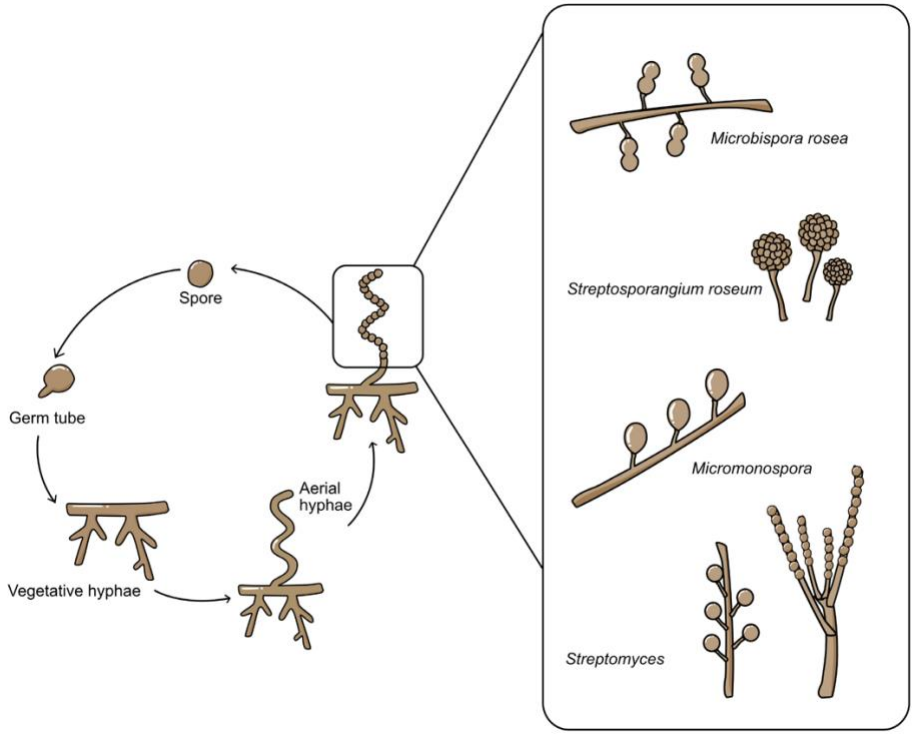

Figure 1. Schematically presented life cycle of sporulating Actinomycetes. Different kinds of conidiospores are shown in the right box (5).

Back in the late 1930s, Selman Waksman, an American microbiologist, identified filamentous Actinomycetes inhabiting terrestrial environments as important sources of antimicrobial compounds. His discovery of numerous antibiotics produced by Actinomycetes, including neomycin and streptomycin, played a crucial role in initiating the "golden age" of antibiotic discovery between 1940 and 1960 (6). Recent studies state that Actinomycetes produce over 10,000 compounds with bioactive properties, making them the largest known source of microbial bioactive substances (45%) (7). Not only do they account for two-thirds of all marketed antibacterials, but they are also known to produce antifungal, anthelmintic, and anticancer agents (3). (Figure 2)

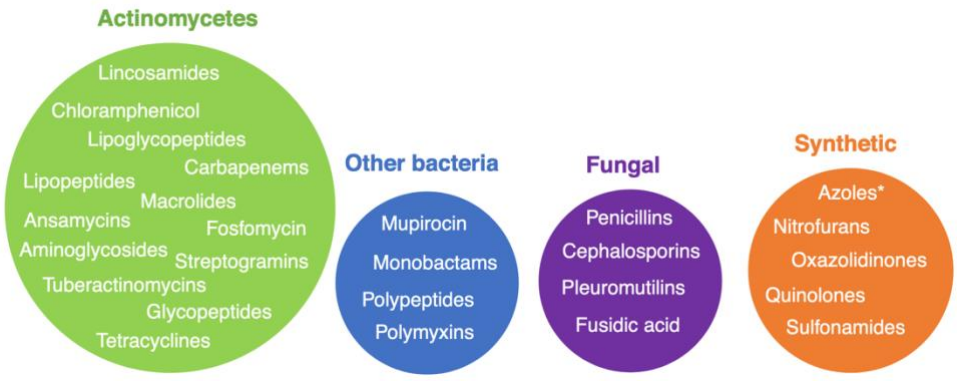

Figure 2. Actinomycetes as the biggest source of antibacterial compounds in comparison to other sources (6).

However, the "golden age" of antibiotic discovery, marked by the rapid identification and isolation of numerous antibiotic classes from soil microbes, soon ended in a decline in novel antibiotic discoveries. From the 1970s onward, analyses of the soil microbe extracts repeatedly resulted in the rediscovery of the already described compounds. This problem occurs at a time of widely spread resistance against available antibiotics. Excessive exposure

to the marketed antibiotics was often highlighted as a main cause of the resistance development. However, the research has found that genes involved in antibiotic resistance evolved long before antibiotics became a part of clinical practice. (2,6,8)

To address this problem, synthetic chemistry has been used as a tool to create analogs of existing antibiotics with improved activity against resistant strains, but this approach is not considered sustainable. Most of the antibiotics in current clinical trials are still either derivatives of already known natural products or synthetic antibiotics rather than completely new classes. To overcome the resistance crisis and expand the antibiotic arsenal, it is necessary to explore novel natural product scaffolds, which are considered optimal starting points for antibiotic development. (6,8)

The metabolites produced by soil-derived microorganisms have been thoroughly researched, not leaving much space for new discoveries. This has led to a growing interest in isolating novel species from different, underexplored ecosystems. For instance, the focus of the current research has largely shifted to rare actinomycetes flourishing in harsh conditions that require the development of special survival mechanisms. Moreover, endophytic actinomycetes are considered one of the most promising sources for future discoveries. (5)

2.1.2.1. *Micromonospora*

Members of the genus *Micromonospora* (family *Micromonosporaceae*, order Micromonosporales, class Actinomycetes) are aerobic, Gram-positive bacteria widely distributed across different natural habitats (Figure 3). The production of carotenoid mycelial pigments is a characteristic of the spore-building *Micromonospora* species that are distinct from other Actinomycetes due to their inability to form aerial hyphae. (9,10)

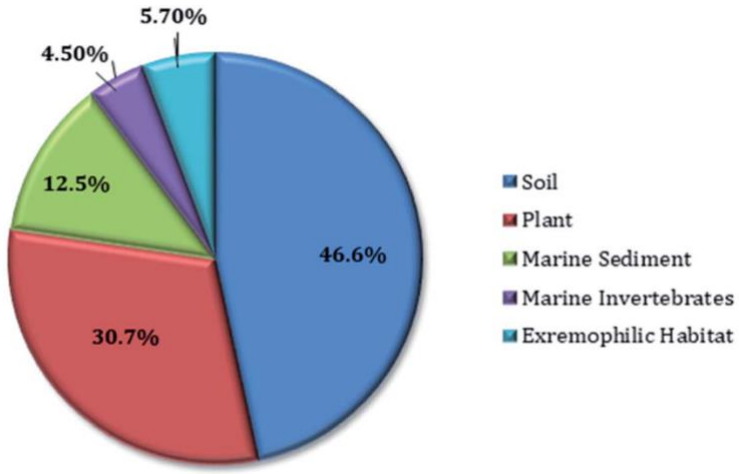

Figure 3. Different natural habitats of *Micromonospora* species (10).

Compounds produced by members of the genus *Micromonospora* have shown various bioactivities, such as antibacterial, antiviral, antifungal, anticancer, antioxidant, and immunosuppressive effects. (9) Some of the examples are (Figure 4):

• Clinically used gentamicin C, a broad-spectrum aminoglycoside antibiotic produced by *Micromonospora purpurea*, consists of five relevant components (C1, C1a, C2, C2a, and C2b), from which gentamicin C1a is considered the most active one. The difference between the five components lies in the positioning of the methyl group, which plays an important role

in their antimicrobial activities. Furthermore, gentamicin C1a derivative, etimicin, is a semi-synthetic aminoglycoside antibiotic that is considered to be valuable in treating bacterial infections caused by aminoglycoside-resistant strains. (11)

• Turbinmicin, a highly oxidized type II polyketide, is produced by *Micromonospora* sp. WMMC-415 isolated from the sea squirt *Ecteinascidia turbinata*. By targeting Sec14p, a protein involved in network dynamics of trans-Golgi, turbinmicin exhibits antifungal activity against *Candida auris* and *Aspergillus fumigatus* pathogens. No currently marketed antifungal treatment targets this crucial protein yet. (12)

• Quinocycline B (kosinostatin) and isoquinocycline B, compounds with both antibiotic and anticancer activities, were produced by a novel *Micromonospora* strain 28ISP2-46T, isolated from deep-sea sponge samples collected from the equatorial Atlantic. These compounds exhibited activity against multidrug-resistant pathogens like *Staphylococcus aureus* and *Acinetobacter baumannii* while also inhibiting DNA topoisomerase IIα. (13)

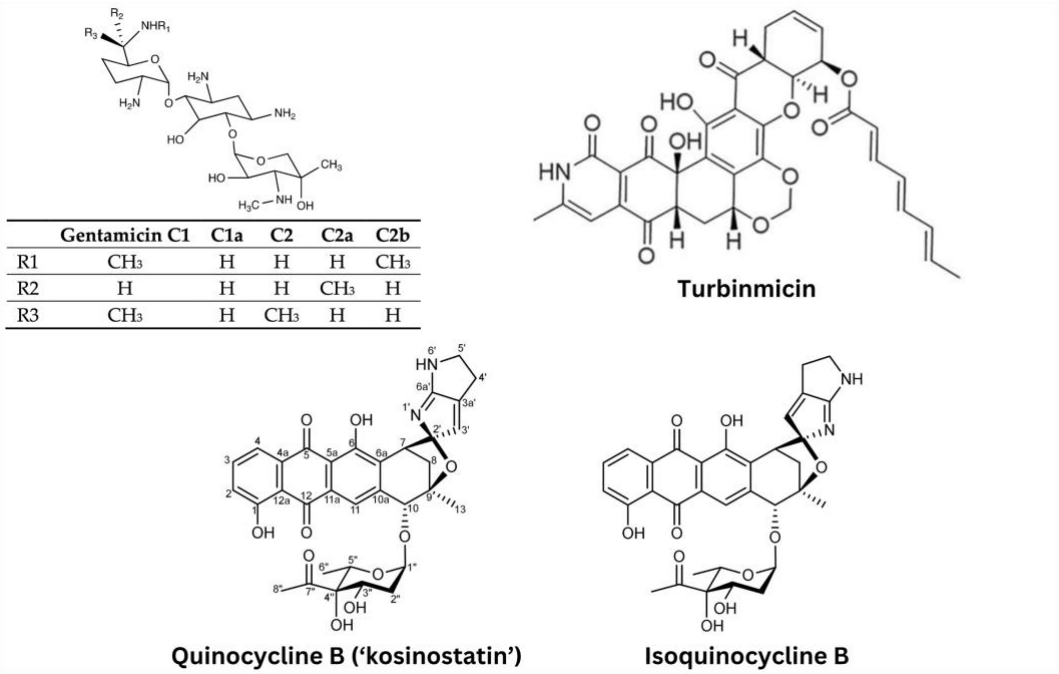

Figure 4. The structural formula of different active compounds isolated from the genus *Micromonospora* (11–13).

The plant rhizosphere, an environment rich in nutrients and carbon sources deriving from the plant, attracts many bacteria, which upon entry into the plant become endophytes. Competition between different bacteria attracted to the plant's rhizosphere makes the *Micromonospora* species inhabiting this environmental niche a possible source of novel bioactive compounds. (14)

2.2. Biosynthetic gene clusters (BGC)

Biosynthetic gene clusters are grouped sets of microbial genes that govern the biosynthesis of secondary metabolites, which are usually exported outside of the cell. (8,15) These clusters typically include core biosynthetic genes that code for enzymes involved in the formation of a compound scaffold, as well as genes encoding tailoring enzymes that modify it. Additionally,

BGCs contain transporter genes for the compound export in some cases, resistance genes for self-protection, and regulatory genes (Figure 5). (1,16)

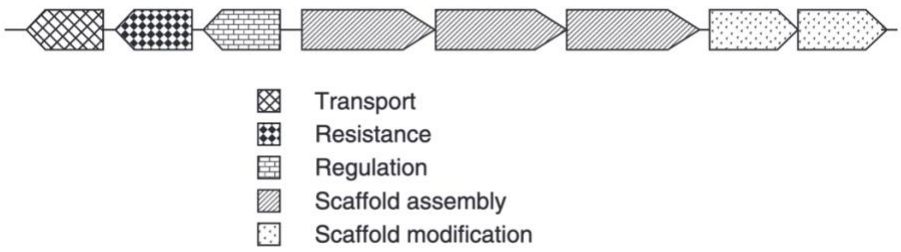

Figure 5. Organization of the biosynthetic gene cluster elements (17).

However, most of these clusters are silent under standard laboratory conditions (16). The complex natural environments are difficult to replicate in the laboratory, but inducing a stress response through modifications of growth parameters or the addition of metals can stimulate the expression of silent BGCs (8,18). Co-culturing is another example of simulating natural habitat conditions, where microorganisms live in complex communities, which force them to compete for resources and produce secondary metabolites (8,19). Additionally, genetic modifications enabled by, for example, CRISPR/Cas9 technology, can lead to transcriptional changes and activation of silent BGCs. (8)

The biosynthesis of secondary metabolites begins with already mentioned “scaffold” enzymes. Prenyltransferases and terpene cyclases belong to this group of compound backbone-building enzymes and are often associated with the production of terpenes. (20)

2.3. Terpenes

Actinomycetes use biosynthetic pathways specified by BGCs to produce terpenes, which are, together with their derivatives, considered to be amongst the largest classes of secondary metabolites with diverse biological activities. Some terpenes have various advantageous effects on human health, as they exhibit such activities as antimicrobial, antiviral, antioxidant, anti-inflammatory, etc. (21)

C5 isoprenoid precursors, dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP), deriving from the mevalonate (MVA) or 2-C-methyl-D-erythritol 4-phosphate (MEP) pathways, are the starting points for terpene biosynthesis (Figure 6). Isoprenyl transferases are responsible for transforming the precursors into isoprenyl diphosphates, whose cyclization is then catalyzed by terpene synthases, resulting in a terpenoid backbone. (21,22) Bacterial terpene synthases lack the high degree of conservation typical for basic biosynthetic enzymes, which, combined with modifications of the scaffold through the addition of diverse functional groups, expands the chemical diversity of these natural products. (21)

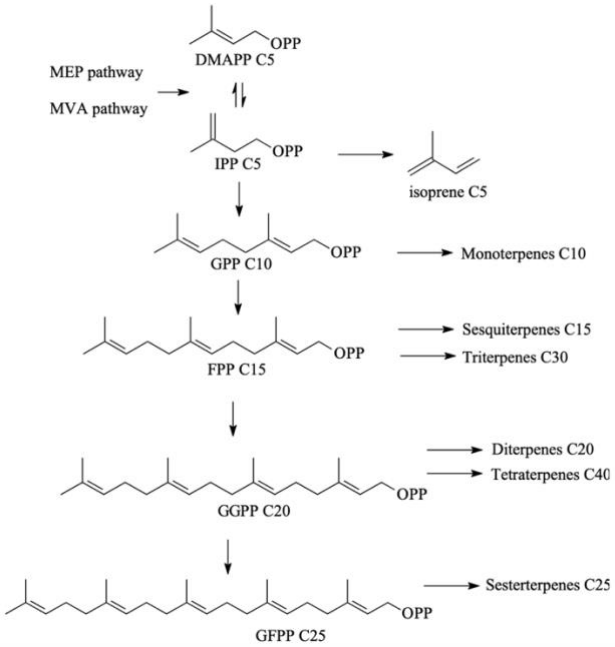

Figure 6. Biosynthetic pathway of different terpenes, OPP: pyrophosphate group (21).

2.3.1. Glycosylated Terpenes

Glycosylation plays an important role in enhancing the stability, solubility, and bioactivity of the molecules, including terpenes, by changing their physical, chemical, and biological characteristics (23). Biological activity of the compounds in nature is often linked to the presence of sugar moieties, which are also significant for biological processes, such as chemical and molecular defence, cell recognition, as well as preservation of cell integrity (23,24).

Glycosyltransferases (GTs) represent a diverse group of enzymes, whose role is to catalyze the transfer of sugar molecules from activated donors to different acceptors, including aglycons. Formation of glycosidic linkages to aglycon compounds is of high importance for the production of valuable natural products. Carbohydrate-Active enZYmes Database (CAZy) classifies GTs into families based on the similarity of their amino acid sequences. However, bacterial, plant, and mammalian GTs show low similarity in their amino acid sequences, even when catalyzing the formation of the same glycosidic linkage. This phenomenon is also seen amongst bacterial GTs. (25)

There are not many reported examples of glycosylated terpenes in comparison to other compound classes deriving from bacteria. However, *Streptomyces*, *Nocardia* and *Salinispora* were found to produce this class of natural compounds, which represent very interesting molecules (Figure 7). (26,27)

- *Streptomyces* sp. Tü 6071 has been identified as a producer of phenalinolactones, which represent a group of C-4 glycosylated diterpenoids with selective antibacterial activity against Gram-positive bacteria. Their diterpenoid core is decorated with unusually oxidized γ-butyrolactone, a pyrrole-carboxylic acid, and trideoxysugar moiety. (26,28,29)

- Another identified glycosylated diterpenoid producing strain is *Nocardia terpenica* IFM 0406, whose product, brasilicardin A, exhibits potent immunosuppressive activities and low

toxicity. The compound's mode of action is unique and is based on inhibiting the amino acid transporter system L, disrupting T-cell activation through deprivation of energy sources. This mode of action is not found in any of the currently marketed drugs, making amino acid transporter system L a promising target for novel immunosuppressive therapies. (30,31)

• *Salinispora tropica* CNB-440 has shown a production of a novel C_{40}-carotenoid, sioxanthin, whose biosynthesis is encoded by four genomic regions. The compound, a glycosylated terpenoid with an aryl moiety, is thought to play a role in oxidative stress protection. Furthermore, sioxanthin's biosynthetic pathway genes are also found within *Micromonospora* genera, but are absent in other actinomycetes outside the *Micromonosporaceae* family. (27)

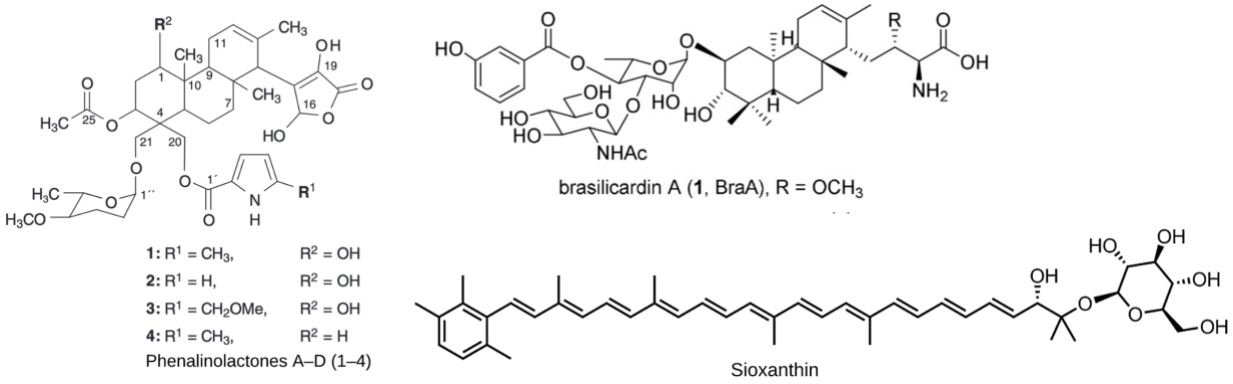

Figure 7. Structural formula of different glycosylated terpenoids isolated from *Streptomyces* (phenalinolactones), *Nocardia* (brasilicardin A) and *Salinispora* (sioxanthin) strains (27,29,31).

2.4. Genome mining

Investigating biodiversity with an aim of identifying novel biological resources is defined as bioprospecting (32). The traditional bioprospecting workflow was previously based on a "top down" method, which started with characterization of the produced compound's biological and chemical properties. On the other hand, a newer "bottom up" approach uses bioinformatics and methods of molecular biology to predict a microorganism's biosynthetic potential based on its genetic information. (33)

Efforts on bacterial bioprospecting can be significantly improved by modern next-generation sequencing (NGS) methods, such as metagenome sequencing, which allows the identification of novel BGCs that encode biosynthesis of secondary metabolites directly from the environmental samples. Cultivability of bacteria does not affect the outcome of this approach (34).
Genomes of actinomycetes were known to be challenging to sequence because of the high G+C contents that polymerase chain reaction (PCR) fails to amplify efficiently, resulting in coverage biases and gaps in assemblies (35). However, progress in DNA sequencing methods that couple long-read data, such as Oxford Nanopore MinION and PacBio SMRT, with the precision of NGS platforms like Illumina, has made acquiring high-quality genome data of actinomycetes much easier (36).

"Genome mining", a new bioinformatics-based approach for identifying BGCs encoding the biosynthesis of natural products, allows the cluster detection by analyzing bacterial genomes (34,37). Due to the vast amount of sequenced, yet uncharacterized BGCs, there is a pressing

need for structural and functional predictions of the BGC metabolites in order to be able to efficiently prioritize the lab research (15).

An important genome mining tool, antiSMASH, is a software that can rapidly and with great accuracy identify BGCs for all known classes of secondary metabolites. It can also compare the identified BGC with the database of already known BGCs for secondary metabolites, such as MIBiG (Minimum Information about a Biosynthetic Gene Cluster) database, and predict the function of the identified gene based on evolutionary similarities. (38,39)

2.5. Heterologous expression

One of the approaches to stimulate the biosynthesis of a compound encoded by a BGC of interest is to clone and heterologously express the cluster in an appropriate exogenous host. The use of the exogenous host is often desirable because the native host either grows slowly, does not produce high amounts of the compound, or is genetically difficult to manipulate. In such cases, a heterologous host can optimize the compound yields. (34,40)

This strategy, which involves sequencing of the genome, its analysis with bioinformatics tools and subsequent cloning and heterologous expression of the identified BGC of interest, has already delivered various novel natural product scaffolds (37). (Figure 8)

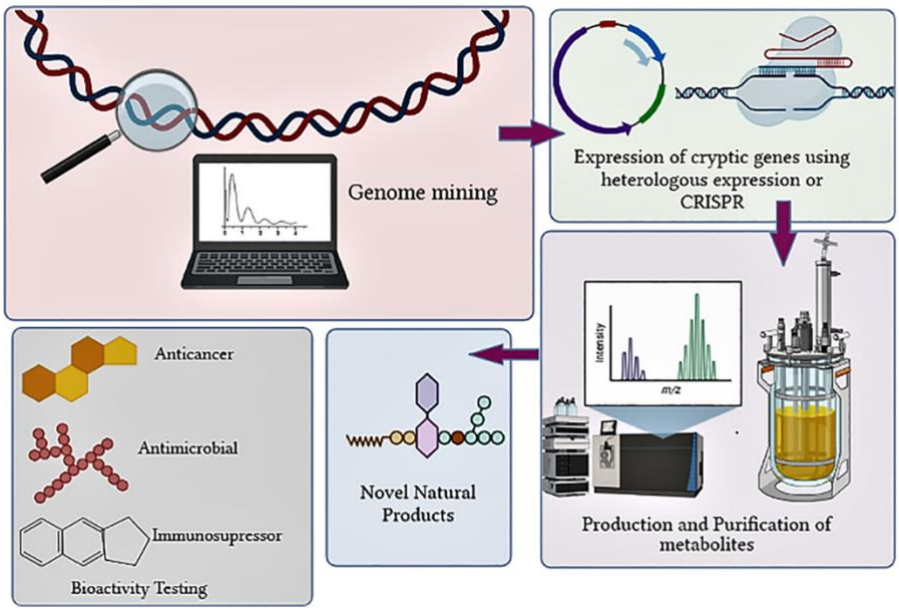

Figure 8. Schematical presentation of the steps leading to the discovery of potentially novel natural products (41).

Actinomycete genomes are typically between 8 and 10 megabases long and harbor 20 to 30 BGCs, usually ranging from 10 to 120 kb in size, that encode secondary metabolite synthesis machineries (42,43). The complexity of the genomes often requires the genomic library construction in order to systematically store and analyze the genetic material. For this purpose, the genetic material needs to be pure and available in sufficient quantity to be cloned into vectors such as plasmids, fosmids or bacterial artificial chromosomes (BACs). Plasmids can carry shorter inserts of 1-10 kb, fosmids hold fragments of 10-50 kb, while BACs are suitable for larger DNA inserts (150-350 kb) (44). To precisely identify the location of the BGC of interest within a library, one of the approaches used is pooled PCR (45).

Once identified and isolated, the BGC of interest can be cloned and expressed heterologously in hosts such as *Escherichia coli, Saccharomyces cerevisiae,* or *Streptomyces coelicolor*, that require less tedious cultivation and genetic manipulation (37). When choosing an appropriate host, it is important to consider the complexity of the compound that needs to be produced, biological similarity between the heterologous host and the native producer, as well as the presence of the host's metabolic capacity to produce a natural compound (46).

Escherichia coli is a well-characterized, easily cultivable, and fast-growing heterologous host, which allows application of a wide range of genetic modification tools. However, *E. coli* is associated with the lack of metabolic precursors and sigma factors needed for foreign promoter recognition, making secondary metabolite production more challenging. Using this bacterium as a host has further disadvantages, such as its inability to produce complex compounds and deliver properly folded proteins. (34,46)

Due to its ability to endure rough industrial conditions and successfully produce compounds deriving from eukaryotes, *Saccharomyces cerevisiae* is another widely used heterologous host (46,47). Its well-characterized genome also allows application of genetic engineering methods, for instance, the employment of either strong constitutive or inducible promoters to increase the production yields of the compound (47). However, the efficiency of *S. cerevisiae* as a heterologous host can be limited by its sensitivity to product toxicity and the need for a precursor supplementation in case of polyketide and nonribosomal peptide biosynthesis in yeast (48).

For the heterologous expression of complex BGCs, *Streptomyces* strains are considered the preferred choice, since their genetic similarity to native producers ensures a successful synthesis of natural products with minimal need for genetic manipulations. The strains also possess essential precursor biosynthesis pathways, allow correct post-translational modifications, and are resistant to a vast number of toxic compounds. However, the slow growth rate and the use of available precursors for the synthesis of endogenous compounds can make a large-scale production and product purification challenging. (34,36,46)

Deletion of BGCs encoding endogenous pathways in heterologous hosts is a strategy used to improve the host's applicability (34). For instance, *Streptomyces coelicolor* M145 was genetically modified to increase levels of secondary metabolite production, leading to the creation of the engineered strain *S. coelicolor* M1154. By deleting the four BGCs, the production of four native antibiotics, coelimycin (Cpk), actinorhodin (Act), calcium-dependent antibiotic (Cda), and undecylprodigiosin (Red), was eliminated. Additionally, point mutations in the transcription and translation machineries contributed to the improved production yields. (49)

2.5.1. Transformation-associated recombination (TAR) cloning

The Transformation-associated recombination (TAR) system is an *in vivo* cloning method based on the natural homologous recombination ability of *Saccharomyces cerevisiae*, which allows BGCs of interest to be assembled on a suitable vector (50).

A TAR vector used in this method contains two "hooks", generated by PCR, with sequences homologous to the flanking sequences of the target region or gene that needs to be cloned (51,52). These "hooks" can be as short as 60 bp (base pairs), but normally fragments around

500 bp are preferred for better assembly efficacy. To allow the integration of the target DNA, prior to the recombination reaction, the vector is linearized, separating the “hooks”. (51)

A TAR vector contains both bacterial and yeast elements, which allows its application in various hosts. The yeast elements, which include a yeast replication determinant and selectable marker, are responsible for the stability and propagation of the construct in *S. cerevisiae*. On the other hand, the BAC cassette’s presence within the vector ensures replication and selection in *E. coli* (37,51). Gene encoding integrase, phage attachment site, origin of transfer, and another antibiotic selection marker ensure conjugation of the final construct from *E. coli* into *Streptomyces* host and its stable integration into the genome (37). An example of a vector used in TAR cloning is pCGW, designed to successfully capture and stably maintain large BGCs encoding the biosynthesis of natural products, with its copy number being inducible to approximately 100 copies per cell in *E. coli* (53).

3. Materials

3.1. Media

*All media were autoclaved at 121°C for 20 minutes. When needed, the pH was adjusted before autoclaving.

3.1.1. *Escherichia coli* strains culture media

Luria Agar (LA) medium
- Tryptone: 10 g
- Yeast extract: 5 g
- Sodium chloride (NaCl): 10 g
- Agar: 15 g
- Distilled water: up to 1 L

Luria Broth (LB) medium
- Tryptone: 10 g
- Yeast extract: 5 g
- NaCl: 10 g
- Distilled water: up to 1 L

3.1.2. *Saccharomyces cerevisiae* BY4742 culture media

Yeast Extract Peptone Dextrose (YEPD/YPD) liquid medium
- Peptone: 20 g
- Yeast Extract: 10 g
- Glucose: 20 g
- Distilled water: up to 1 L

Yeast Extract Peptone Dextrose (YEPD/YPD) agar medium
- Peptone: 20 g
- Yeast Extract: 10 g
- Glucose: 20 g
- Agar: 15 g
- Distilled water: up to 1 L

Y1376 liquid medium (Yeast Synthetic Drop-out liquid medium) (without leucine)
- Yeast Synthetic Drop-out Medium Supplement (Y1376): 1.62 g
- Yeast nitrogen base without amino acids: 6.7 g
- Distilled water: up to 1 L

- After autoclaving, add: 40 mL/L of 50% (w/v) glucose solution, sterilized through a 0.2 µm filter.

Y1376 agar medium (Yeast Synthetic Drop-out agar medium) (without leucine)
- Yeast Synthetic Drop-out Medium Supplement (Y1376): 1.62 g
- Yeast nitrogen base without amino acids: 6.7 g
- Bacteriological agar: 20 g

• Distilled water: up to 1 L

• After autoclaving, add: 40 mL/L of 50% (w/v) glucose solution, sterilized through a 0.2 µm filter.

3.1.3. *Streptomyces* strains culture media

2XYT medium
• Tryptone: 16 g
• Yeast extract: 10 g
• NaCl: 5 g
• Distilled water: up to 1 L

Soy Flour Mannitol (SFM) Agar medium
• D-Mannitol: 20 g
• Soy Flour: 20 g
• Agar: 20 g
• Tap water: 1 L

3.1.4. Fermentation media for *Streptomyces* and *Amycolatopsis* strains

NL111V medium (pH 7.2)
• Beef extract powder: 10 g
• Calcium carbonate ($CaCO_3$): 5 g
• Malt extract: 50 g
• Tap water: 0.5 L

SM17 medium (pH 6.4)
• Glucose: 2 g
• Glycerol: 40 g
• Starch: 2 g
• Soy flour: 5 g
• Peptone: 5 g
• Yeast extract: 5 g
• NaCl: 5 g
• $CaCO_3$: 2 g
• Tap water: 1 L

3.1.5. Seeding media for *Streptomyces* and *Amycolatopsis* strains

***Amycolatopsis* seeding medium** (used for: *A. japonicum* and *S. albus*)
• Glucose: 7.5 g
• Glycerol: 7.5 g
• Soya peptone: 7.5 g
• NaCl: 1.5 g
• Yeast extract: 2.5 g
• Distilled water: 0.5 L

YEME medium (used for: *S. coelicolor*)
- Yeast extract: 1.5 g
- Peptone: 2.5 g
- Malt extract: 1.5 g
- Glucose: 5 g
- Sucrose: 170 g
- Distilled water: up to 0.5 L

- After autoclaving, add 1 mL $MgCl_2$ x 6 H_2O (2.5 M).

3.2. Antibiotic stock solutions

Table 1. Antibiotic stock solutions used for different organisms during the project.
Kanamycin and nalidixic acid stock solutions were filtered using a sterile filter with 0.2 μm pore size. All solutions were stored at -20°C and handled on ice.

Antibiotic (Abbreviation)	Solvent	Concentration of stock solution (mg/mL)	Concentration used for *E. coli* strains (μg/mL)	Concentration used for *Streptomyces* and *Amycolatopsis* strains (μg/mL)
Chloramphenicol (Clm)	Ethanol 96%	30	15	-
Kanamycin (Km)	H_2O	50	50	50
Nalidixic acid (Nal)	0.1 M NaOH	30	30	30

3.3. Solutions

3.3.1. Solutions for fosmids isolation

Solution I (pH 8)
- 50 mM Tris-HCl (pH 8): 1.51425 g
- 10 mM EDTA (ethylenediaminetetraacetic acid): 0.73061 g
- Distilled water: 250 mL
- RNase (add 1 μL fresh RNase per 100 μL Solution I before each use)

Solution II
- 200 mM NaOH (sodium hydroxide): 1.99985 g
- 1 % SDS (sodium dodecyl sulfate): 2.5 g
- Distilled water: 250 mL

Solution III (pH 5.5)
- 3M potassium acetate (pH 5.5): 38.2785 g
- Distilled water: 130 mL

*Solutions I and III were kept at 4°C, while Solution II was stored at room temperature.

3.3.2. Solution for glycerol stock preparation

Glycerol 20 %
- Glycerol: 200 mL
- Distilled water: up to 1000 mL

3.3.3. Solution for colony PCR of *S. cerevisiae* transformants

Sodium hydroxide (NaOH) 20 mM
- NaOH: 0.08 g
- Distilled water: up to 100 mL

3.3.4. Solution for growth induction

L-arabinose 1 M
- L-arabinose: 3 g
- Distilled water: up to 20 mL

3.3.5. Solution for DNA rehydration

Tris-HCl 10 mM (pH 8)
- Tris-HCl (1M): 20 µL
- Distilled water: up to 2 mL

3.3.6. Solutions for yeast transformation

Lithium acetate 1 M
- Lithium acetate dihydrate: 5.101 g
- Distilled water: up to 50 mL

PEG solution 50 %
- Polyethylene glycol (MW: 3350): 100 mL
- Distilled water: 100 mL

3.4. Gel electrophoresis material

TBE buffer (10X)
- Tris base: 108 g
- Boric acid: 55 g
- 0.5 M EDTA (pH 8.0): 40 mL
- Distilled water: 900 mL

Agarose gels
During the project, different agarose concentrations (0.5 % and 0.8 % w/v) were used to prepare gels. Agarose was dissolved in TBE buffer (1X) and 2.5 µL of GelRed Nucleic Acid Stain (10.000X)/50 mL was added to visualize the bands under ultraviolet light.

3.5. Strains

Table 2. Different microbial strains used for molecular biology and as heterologous hosts.

Strain	Purpose of use and characteristics	Growth conditions and media
***Escherichia coli* EPI300**	• Creation of genomic library • Chemically competent cells of the strain used for capture vector construction and production of its higher amounts • Electrocompetent cells of the strain used to transform pYES-C1.8BGC construct and produce its higher amounts	37°C overnight, LB medium
***Escherichia coli* ET12567**	• Electrocompetent cells of the strain used to transform pYES-C1.8BGC construct and produce unmethylated DNA (-*dcm*) • *E. coli* ET12567 (pUB307) cells used to transfer vectors carrying BGC into a heterologous host (tri-parental conjugation)	37°C overnight, LB medium
***Saccharomyces cerevisiae* BY4742**	• Yeast-based assembly of capture vector and BGC 1.8 (selection marker: *leu2*Δ0)	30°C, YPD medium for wildtype and Y1376 synthetic drop-out medium for engineered strain
***Streptomyces coelicolor* M1154**	• Engineered host for heterologous expression and production of secondary metabolites in fermentation process (Δ*act*, Δ*red*, Δ*cpk*, Δ*cda* and point mutations in *rpoB* and *rpsI*)	• 30°C, SFM agar medium (culture medium) • 28°C, YEME medium (seeding medium)
***Streptomyces albus* J1074**	• Wildtype host for heterologous expression and production of secondary metabolites in fermentation process	• 30°C, SFM agar medium (culture medium) • 28°C, *Amycolatopsis* medium (seeding medium)
Amycolatopsis japonicum	• Wildtype host for heterologous expression and production of secondary metabolites in fermentation process	• 30°C, SFM agar medium (culture medium) • 28°C, *Amycolatopsis* medium (seeding medium)

3.6. Vectors

Table 3. Vectors used during the project for different purposes. *Selection markers are highlighted in bold/italic.

Vector name	Size (bp)	Characteristics	Purpose
pCC1FOS (Epicentre)	8139	***CmR***, *redF*, *ori2*, *oriV*, *repE*, *parA*, *parB*, *parC*, cos, loxP, *lacZ*	Fosmid used for the genomic library construction
pYES (not published, created by MSc. Dr. Olha Schneider by modifying the pCGW vector (54))	15240	***LEU2***, CEN6-ARS4, *aac(3)IV*, *sopA*, *sopB*, *sopC*, *incC*, *oriV*, *ori2*, *repE*, cat promoter, ***cat***, ***aph(3´)-II***, *incP*, *traJ*, attP, phage phi-C31 integrase	Plasmid used as a capture vector for BGC 1.8 assembly

3.7. Primers

Table 4. Primers used during the project for different purposes. *Endonuclease restriction sites within the sequences are marked in bold/italic.

Primer name	Sequence	Product size (bp)	Used annealing T (°C)	Function
pCC1FOS-FP	GGATGTGCTGCAAGGCGATTAAGTTGG			Sequencing of pCC1FOS fosmid
pCC1FOS-RP	CTCGTATGTTGTGTGGAATTGTGAGC			
RLA083-1.8-A-Fw	AAGAACCGGCTCATGAAC	885	63	Screening of the *Micromonospora* sp. RLA083 genomic library for BGC 1.8
RLA083-1.8-A-Rv	TACGTGTACGCGTTCGAC			
RLA083-1.8-B-Fw	ACACCTTGATCCGATACG	804	61	
RLA083-1.8-B-Rv	AAGAATCGTCGCCGAATC			
RLA083-1.8-C-Fw	GTCGATGTGGACCGATTG	637	63	
RLA083-1.8-C-Rv	GTTCATCGTCGCCATCTG			
RLA083-1.8-D-Fw	GCTGCTGGTCGAATTCTC	969	63	
RLA083-1.8-D-Rv	GACGAGCATCGAGTAGTC			
C1.8-Lflank-Fw	GTCA***GCATGC***TTGATGGTGGCGGTGACG	606	64	Creation of flanking regions for pYES-based capture vector
C1.8-Lflank-Rv	GTCA***GTTTAAAC***GGATGGTCATCCGATCTC			
C1.8-Rflank-Fw	GTCA***GTTTAAAC***CAGGAGCTCCGGCATGAC	686	66	
C1.8-Rflank-Rv	GTCA***GACGTC***ATCACGCAGCAGGTTCAG			
RLA083-1.12-A-Fw	ACGATGACCGTGGTCAAC	930	66	Screening of the *Micromonospora* sp. RLA083 genomic library for BGC 1.12
RLA083-1.12-A-Rv	ACCATGATGCCCTTCCTG			
RLA083-1.12-B-Fw	GGCATGCGAGAAAGTGAG	836	66	
RLA083-1.12-B-Rv	ATCCAGGCGTTCAACTCC			
RLA083-1.12-C-Fw	AGAACGTGCGGCAACATC	712	66	
RLA083-1.12-C-Rv	ATGGCGGAGTACATGACC			
RLA083-1.12-D-Fw	TTCGACCTCATCGACAAC	970	61	
RLA083-1.12-D-Rv	ACCGCGAAGGAGAAGATG			
RLA083-1.12-E-Fw	ACGACCCGTGATGACAAC	722	63	
RLA083-1.12-E-Rv	GAGTTGGACTCCCAGTTG			

3.8. Enzymes

Table 5. List of enzymes used during the project.

Restriction enzymes	Used for:	Expected fragment sizes:
***Nde*I, *Eco*RI – HF, *Not*I – HF**	Final confirmation of sequencing results accuracy for pCC1FOS-C1.8	***Nde*I**: 18.3 kb, 14.4 kb, 9.1 kb and 3.2 kb ***Eco*RI – HF**: 32.3 kb, 11.1 kb and 1.6 kb ***Not*I – HF**: 20.8 kb, 11 kb, 7.5 kb and 5.7 kb
***Sph*I, *Pme*I**	Digestion of the left capture vector flank	589 bp
***Pme*I, *Aat*II**	Digestion of the right capture vector flank	673 bp
***Sph*I, *Aat*II**	Digestion of the pYES plasmid for capture vector creation	15.217 bp
***Hind*III, *Bam*HI**	Checking of capture vector construct's accuracy	13.4 kb, 2.06 kb and 969 bp
***Pme*I**	Linearization of the capture vector	16.479 bp
***Xba*I, *Spe*I – HF**	Releasing of BGC 1.8. from pCC1FOS vector	38.820 bp

***Nde*I, *Eco*RI – HF, *Nco*I**	Checking of the assembled capture vector with BGC 1.8	***Nde*I**: 33.8 kb and 18.3 kb ***Eco*RI – HF**: 37.3 kb, 13.1 kb and 1.6 kb ***Nco*I**: 22.3 kb, 6.5 kb, 6.2 kb, 4.5 kb, 4.0 kb, 3.4 kb, 2.8 kb, 1.6 kb and 560 bp
Other enzymes	**Used for:**	
RNase A	RNA digestion for sample purification	
Taq Polymerase	PCR	
T4 DNA Ligase	Ligation reaction	

*All enzymes are from New England Biolabs® company.

3.9. Commercial kits used during the project

Kit for gDNA extraction
Wizard® Genomic DNA Purification Kit – Quick protocol from Promega

Kit for the construction of genomic library
CopyControl Fosmid Library Production Kit with pCC1FOS Vector (Epicentre)

Kit for plasmid isolation
Wizard® Plus SV Minipreps DNA Purification System Kit (Promega)

Kit for the isolation and purification of DNA from PCR reactions
Monarch® PCR and DNA Cleanup Kit (5 µg) (New England BioLabs)

Kit for the isolation and purification of DNA from agarose gel
Monarch® DNA Gel Extraction Kit (New England BioLabs)

4. Methods

4.1. *In silico* methods

4.1.1. antiSMASH

In this project, genome sequences were analyzed by antiSMASH (55). The online version of antiSMASH 6.1.1. (https://antismash.secondarymetabolites.org) was used to identify the BGCs for secondary metabolites and their potential functions by uploading the strain's genome sequence to the server. NCBI website (56) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was utilized to perform BLAST searches to validate BGCs and core gene software predictions.

4.1.2. Primer design

All primers for this project were designed by Dr. Jaime Felipe Guerrero Garzón, BSc MSc, using the Clone Manager software and synthesized by Eurofins Genomics company.

4.2. General methods

4.2.1. Touchdown PCR

The PCR method of choice during this project was the touchdown PCR (Table 6). In each cycle, the annealing temperature decreased by 1°C until the optimal annealing temperature was reached. The next 27 cycles were performed using the optimal annealing temperature, calculated for every primer individually with BioLab's program Tm calculator (Table 7). All the reactions were carried out using the Eppendorf® Master Cycler Nexus X2 thermocycler.

Table 6. PCR Master Mix of components for performing a single PCR reaction.

Master mix **per reaction**	
Taq Polymerase Buffer (10x)	1.25 µL
dNTPs Solution Mix	0.25 µL
Taq DNA Polymerase	0.25 µL
dH_2O	9.25 µL
Primer forward (10 mM)	0.5 µL
Primer reverse (10 mM)	0.5 µL
Sum	12 µL
DNA template	0.5 µL
Volume per reaction	12.5 µL

Table 7. Settings used for touchdown PCR reactions.

Reaction phases	**Temperature (°C)**	**Time (seconds)**	**Number of cycles**
Initial denaturation	98	60	1
Denaturation	98	10	1 cycle for every 1°C temperature drop
Annealing	from 69 °C to optimal annealing T (1°C drop per cycle)	30	
Elongation	72	30	
Once the optimal annealing T was reached, the following settings were applied:			

Denaturation	98	10	27
Annealing	optimal annealing T	30	
Elongation	72	30	
Final extension	72	600	1
Hold	4	indefinite	1

4.2.1.1. Primer preparation

The amount of distilled water that needed to be added to each primer to reach the concentration of 100 mM was listed in a table delivered together with the primers. To achieve the final concentration of 10 mM required for PCR reactions, 1:10 dilutions were prepared. The primer dilutions were stored at -20°C until needed. Before being used in certain steps of the project, all primers were tested by performing touchdown PCR and gel electrophoresis. These tests were performed to check whether the product sizes and annealing temperatures matched the software-calculated values.

4.2.2. Restriction assays

Restriction assays were used at various steps of this project, and different endonuclease enzymes were employed for cutting DNA at specific sites (Table 5 and Table 8). These reactions were performed to confirm sequencing results, cut the flanks and plasmids to be complementary for ligation, linearize vectors and release DNA fragments from constructs for assembly reactions, and test the accuracy of the assembled constructs. The restriction digestion reactions were incubated in a water bath, with the temperature of 37°C and incubation time set from 1 to 4 hours. The results were visualized using gel electrophoresis to verify whether the size of the restriction-digested DNA fragments matched the expected values.

Table 8. Restriction assay reaction mix.

Reagent	**Volume (µL)**
DNA sample	5
Restriction enzyme 1	1
Restriction enzyme 2	1
Buffer 10X (rCutSmart)	2
dH2O	up to 20

4.2.3. Gel Electrophoresis

Agarose gel electrophoresis was the method of choice for many steps throughout the project, such as: confirming if the isolation of the desirable concentration of DNA was successful, checking the DNA's integrity, visualizing the PCR, ligation, and restriction digestion results, and testing the success of a yeast-based DNA assembly. The agarose gels were prepared as described in the materials section 3.4. Gel electrophoresis material and the samples were mixed with Blue Loading Dye 10X from Thermo Fisher Scientific before being applied onto the gel. The Quick-Load 1 kb Extend DNA ladder from New England Biolabs® was used as a molecular weight marker to estimate the concentration and size of the tested DNA fragments in the agarose gel. The gels were electrophoresed at 110 V until the desired separation of the bands was observed, or at 35 V if the gel was run overnight.

4.2.4. Preparation of microbial cell glycerol stocks

1.5 mL was taken from an overnight culture and pipetted into a clean 1.5 mL Eppendorf® tube, which was then centrifuged for 5 minutes at 13000 rpm. The resulting supernatant was decanted, while the pelleted cells were resuspended in 1 mL of sterile 20 % glycerol. The mixture was gently pipetted to mix, and the resuspended cells in glycerol were then transferred into a cryotube, in which they were frozen at -80°C until further use.

4.3. Creation of the genomic library

4.3.1. Isolation of genomic DNA (gDNA)

For the isolation of the gDNA from *Micromonospora* sp. RLA083, the protocol from the Wizard® Genomic DNA Purification Kit (Promega) was followed. The mentioned protocol was originally created to obtain gDNA for PCR. However, for the purposes of creating a genomic library, all the specified amounts in the protocol should be up-scaled by ten. Apart from this change to the protocol, an additional purification step was added, which meant washing the aqueous phase with 3 mL of phenol-chloroform-isoamyl alcohol (25:24:1, v/v) after protein precipitation. Instead of a Rehydration Solution, 10 mM Tris-HCl (pH 8) buffer was used to rehydrate the DNA overnight at 4°C.

4.3.2. Creation steps of the genomic library

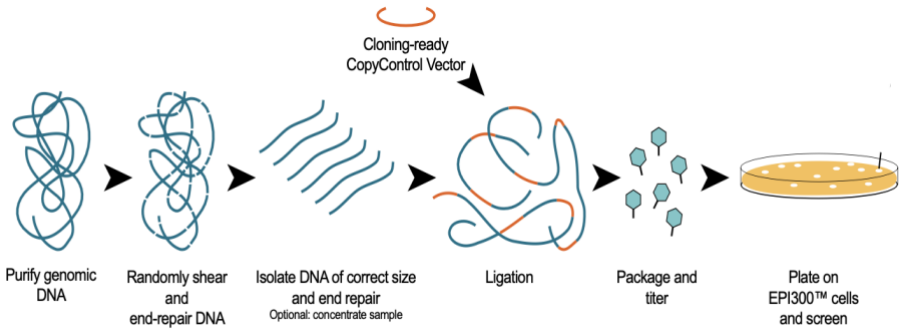

Figure 9. The single steps of the creation process of the genomic library (57).

The genomic library of *Micromonospora* sp. RLA083 was constructed using the CopyControl Fosmid Library Production Kit with pCC1FOS Vector (Epicentre). The kit's protocol was followed as suggested in Figure 9, with certain changes to specific steps.
The gDNA was isolated and purified as described in the previous section.
The shearing of the gDNA to create approximately 40 kb fragments was unnecessary, since the protocol required at least 10 % of the gDNA to co-migrate with the Fosmid Control DNA in the agarose gel to be deemed suitable for further use.
Using a formula provided in the kit's protocol, it was estimated that approximately 1000 clones were needed to achieve a 99 % likelihood of a certain DNA sequence being present in the fosmid library containing 40 kb inserts.
The changes to the protocol also included different dilutions used for titering the packaged fosmid clones, as well as the choice of medium and antibiotic concentration for this step. To ensure the growth of single colonies, *E. coli* EPI300 cells were not only mixed with undiluted packaged phage particles, but also with the following dilutions:
1. 10 µL of packaged phage particles diluted in 90 µL Phage Dilution Buffer (PDB),
2. 10 µL of the 1. dilution further diluted in 990 µL PDB,

3. 100 µL of the 2. dilution further diluted in 900 µL PDB.
The mixtures were pipetted onto LA medium plates containing Clm_{15} (15 µg/mL chloramphenicol). Once the colonies grew, the optimal titer of the packaged phage particles was determined to achieve the needed clone number for the creation of the genomic library. All colonies were picked using a pipette tip and transferred into the wells of eleven 96-well plates containing LB medium with a respective antibiotic. For the creation of the backup library, colonies from the original library were inoculated on Petri dishes containing LA medium with Clm_{15}. The freshly grown *E. coli* EPI300 colonies containing DNA fragments from the strain's genome were picked with the tips of a multichannel pipette and inoculated in the wells with LB medium and appropriate antibiotic of another eleven 96-well plates. Both the original and the backup library were incubated in a shaker at 37°C and 200 rpm for the colonies to grow. Once sufficient growth of the colonies was reached, 80 µL of sterile 50 % glycerol was added to each well of each 96-well plate before storing the libraries at -80°C.

4.4. PCR-based screening of the genomic library

4.4.1. Pooled PCR screening of the genomic library

Touchdown PCR was used to screen the genomic library. Fosmid DNA from library's 96-well plates was isolated and purified as described in 4.4.2. Each plate of the genomic library was screened for the presence of the desired biosynthetic gene cluster. Once a signal for this DNA sequence was detected on a specific plate, this 96-well plate was further screened to identify the exact location of the BGC's fragment within it. This was possible by dividing the identified plate into smaller segments and testing each of them for a positive signal. First, the quarters of the plate, then the identified quarter's rows or columns, and finally the row's or column's single clones were screened for the BGC's presence (Figure 10). For the screening of the library, the strain's gDNA diluted with distilled water (1:50) served as a positive control, while pure distilled water was used as a negative control. All PCR results were visualized by performing gel electrophoresis, where the bands served as an indication of a successful reaction.

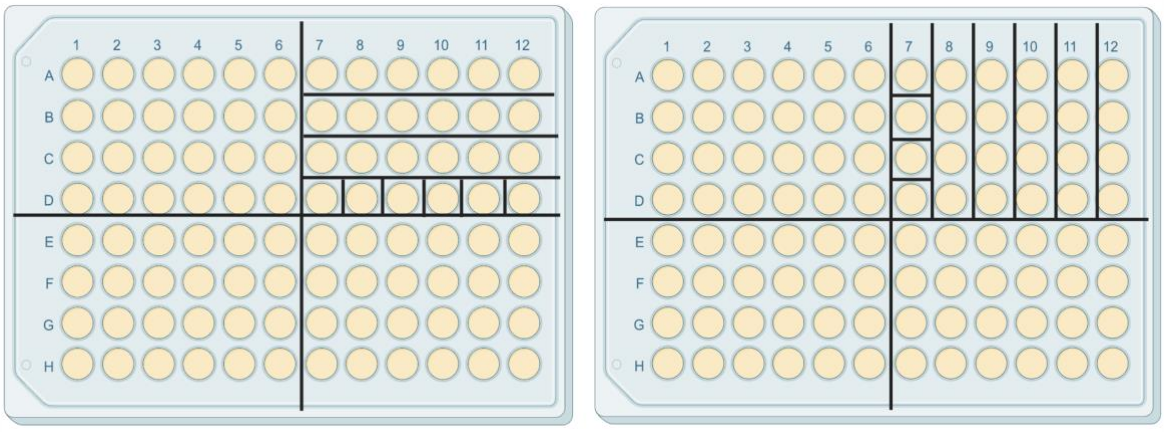

Figure 10. Division of the 96-well plate to determine the BGC's exact location within the plate using the pooled PCR screening method.

4.4.2. Isolation of fosmid DNA from *E. coli* cells

The fosmid DNA isolation from *E. coli* cells was performed by following the three-solution protocol by Birnboim and Doly (58) with slight modifications to the original steps. The following subsections describe the precise isolation procedures from solid and liquid cultures.

4.4.2.1. Fosmid DNA isolation from *E. coli* cells grown on solid medium

2 mL of sterile LB medium was pipetted onto freshly grown colonies, which were then gently washed down using pipette tips. The Petri dish was held at an angle so the detached cells could be suspended in the LB medium at the bottom. Once the cell suspension was transferred into fresh tubes, the centrifugation for 5 minutes at 13000 rpm was performed, after which the resulting supernatant was removed. In the next step, 200 µL of Solution I (3.3.1.), which had been previously stored in the fridge and mixed with 2 µL of RNase (10 mg/mL) right before use, was applied to every tube. The pellet was resuspended using a vortex. 400 µL of Solution II (3.3.1.) stored at room temperature was added to resuspended contents in each tube. The tubes were inverted until the cloudy mixture became transparent. Then, 300 µL of ice-cold Solution III (3.3.1.) was pipetted into the tubes, which were then inverted until a white precipitate formed. After centrifuging the tubes for 5 minutes at 13200 rpm, the supernatant, which contained the fosmid DNA, was moved to a clean 2 mL Eppendorf® tube. 1 mL of ice-cold isopropanol was added to achieve fosmid DNA's precipitation, followed by the tubes' inversion several times until a visible smear was noticeable. After a 10-minute incubation of tubes on ice, they were centrifuged again for 10 minutes at 132000 rpm. The pelleted fosmid DNA was washed with 200 µL of ice-cold 70 % ethanol by gently pipetting, followed by another centrifugation for 5 minutes at 13200 rpm. The alcohol was carefully decanted, and the tubes containing precipitated DNA were kept open in the laminar airflow hood until the ethanol evaporated entirely. 100 µL of distilled water was used to resuspend the air-dried pellet and the resuspended DNA was stored overnight at 4°C to dissolve. The next day, the tubes were moved to the -20°C fridge.

4.4.2.2. Fosmid DNA isolation from *E. coli* cells grown in liquid medium

2 mL of each overnight culture was transferred into clean 2 mL Eppendorf® tubes. They were centrifuged for 1 minute at 13000 rpm, followed by a supernatant removal. These steps were repeated as long as there was remaining overnight culture to be pelleted. Once the cells were pelleted, the same isolation procedure as the one described in 4.4.2.1. was followed.

4.5. DNA sequencing

All the DNA sequencing for this project was carried out by the company Eurofins Genomics. The samples were prepared following the instructions in the company's manual, which also indicated the required sample concentration. All the samples were labeled with stickers carrying a unique QR code to avoid any possible confusion.

4.6. Capture vector design

4.6.1. Preparation of the capture vector components

For the isolation and purification of the pYES plasmid from *E. coli* EPI300 cells, the protocol from the Wizard® Plus SV Minipreps DNA Purification System Kit (Promega) was followed. Tuchdown PCR was used to amplify DNA fragments for the construction of the capture vector's flanking regions. The results were visualized by performing gel electrophoresis, after which the protocol from the Monarch® PCR and DNA Cleanup Kit (5 µg) (New England BioLabs) was followed to isolate and purify the flanks from the PCR reaction mix. Restriction reaction with endonuclease enzymes was used to linearize the pYES plasmid and cut the capture vector flanks at specific sites (Table 5). The restriction assay products were run in the agarose gel to verify the success of the digestion reactions. Gel slices containing digested fragments were cut using a knife and following the protocol from the Monarch® DNA Gel Extraction Kit (New England BioLabs), both the pYES plasmid and the flanks were isolated and purified from the agarose gel slices for the ligation reaction.

4.6.2. Ligation

The ligation reaction was used to join the pYES plasmid with the right and left flanks at their complementary sites, forming a capture vector. The reaction, catalyzed by the T4 DNA Ligase, was performed in the T4 DNA Ligase Buffer (10X), and the reaction mix was incubated for 1 hour at room temperature.

Table 9. Ligation reaction with T4 DNA Ligase components mix.

Reagent	**Volume (µL)**
T4 DNA Ligase	1
T4 DNA Ligase Buffer (10X)	2
Digested pYES vector	3
Digested DNA fragment 1 (left flank)	7
Digested DNA fragment 2 (right flank)	7
Total volume	20

4.6.3. Transformation in *E. coli* EPI300 chemically competent cells

The chemically competent *E. coli* EPI300 cells were prepared by Mag. Silvia Löbsch. After 15 minutes of thawing the cells on ice, 5 µL of the ligation reaction was added to the thawed cells and mixed by gently pipetting. The empty circular pYES plasmid was used as a positive control, whereas the linear pYES plasmid served as a negative control. After a half-hour incubation period on ice, the samples were exposed to 42°C for 45 seconds in a water bath. The heat-shocked cells were then allowed to cool in an ice box for 5 minutes. Working in a sterile hood, 950 µL of LB medium was applied to cooled tubes. The contents were then pipetted 2-3 times up and down to mix and finally incubated in a shaker with the following settings: 1 hour, 37°C, and 200 rpm. The cell suspension in LB medium was then spread over Petri dishes containing LA medium with Km_{50} (50 µg/mL kanamycin) using an inoculation spreader. The inoculated plates were incubated overnight at 37°C.

4.6.4. Plasmid DNA isolation from *E. coli* cells

Freshly grown colonies from the overnight plates were picked and transferred into a 5 mL LB medium, previously supplemented with Km_{50} and 1 M arabinose. Shaking it at 200 rpm, the

culture was incubated overnight at 37°C. The Wizard® Plus SV Minipreps DNA Purification System Kit (Promega) protocol was followed to isolate and purify the construct from the liquid cell culture.

4.7. TAR cloning in yeast

4.7.1. Plasmid linearization and BGC isolation

The capture vector was cut with an endonuclease enzyme to be linearized (Table 5). At the same time, the fosmids containing the BGC were also digested with appropriate enzymes (Table 5) to release the BGC from the construct, as described in 4.2.2. Once the restriction assay results were visualized by performing gel electrophoresis on a 0.5 % agarose gel, the digested fragments were isolated from the gel following the steps specified in the subsection 4.7.1.1.

4.7.1.1. Isolation and purification of digested DNA from 0.5 % agarose gel

Once the digested DNA fragments were separated in the gel and the gel slices containing those fragments excised, this protocol described by MSc Sonali Vaidya (59), with slight modifications, was applied to recover the DNA fragments from 0.5 % agarose gel. The excised gel slice was transferred into a tarred 2 mL Eppendorf® tube, chopped using an inoculation loop, and weighed. The gel slice weight was essential to determine the needed volumes of the isolation components. One volume of phenol was added to the tube's contents, which were mixed by vortexing for approximately 10 seconds until the formation of a milky-white mixture was visible. Right after that, the tube was exposed to -80°C for 20 minutes and then centrifuged at room temperature with the following settings: 15 minutes, 13600 rpm. The resulting upper phase, which contained the DNA fragments, was moved to a new 2 mL Eppendorf® tube. 0.1 volume of Neutralization Solution from the Wizard® Plus SV Minipreps DNA Purification System Kit (Promega) and one volume of phenol/chloroform were added to the tube's contents. The contents were then gently mixed by inversion, specifically avoiding the use of vortex, and finally spun in a centrifuge with the following settings: 15 minutes, 13600 rpm, and 4°C. Once again, the upper phase was transferred to a new tube, to which 0.1 volume of 3 M sodium acetate (pH 7) was added and mixed. The next addition to the tube's contents meant 2.5 volumes of ethanol. Once properly mixed without vortexing, the tube was allowed to cool at -80°C for 10 minutes. The frozen tube was spun at 15000 rpm for 20 minutes in a centrifuge set to 4°C. The supernatant was carefully removed, allowing the thorough washing of the pellet through the addition of 100 µL of ice-cold 70 % ethanol. Another centrifugation with the same settings was performed, after which the ethanol was decanted. The pelleted DNA fragments were washed and centrifuged one more time. After the final centrifugation of the tube, the alcohol used for the washing step was removed, allowing the pellet to get air-dried under laminar airflow. Finally, the DNA fragments were resuspended in 15 µL of Nuclease Free Water from Wizard® Plus SV Minipreps DNA Purification System Kit (Promega), and the sample was kept at 4°C overnight.

4.7.2. Yeast transformation

As outlined by MSc Sonali Vaidya (59), the yeast transformation was performed with minor changes as follows: YPD agar plates (3.1.2.) were inoculated with *Saccharomyces cerevisiae* (*S. cerevisiae*) BY4742 culture (Table 2). Plates were incubated at 30°C for 3 days. Single strain colonies were inoculated in 10 mL of liquid YPD medium (3.1.2.). The liquid yeast

culture was grown at 30°C overnight. The next day, 2.5 mL of the freshly grown yeast cells was pipetted into 50 mL of the liquid YPD medium. The new liquid culture was incubated in a shaker (200 rpm, 30°C) for approximately 4 hours until the optical density, measured with a UV/Vis spectrophotometer, showed sufficient growth of the cells (OD_{600} between 0.85 and 0.9).

In the next step, the master mix for transformation was prepared (Table 10). The SS DNA (single stranded salmon sperm DNA in water) was allowed to defrost at room temperature before being heated at 100°C for 10 minutes using a heating block. After exposing it to heat, the SS DNA was cooled on ice for 10 minutes.

Table 10. Components of the transformation master mix.

Components per reaction	Volume in µL
50 % PEG (MW: 3350)	240
1M lithium acetate	60
SS DNA	50

Once the *S. cerevisiae* BY4742 cell growth reached the desirable OD_{600}, the liquid culture was moved to a 50 mL Falcon® tube to be centrifuged at 6000 rpm and 4°C for 5 minutes. Next, the supernatant was decanted, allowing the pelleted cells to be resuspended in 2.5 mL of sterile water. 500 µL of the resuspended cells was transferred to each of five fresh 1.5 mL Eppendorf® tubes, and the tubes were centrifuged at 6000 rpm for 1 minute. After removing the supernatant, 400 µL of sterile 100 mM lithium acetate solution was used to resuspend the resulting cell pellet by pipetting. Another centrifugation with the same settings followed, after which the supernatant was removed, and the pelleted cells were kept in an ice box until the next step.

Table 11. Yeast-based DNA assembly reaction components.

Reaction components	Volume in µL
Linear capture vector	2
BGC fragment	5
Distilled water	up to 36
Negative control	
Linear capture vector	2
Distilled water	up to 36
Positive control	
Circular capture vector	2
Distilled water	up to 36

Linearized capture vector and the BGC fragment were introduced into *S. cerevisiae* BY4742 cells to perform a yeast-based DNA assembly, implementing the following steps.
350 µL of Transformation Master Mix (Table 10) was added to each tube containing yeast cell pellet, and it was crucial not to mix them. Next, the DNA Assembly Reaction Mix (Table 11) was added to the tubes. For proper mixing of all components, the tubes were vortexed for 10-20 seconds and then incubated for 45 minutes at 42°C using a heating block. After incubation, the tubes were spun for 1 minute at 6000 rpm, and the formed thick supernatant was removed. 500 µL of the liquid Y1376 drop-out medium (3.1.2.) was added to each tube, and the pellets were resuspended by gently pipetting. 500 µL of each resuspended sample was then plated on solid Y1376 agar plates (3.1.2.), and the plates were left in an incubator for two days at 30°C.

4.7.3. Colony PCR of yeast transformants

To confirm that the DNA assembly reaction was successful, colony PCR of the freshly grown *S. cerevisiae* BY4742 transformants was performed following the slightly changed version of the method outlined by MSc Sonali Vaidya (59). Twenty transformant clones, along with one positive and one negative control clone, were picked using a sterile toothpick and inoculated on Y1376 drop-out medium (3.1.2.) plates. The plates were incubated at 30°C for three days. A portion of the grown colonies was collected from the agar surface and transferred into tubes, previously supplemented with 20 µL of NaOH (20 mM) each. Employing a heating block, the tubes were incubated at 95°C for 45 minutes and then spun at 13600 rpm for 10 minutes. Colony PCR was performed using 2 µL of the supernatant from every tube as a DNA template. Afterward, the gel electrophoresis with 0.8 % agarose gel was carried out to confirm the transformation reaction's success.

4.7.4. Plasmid isolation and purification from yeast

With minor modifications, the protocol explained by MSc Sonali Vaidya (59) was applied, and 10 mL of yeast synthetic drop-out medium (3.1.2.) was inoculated with the *S. cerevisiae* clone that showed positive signals in the gel electrophoresis following colony PCR. The liquid culture was shaken at 200 rpm and 30°C overnight. The next day, 2.5 mL of the overnight culture was centrifuged for 3 minutes at 6000 rpm. The resulting supernatant was carefully removed, whereas the pelleted yeast cells were resuspended in 500 µL of 50 mM Tris-HCl (pH 8). In the next step, 50 µL of enzyme lyticase was added and mixed with the pelleted cells by inverting the tube's contents. The resulting enzyme-cell mixture was exposed to 30°C in a water bath for 45 minutes. During incubation, the tube's contents were inverted every 15 minutes to ensure proper mixing. Following incubation, the tubes were centrifuged at 10000 rpm for 5 minutes, and the resulting supernatant was immediately removed. The adapted version of the Wizard® Plus SV Minipreps DNA Purification System Kit (Promega) protocol was used for the further isolation of plasmid DNA from yeast. The resulting pellet was resuspended in 250 µL of the kit's Cell Resuspension Solution by vortexing the tube. Once the cells were resuspended, 500 µL of the kit's Cell Lysis Solution was pipetted to the tube, followed by mixing the contents by inversion and then incubating the tube at room temperature for 5 minutes. Next, 10 µL of Alkaline Protease Solution was added to the tube's contents, which were mixed and incubated under the same conditions. Finally, 700 µL of Neutralization Solution was added, after which the tube was mixed by inversion and centrifuged with the following settings: 13400 rpm, 10 minutes. The clear supernatant was then moved to a kit's spin column and centrifuged for 1 minute using the maximum speed program. To wash the pelleted capture vector construct, two cleaning steps with 750 µL of Column Wash Solution were performed, with centrifugation for 1 minute following each addition. Once the flow-through was removed, the spin column was centrifuged for another 2 minutes for the pelleted capture vector construct to dry entirely. 60 µL of Nuclease-Free Water, previously warmed at 50°C, was added to the pellet. To allow the plasmid DNA to dissolve, the contents were incubated at 37°C for 15 minutes using a heating block. The isolated capture vector construct was stored at -20°C until needed for further project steps.

4.8. Transformation of plasmid DNA in electrocompetent *E. coli* cells

4.8.1. Electrocompetent *E. coli* cell preparation

Electrocompetent cells of two *E. coli* strains, *E. coli* EPI300 and *E. coli* ET12567, were prepared following a slightly modified version of the method described by Dr. Jaime Felipe Guerrero Garzón (60):
An *E. coli* cell culture was initiated in 10 mL of LB medium supplemented with a respective antibiotic concentration, and shaken at 200 rpm and 37°C overnight. The next day, a new liquid culture was started by inoculating 100 mL of LB medium containing the same antibiotic with 1 mL of freshly grown overnight culture. The liquid culture was shaken under the same conditions as previously described for 3-4 hours until the growth of the cells reached the desired OD_{600} between 0.4 and 0.6. The optimally grown cells were then divided into two 50 mL Falcon® tubes and centrifuged with the following settings: 4000 rpm, 4°C, for 5 minutes. After centrifugation, the supernatant was removed, allowing the resuspension of the pelleted cells in 40 mL of ice-cold 10 % glycerol. The same settings (4000 rpm, 4°C, 5 minutes) were applied to centrifuge the glycerol cell suspension. The resulting supernatant was immediately removed, while the pelleted cells were resuspended in 30 mL of ice-cold 10 % glycerol. After the final centrifugation under the same conditions and removal of supernatant, the cells were resuspended in 500 µL of 10 % glycerol that remained. For the transformation process, 50 µL aliquots were prepared and stored at -80°C until needed.

4.8.2. Electrocompetent *E. coli* cell electroporation

This step of the project was performed based on the protocol described by Dr. Jaime Felipe Guerrero Garzón (60), with minor changes. 1.5 µL of vectors harboring BGCs, obtained in the DNA assembly in yeast, was added to one 50 µL aliquot of the prepared electrocompetent *E. coli* EPI300 or *E. coli* ET12567 cells. The contents were mixed carefully to avoid the formation of bubbles. 50 µL of the mixture was transferred into a 0.2 cm electroporation cuvette, which had been previously kept at -20°C. The electroporation was carried out on a BioRad GenePulser II with the following configurations: 200 Ω, 25 µF, 2.5 kV, and the expected time exposure between 4.5 and 4.9 ms. After electroporation, the shocked cells were immediately mixed with 950 µL of ice-cold LB medium by gently pipetting. The cell suspension was then shaken at 200 rpm for 2 hours at 37°C. Finally, the cells were spread onto Petri dishes with LA medium and appropriate antibiotics with an inoculation spreader and incubated overnight at 37°C.

4.9. Conjugation into *Streptomyces* and *Amycolatopsis*

4.9.1. Harvesting of spores and glycerol stock preparation

To harvest spores, the method presented in the doctoral thesis of Dr. Jaime Felipe Guerrero Garzón (60) was applied:
Once the spores grew sufficiently, using a sterile plastic pipette, 5 mL of sterile 20 % glycerol was transferred onto the plate containing spores, and its surface was scraped carefully using the tip of the 5 mL pipette. The resulting spore suspension in glycerol was then pipetted into the sterile syringes containing cotton wool as a filter to remove mycelium. Finally, aliquots of the filtrate were transferred into cryotubes and stored at -80°C.

4.9.2. Tri-parental conjugation

The capture vector harbouring BGC was conjugated from *E. coli* ET12567 cells into three other strains: *Streptomyces coelicolor* M1154, *Streptomyces albus* J1074, and *Amycolatopsis japonicum*, based on the protocol outlined by Dr. Jaime Felipe Guerrero Garzón (60). LA medium plates with kanamycin and chloramphenicol were inoculated with: first, *E. coli* ET cells carrying the helper plasmid pUB307 for conjugation between bacteria; second, *E. coli* ET cells with incorporated capture vector carrying BGC construct; and third, *E. coli* ET cells harbouring the empty pYES plasmid. The plates were incubated at 37°C overnight. The freshly grown cell colonies were scraped from the plate surface using an inoculation loop and suspended in 500 µL 2XYT medium (3.1.3.).

50 µL of the *Streptomyces* and *Amycolatopsis* glycerol spore suspension was kept on ice until thawed and then added to 350 µL of 2XYT medium in 1.5 mL Eppendorf® tubes. To induce the germination of spores, the tubes were incubated at 50°C for 10 minutes using a heating block. The heat-shocked spore suspension was left at room temperature for approximately 10 minutes to cool down.

Table 12. Conjugation reaction components for all three strains: *Streptomyces coelicolor* M1154, *Streptomyces albus* J1074, and A*mycolatopsis japonicum*.

Reaction components for each strain	**Volume (µL)**
E. coli ET/pUB307	50
E. coli ET/vector-BGC OR *E. coli* ET/empty pYES	50
Heat-shocked spore suspension	400
Total volume per reaction	500

The conjugation reaction tubes (Table 12) underwent a centrifugation at 4000 rpm for 1 minute. After the removal of 250 µL of the resulting supernatant, the remaining 250 µL of reaction volume were used to resuspend the pellet. The suspension was spread over the surface of SFM (3.1.3.) plates with $MgCl_2$ (1 mL $MgCl_2$/100 mL SFM medium) and the plates were incubated at 30°C for 16-18 hours.

4.9.3. First antibiotic selection of the transconjugants

After the 16–18-hour incubation, the first antibiotic selection of the transconjugants was performed. The antibiotic solution used for this step contained 30 µL of nalidixic acid and 30 µL kanamycin in 1 mL of distilled water. The 1 mL antibiotic solution was spread evenly over the surface of each plate, and the plates were kept in the laminar airflow hood until completely dry. The incubation of plates at 30°C was continued.

4.9.4. Second antibiotic selection of the transconjugants

Once the single colonies grew, the second selection was performed. For the second selection, the chosen colonies were picked using a sterile toothpick and spread on the LA plates containing kanamycin (50 µg/mL) and nalidixic acid (30 µg/mL). The plates were then incubated at 30°C for a couple of days. After the second selection, a part of the grown mass from each clone was picked with a sterile inoculation loop and spread evenly across the plates containing LA medium with kanamycin (50 µg/mL) and nalidixic acid (30 µg/mL). Sporulation took place in the incubator at 30°C, and once it reached the desired level, spore suspensions were prepared and stored as described in section 4.9.1.

4.10. Fermentation and methanol extraction

4.10.1. Preparation of seeding cultures

Seeding cultures for recombinant *Streptomyces* and *Amycolatopsis* strains were prepared as follows: for both *A. japonicum* and *S. albus*, 10 mL of Amycolatopsis seeding medium (3.1.5.), and for *S. coelicolor*, 10 mL of YEME medium (3.1.5.), was inoculated with 200 µL spore suspension. The autoclaved 100 mL baffled flasks containing the seeding cultures were incubated in a shaker at 28°C and 200 rpm.

4.10.2. Fermentation

Once sufficient growth of the seeding cultures was reached, fermentation cultures for each strain were started using two media: NL111V and SM17 (3.1.4.). 25 mL of each medium was inoculated with 2 mL of each seeding culture in sterile 250 mL baffled flasks, which were then incubated at 28°C and 200 rpm. Following fermentation, its products were carefully transferred into round flasks without touching the walls and in the final step freeze-dried. In this step, the samples were frozen using an ethanol bath, connected to the equipment, and left overnight to dry. After freeze-drying, the samples were stored at -20°C until further use.

4.10.3. Methanol extraction of the fermentation products

20 mL of methanol was added to every previously freeze-dried fermentation sample. To initiate the methanol extraction of the fermentation products, the samples were shaken at room temperature for at least 30 minutes. Once the extraction process was done, centrifugation of the samples at 4000 rpm for 5 minutes was performed. The 2 mL of each resulting upper phase was transferred into a clean 2 mL Eppendorf® tube. In the next step, the extracts were dried using a low boiling points method with the temperature set to 40°C, which is the evaporation temperature of methanol. Once the samples were dry, 200 µL of methanol was pipetted to each tube, concentrating the samples 10X, and the contents were mixed by gently pipetting. The final centrifugation of the samples was carried out at maximum speed for 3 minutes, after which 60 µL of each resulting upper phase was transferred into HPLC vials for methanol extracts to be analyzed by High Performance Liquid Chromatography (HPLC) and Liquid Chromatography-Mass Spectrometry (LC-MS).

4.11. Analytical methods

4.11.1. High Performance Liquid Chromatography (HPLC)

The methanol extracts of secondary metabolites produced by different recombinant strains during fermentation were analyzed using the HPLC method (Shimadzu HPLC unit) as described by Dr. Jaime Felipe Guerrero Garzón (60). Two mobile phases, aqueous 0.1 % formic acid (A) and acetonitrile (B), were utilized to carry out the reverse-phase chromatography, where the change in acetonitrile concentration was used as a gradient. The acetonitrile concentration was gradually increased from 5 % to 95 % over 45 minutes, followed by keeping it at 95 % for 10 minutes (washing step) and finally reducing it back to 5 % over 10 minutes (re-equilibration step). The Luna® C^{18} column from Phenomenex with the set temperature at 25°C served as a stationary phase. Its dimensions of 250 mm x 4.6 mm allowed the passage of particles up to 5 µm in size. The flow rate was set to 0.5 mL/min, while the detection was carried out using two types of detectors: PDA (Photodiode Array)

detector, which detected the absorption of UV light in the range of 190 to 800 nm, and ELSD-LT II detector (Evaporative Light Scattering Detector).

4.11.2. Liquid Chromatography-Mass Spectrometry (LC-MS)

The LC-MS measurements and analysis were done by Dr. Martin Zehl at the Mass Spectrometry Center, University of Vienna.
A Vanquish Horizon UHPLC system (Thermo Fisher Scientific) connected to the ESI source of a timsTOF fleX mass spectrometer (Bruker Daltonics) was used for LC-MS analysis of secondary metabolites' methanol extracts. An Acquity Premier HSS T3 column with VanGuard FIT, 2.1 x 150 mm in size, was used as a stationary phase, allowing the passage and separation of particles sized up to 1.8 µm. The temperature of the column oven was set at 40°C, while the column's flow rate was kept at 0.5 mL/min. Two mobile phases were used to perform this analysis, water as a mobile phase A, and acetonitrile/water (9:1) mixture as a mobile phase B, both modified with 0.1 % formic acid. (Table 13)

Table 13. Mobile phase gradient used for the analysis of methanol extracts with LC-MS.

Mobile phase B	Time	Process
0 % - 20 %	in 10 minutes	separation and elution of the sample components
20 % - 100 %	in 25 minutes	
100 %	for 4 minutes	isocratic column cleaning
0 %	for 6 minutes	re-equilibration

The high-resolution ESI-MS spectra were captured in the m/z range between 100 and 2500, using positive ion mode.

4.12. Bioassays

To test if the fermentation products show any antimicrobial activity, disc diffusion assays with three different strains (Table 14) were performed.

Table 14. Microorganisms used for disc diffusion assays to detect the potential bioactivity of secondary metabolites produced by recombinant strains during fermentation.

Strain	Microorganism type	Media	Incubation T (°C)
Staphylococcus carnosus	Gram-positive bacteria	LB medium	37
Pseudomonas putida	Gram-negative bacteria	LB medium	37
Saccharomyces cerevisiae	Yeast	YPD medium	30

30 µL of methanol extracts were pipetted onto sterile bioassay discs, which were then left to dry for approximately 20 minutes. Each LA medium (3.1.1.) plate was inoculated with 500 µL of an overnight culture of either *S. carnosus* or *P. putida*, while the YPD(A) medium (3.1.2.) plate was inoculated with 500 µL of *S. cerevisiae* overnight culture. Using sterile tweezers, the dried bioassay discs were transferred onto inoculated plates. The plates were then incubated overnight at the appropriate temperature for each strain (Table 14). The following day, the plates were checked for inhibition of the growth of strains in the region where the bioassay discs were placed.

5. Results

5.1. antiSMASH genome analysis

Using the open-source bioinformatics tool called antiSMASH, the genome of *Micromonospora* sp. RLA083 was analyzed for the presence of BGCs, resulting in the identification of 24 potential BGCs.

5.1.1. *Micromonospora* sp. RLA083 BGC 1.8

The BGC 1.8 was chosen to be worked with because of its low similarity of 38% to the already known BGCs, indicating its potential for specifying biosynthesis of a novel compound. antiSMASH 6.1.1. suggests that BGC 1.8 encodes the biosynthesis of a glycosylated terpene. This was another reason for choosing this BGC for the project since terpenes previously showed a wide range of biological activities, some of them beneficial for human health. The BGC 1.8 is approximately 30 kb in size, with gene products terpene cyclase, prenyl transferase, and polyprenyl synthetase, which most probably contribute to the synthesis of the core molecule. The software also suggests the presence of biosynthetic genes, such as aminotransferase, monooxygenase, dioxygenase, glycosyltransferases, which potentially add up to the structural complexity. Additionally, BGC 1.8 seems to contain the gene encoding an ABC transporter, which is usually involved in the efflux of the final compound from the cell. (Figure 11)

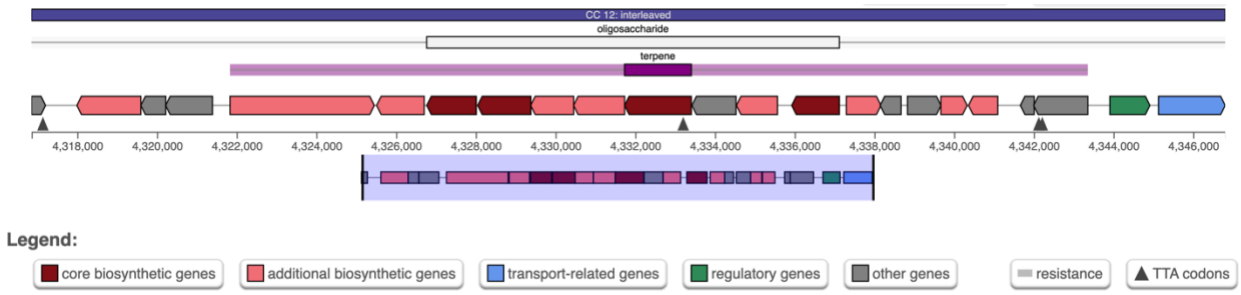

Figure 11. antiSMASH-generated overview of biosynthetic gene cluster RLA083_1.8.

5.1.2. *Micromonospora* sp. RLA083 BGC 1.12

The second cluster from *Micromonospora* sp. RLA083 chosen for this project was BGC 1.12 (size ~80kb). Its 33% similarity to known clusters also suggests a possibility of a novel compound production. antiSMASH analysis predicts that BGC 1.12 encodes a type I polyketide synthase (PKS) consisting of six modules. The initial module appears to be specific for methylmalonyl-CoA, whereas the remaining five modules are specific for malonyl-CoA. Through domain annotation analysis, it was assumed that these modules incorporate 13 C atoms into the compound's backbone, while dehydratase (DH) and ketoreductase (KR) domains in some of the modules suggest potential formation of four double bonds. Moreover, the final compound could potentially contain two nitrogen atoms due to the presence of two genes encoding aminotransferases. Additionally, two genes within the BGC 1.12 encoding ketoreductases are assumed to influence the final compound's structure, while the thioesterase (TE) domain's presence indicates that the product might undergo a cyclization. Therefore, the final product could either form a macrocycle or be partially cyclized. (Figure 12)

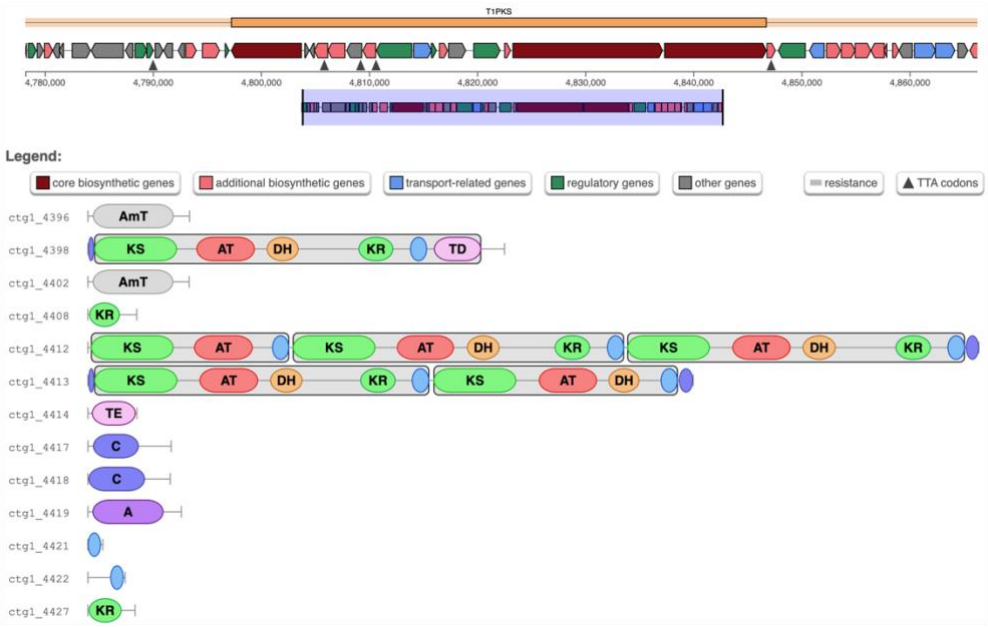

Figure 12. antiSMASH-predicted biosynthetic gene cluster RLA083_1.12 overview and PKS module annotation.

5.2. *Micromonospora* sp. RLA083 genomic library

The antiSMASH analysis of the *Micromonospora* sp. RLA083 genome showed the presence of several BGCs coding for diverse secondary metabolites, making this strain a promising candidate for genome mining. To properly examine the strain's potential for producing novel compounds via heterologous expression of identified BGCs, the fosmid genomic library was created as described in section 4.4. using the pCC1FOS vector (Figure 13). Since the pCC1FOS fosmid can incorporate DNA fragments up to approximately 40 kb, and the C1.8BGC of interest is relatively small (~30 kb), chances were high that the whole cluster would be contained within a single fosmid library clone. The library creation process yielded 1562 *E. coli* EPI300 recombinant clones, each carrying a DNA fragment from the *Micromonospora* sp. RLA083 genome. For a 99% likelihood of a certain DNA sequence being present in the fosmid library containing 40 kb inserts, it was estimated that approximately 1000 clones were required. To ensure sufficient library diversity, 1056 randomly chosen colonies were picked, resulting in the creation of eleven 96-well genomic library plates. These plates represent an extensive resource for the screening and further characterization of the strain's BGCs.

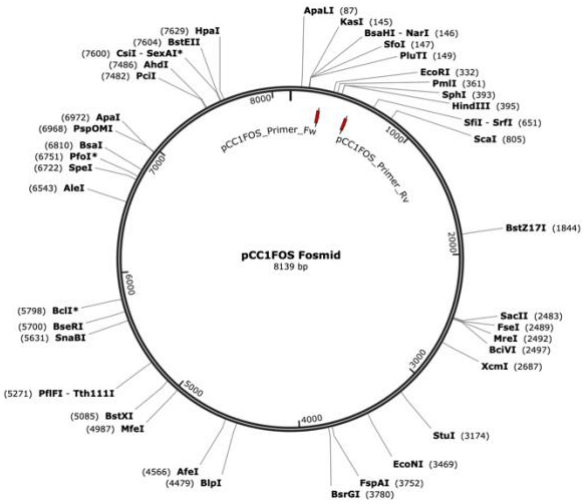

Figure 13. pCC1FOS fosmid vector (8139 bp) used to create the *Micromonospora* sp. RLA083 genomic library with its restriction and primer binding sites.

5.3. Pooled PCR screening of the genomic library

The screening of the *Micromonospora* sp. RLA083 genomic library was done in order to identify the clones that contain BGC 1.8 and BGC 1.12 fragments.

5.3.1. BGC 1.8 pooled-PCR screening

The screening for BGC 1.8 was performed in several stages, first screening all eleven genomic library plates until the one with all positive signals for the cluster was found, and then aiming to precisely locate the four fragments representing this BGC (A, B, C, and D) within the detected plate (4.4.1.). The expected PCR fragment sizes are listed in the Materials section (Table 4).

Pooled PCR of all eleven plates of the *Micromonospora* sp. RLA083 genomic library was first done with the primers amplifying fragments A and D for the BGC 1.8. Plates 2 and 9, and possibly plates 1, 3, and 10, showed positive signals for fragment A, while plates 1, 3, 5, 9, 10, and 11 showed amplifications for fragment D (Figure 14).

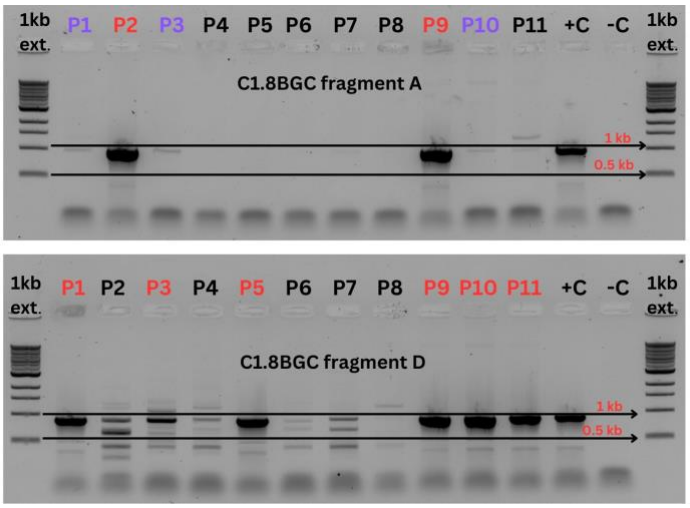

Figure 14. Gel electrophoresis showing pooled PCR screening results for eleven *Micromonospora* sp. RLA083 genomic library plates with primers amplifying fragments A and D for BGC 1.8. The plates marked in red are

those yielding amplification for either fragment A or D, whereas the purple ones showed a positive signal that cannot be taken as a correct one with high certainty.

Plates 1, 2, 5, 9, 10, and 11 were selected to be screened with primers amplifying fragments B and C for BGC 1.8. The reason for excluding other plates from screening was that they showed no amplification for either the beginning or the end of the cluster. The gel electrophoresis results showed the amplification of fragment B of the cluster in plates 1, 5, 9, 10, and 11, while fragment C was amplified using templates from plates 1, 2, 9, and 11. (Figure 15)

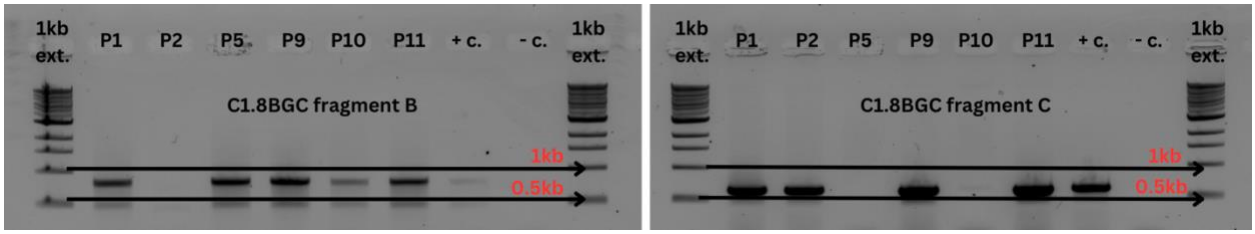

Figure 15. Gel electrophoresis showing pooled PCR screening results of the selected *Micromonospora* sp. RLA083 genetic library plates with primers B and C for BGC 1.8.

Since the only two plates giving amplifications of all four fragments of the BGC 1.8 were plates 1 and 9, the quarters of each were screened with primers amplifying fragments A, B, C, and D. Both plates showed positive signals for fragments A, B, and C. However, plate 1 did not appear to have an amplification for fragment D this time, while plate 9 had a barely visible amplification for the mentioned fragment of the cluster. All quarters of both plates were screened again with primers for fragment D to test if applying samples with a multichannel pipette could have caused too little sample concentration in the wells to be detected. Once it was confirmed that both plates had signals for fragment D, it was decided to proceed with the screening of the 2nd quarter of plate 9 (P9.2). The reason behind it is that plate 9 had all four signals in the same plate quarter in comparison to plate 1, making the possibility of a whole cluster being present in a single clone much higher (Figure 16).

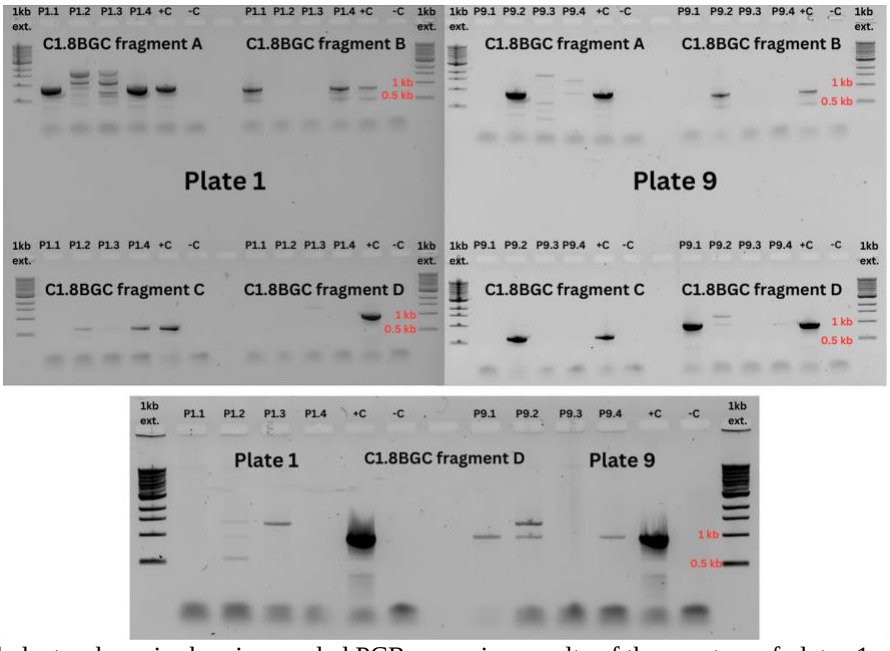

Figure 16. Gel electrophoresis showing pooled PCR screening results of the quarters of plates 1 and 9 from *Micromonospora* sp. RLA083 genomic library with four sets of primers for BGC 1.8. The two gel images above

show the screening results of the two plates lacking a (sufficient) signal for fragment D, while the gel image below shows the results of a repeated PCR reaction with primers D for all quarters of plates 1 and 9.

Once the quarter of the plate yielding all four BGC 1.8 fragments was successfully identified, the rows of the detected quarter were screened. The gel electrophoresis results led to the conclusion that all cluster fragments were contained within the 4th row of the 2nd quarter in plate 9 (P9.2.4) (Figure 17).

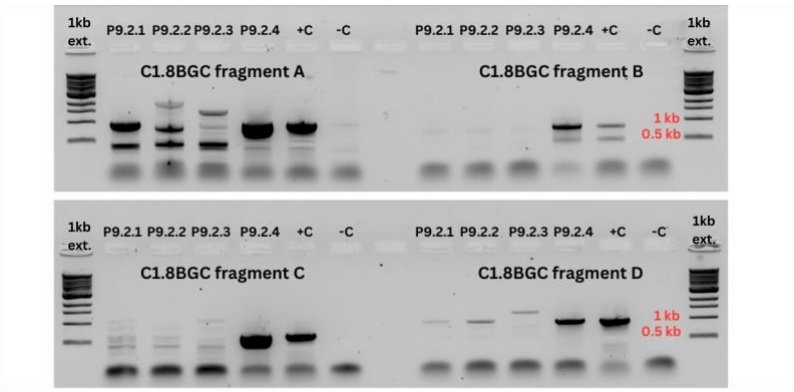

Figure 17. Gel electrophoresis showing pooled PCR screening results of the 2nd quarter's rows from plate 9 of *Micromonospora* sp. RLA083 genomic library with four sets of primers for BGC 1.8.

In the final screening step, the single clones from the 4th row of the 2nd quarter of plate 9 were screened for the presence of the cluster fragments. The gel electrophoresis results indicated a potential presence of the entire BGC 1.8 in a single clone. However, sequencing of samples 9.2.4.1, 9.2.4.2, 9.2.4.3, and 9.2.4.4 revealed no presence of fragments A and D. According to these results, during the construction of the *Micromonospora* sp. RLA083 fosmid library, cluster 1.8 was cut and integrated into the pCC1FOS fosmid without the two regions amplified by primers specific for fragments A and D. The results indicate that the integrated cluster fragment ranges from 4365th to 41396th bp of the ~47 kb BGC 1.8, corresponding to a total size of approximately 37 kb (Figure 18). The ~37 kb BGC 1.8 fragment was identified in clone 9.2.4.1, which was isolated and purified from the detected *E. coli* EPI300 clone following the protocol described in the Methods section and used for further project steps (Figure 19).

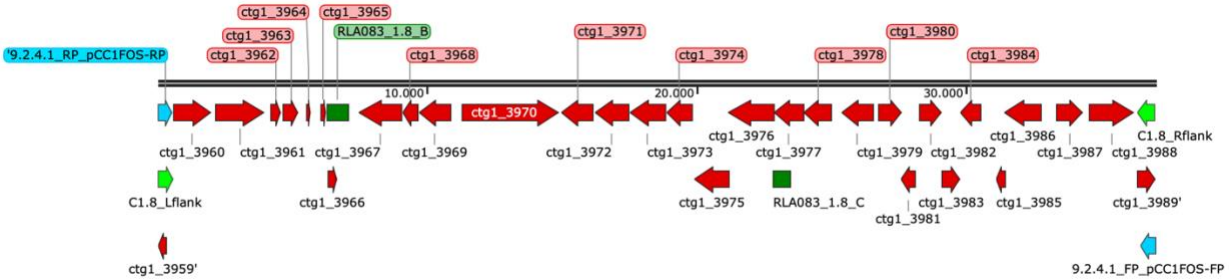

Figure 18. The *Micromonospora* sp. RLA083 BGC 1.8 cloned cluster fragment without the expected fragments A and D (37.032 bp), with integration points of the cluster fragment into the pCC1FOS fosmid highlighted in light blue and primer binding sites for B and C in dark green. Light green arrows indicate binding sites for the pYES capture vector's flanks for subsequent heterologous expression.

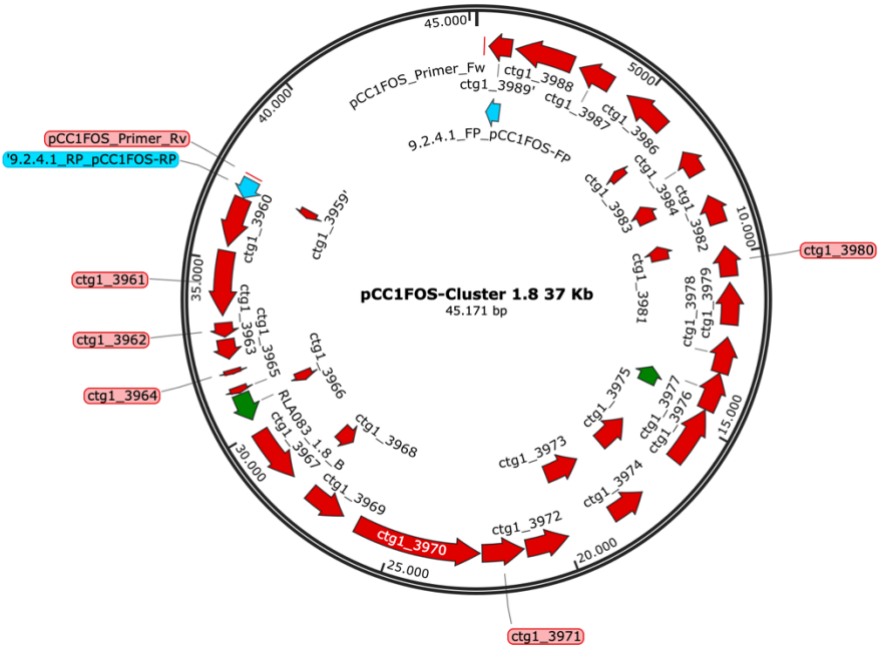

Figure 19. pCC1FOS fosmid vector (8139 bp) with cloned BGC 1.8 fragment 9.2.4.1 (37.032 bp) with a total construct size of 45.171 bp.

5.3.2. BGC 1.12 pooled-PCR screening

The screening of the *Micromonospora* sp. RLA083 genomic library for the presence of BGC 1.12 was performed with five sets of primers (A, B, C, D, and E) (Table 4). Due to the BGC's large size of ~80kb, it was expected that the cluster will be split in at least two parts during library generation, since the vector can only harbor DNA fragments of approximately 40 kb. The screening process was conducted as described in the Methods, starting with all eleven genomic library plates being tested with primers A, C, D, and E, since it was expected that fragments ABC would likely be in one piece. All fragment sizes are listed in the Materials section. Plates 2, 4, 6, 7, and 11 showed amplifications for fragment A, 1, 2, 5, 6, 9, and 10 for fragment C, 1, 3, 5, 8, 9, 10, and 11 for fragment D, and 1, 3, and 8 for fragment E. (Figure 20)

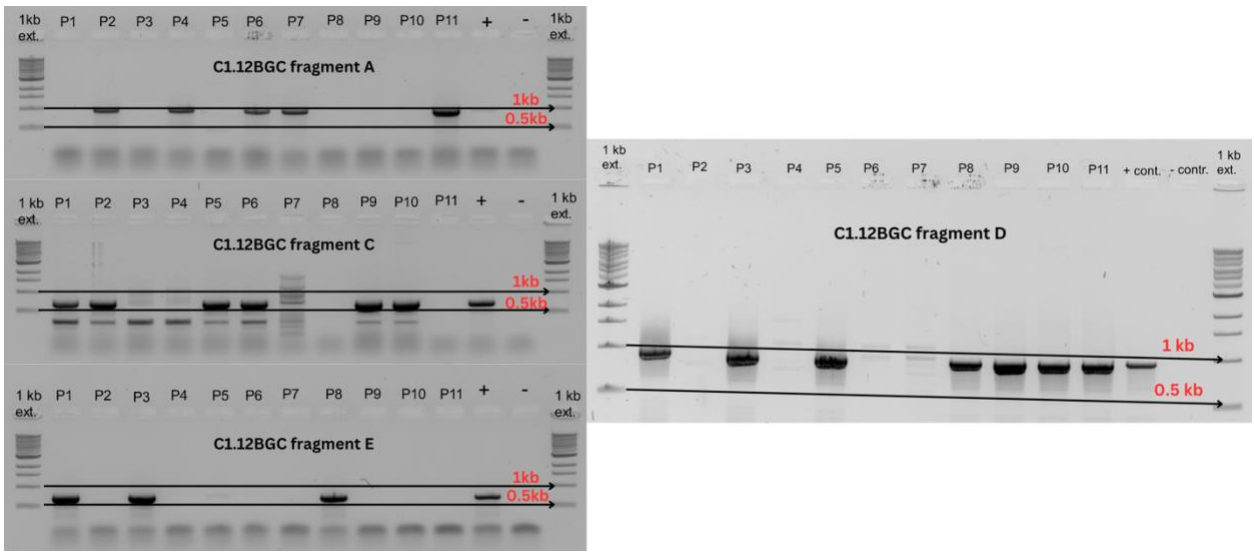

Figure 20. Gel electrophoresis showing pooled PCR screening results of eleven *Micromonospora* sp. RLA083 genomic library plates with primers A, C, D, and E for BGC 1.12.

Plates 2 and 6 showed positive signals for both BGC 1.12 fragments A and C, and the quarters of these plates were tested with primers A, B, and C. All three fragments were amplified from the DNA pool of the 3rd quarter of plate 2 (P2.3). Next, the columns of the 3rd quarter of plate 2 were screened with primers amplifying A, B, and C, identifying ABC fragments of the cluster in the 4th column (P2.3.4). (Figure 21)

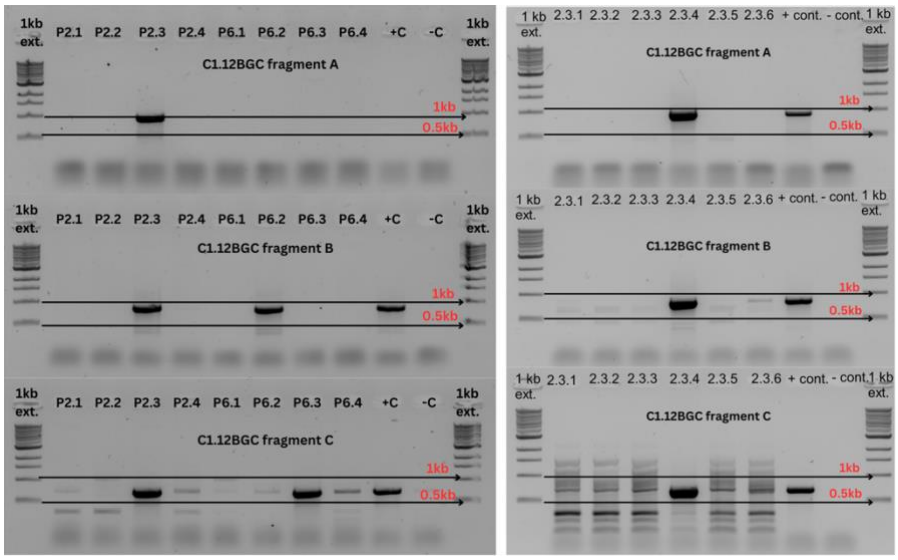

Figure 21. Gel electrophoresis showing pooled PCR screening results of quarters of plates 2 and 6 (*left*) and columns of the 3rd quarter of plate 2 (*right*) with primers A, B, and C for BGC 1.12.

Finally, the single clones from the 4th column of the 3rd quarter of plate 2 were screened for the presence of all, A, B, and C, fragments, which were then confirmed to be contained within the 3rd fosmid clone (P2.3.4.3). (Figure 22)

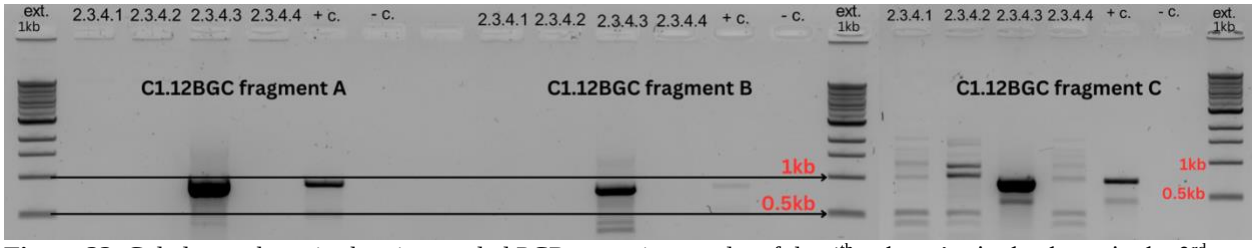

Figure 22. Gel electrophoresis showing pooled PCR screening results of the 4th column's single clones in the 3rd quarter of plate 2 with primers A, B, and C for BGC 1.12.

On the other hand, plates 1, 5, 9, and 10 showed positive signals for both BGC 1.12 fragments C and D, and their quarters were tested for the CD fragment's presence. However, the screening proceeded with plate 10 since it had amplifications for both fragments in the same plate's quarter (P10.1). The 1st quarter of the genomic library's plate 1 was tested with primers C and D, and both fragments' amplifications were identified in the quarter's 5th column (P10.1.5). (Figure 23)

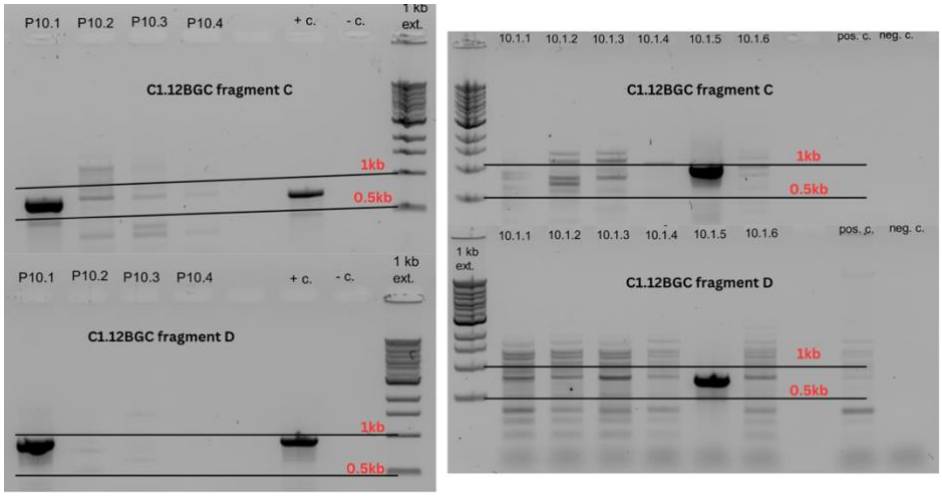

Figure 23. Gel electrophoresis showing pooled PCR screening results of quarters of plate 10 (*left*) and columns of the 1st quarter of plate 10 (*right*) with primers C and D for BGC 1.12.

Next, the 5th column's single clones were tested with primers C and D for the final identification of the clone containing this part of the BGC 1.12. The C and D fragments were detected in the first two fosmid clones of the 5th column. Due to many unspecific signals and no positive control, it was necessary to have the final confirmation by sequencing, which proved the PCR results to be correct. The BGC's C and D fragments were identified in clone 10.1.5.1. (Figure 24)

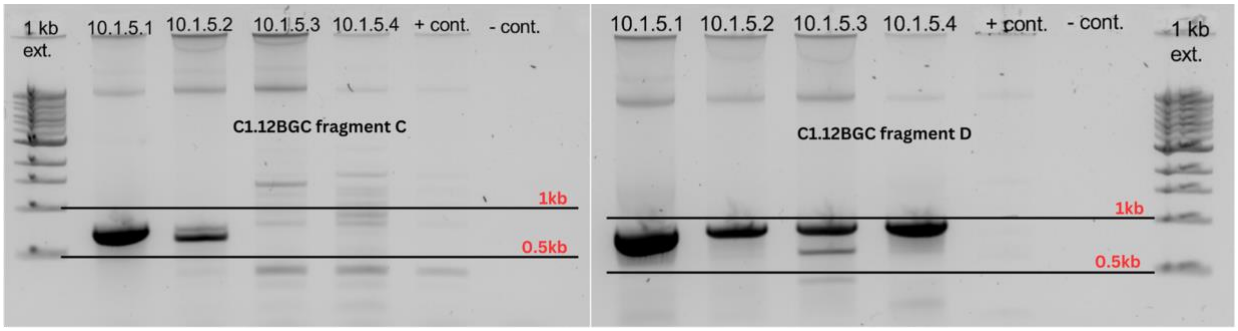

Figure 24. Gel electrophoresis showing pooled PCR screening results of the 5th column's single clones in the 1st quarter of plate 10 with primers C and D for BGC 1.12.

In the first round of screening of all eleven plates, plates 1, 3, and 8 showed positive signals for BGC 1.12 fragments D and E. The quarters of these plates were tested with primers for fragments D and E, and the 3rd quarter of plate 3 containing both fragments was chosen to proceed with (P3.3). Its columns were then screened for the D and E fragments' presence, resulting in three positive columns for both fragments. Due to the strongest amplification, the 3rd column of the 3rd quarter of plate 3 (P3.3.3) was screened for the single fosmid clone with D and E fragments. The final fragment of the BGC 1.12 was detected in the 3rd column's 3rd clone (P3.3.3.3). (Figure 25) The screening process for the D and E fragments was followed by the lack of the positive control's amplification in the gel, which required the results confirmation by sequencing.

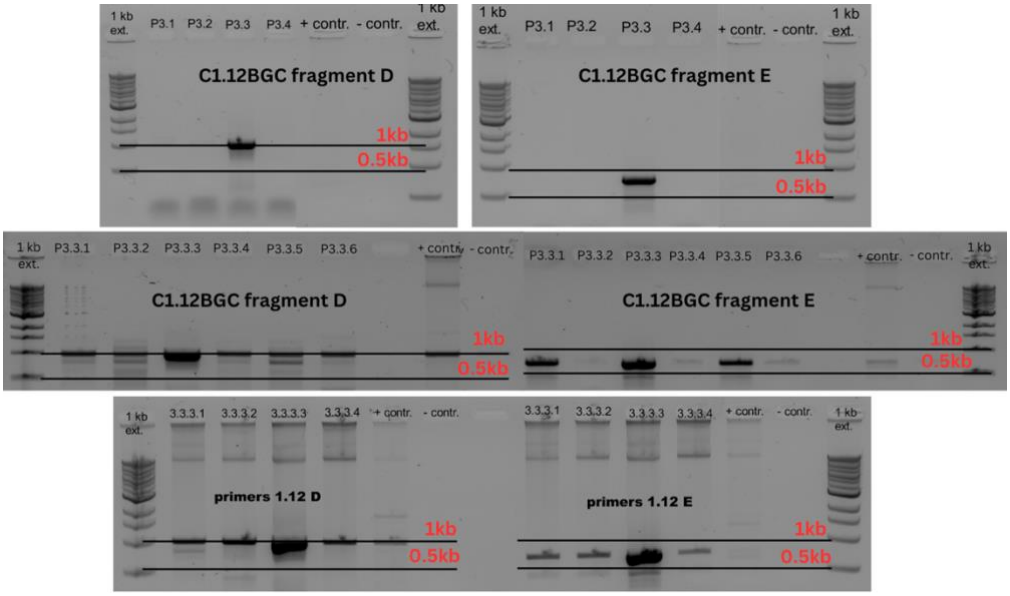

Figure 25. Gel electrophoresis showing pooled PCR screening results of the plate 3 quarters, columns of the 3rd quarter of plate 3, and 3rd column's single clones in the 3rd quarter of plate 3 with primers D and E for BGC 1.12.

The sequencing results confirmed that the BGC 1.12 was contained within three fosmid clones: A, B, and C fragments in 2.3.4.3 clone, CD fragment in 10.1.5.1 clone, and finally D and E fragments in 3.3.3.3 clone. (Figure 26)

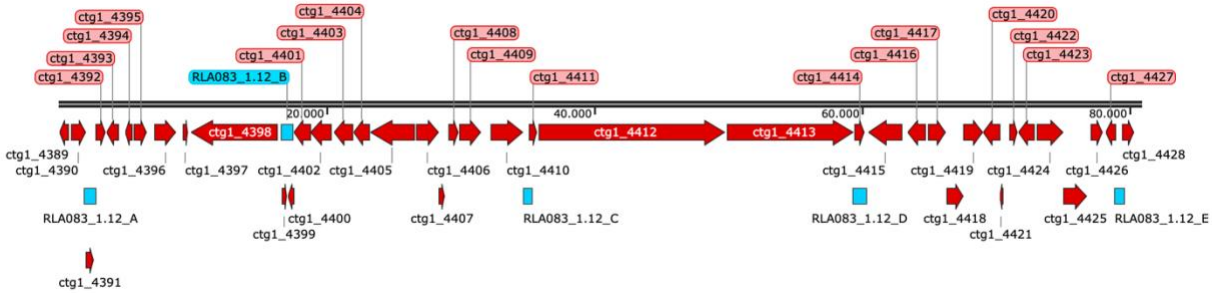

Figure 26. The *Micromonospora* sp. RLA083 BGC 1.12 (80.733 bp). The starting points of primers A, B, C, D, and E for BGC 1.12 are marked in blue.

5.4. Construction of the pYES-based capture vector

5.4.1. Amplification of the capture vector flanks and plasmid isolation

To assemble a capture vector for the BGC 1.8 fragment (further also referred to as 9.2.4.1), first, the DNA fragments homologous to the ends of the cluster were amplified by PCR (Figure 27). The pYES plasmid (Figure 28) employed for capture vector creation was successfully isolated and purified as previously described in the Methods section.

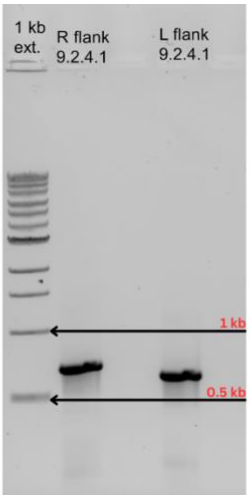

Figure 27. Gel electrophoresis results of the isolated, by PCR amplified, left (606 bp) and right (686 bp) capture fragments.

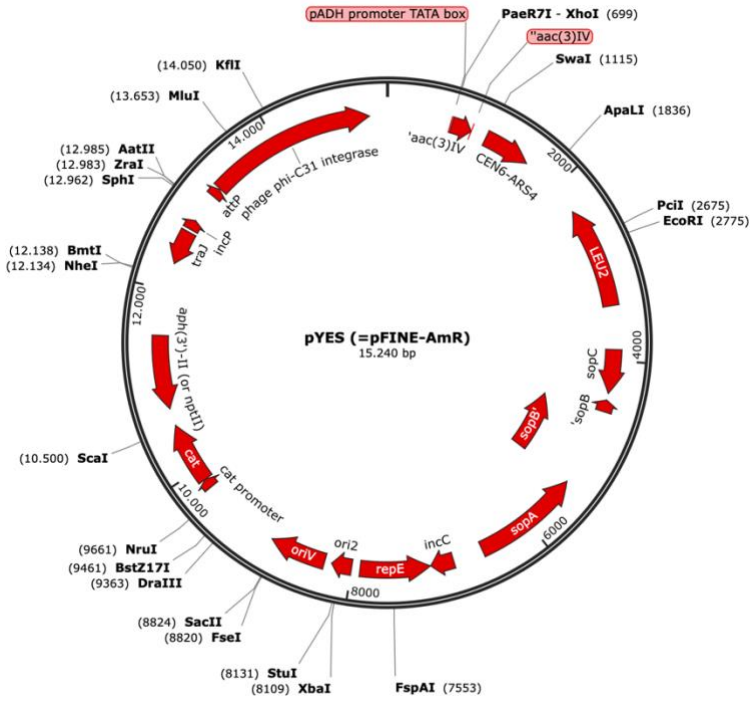

Figure 28. Map of the pYES plasmid used for constructing the BGC 1.8 fragment capture vector. This vector's components are: *LEU2*, CEN6-ARS4, *aac(3)IV*, *sopA*, *sopB*, *sopC*, *incC*, *oriV*, *ori2*, *repE*, cat promoter, *cat*, *aph(3´)-II*, *incP*, *traJ*, attP, phage phi-C31 integrase. Endonuclease restriction sites are also shown.

5.4.2. Construction of capture vector for BGC 1.8

Both capture vector flanks and pYES plasmid were digested with restriction enzymes to attain complimentary ends that will allow successful ligation reaction, creating a capture vector (Table 5). The left flank (1.8_Lflank) was cut with *Sph*I and *Pme*I endonuclease enzymes, resulting in the 589 bp fragment, while the right flank (1.8_Rflank) was digested with *Pme*I and *Aat*II enzymes, yielding a 673 bp fragment. For the digestion of the pYES plasmid, restriction enzymes *Sph*I and *Aat*II were used to attain a 15.217 bp linearized plasmid. The digested fragments were then separated in the agarose gel and successfully isolated and purified for further project steps.

To assemble a capture vector, a ligation reaction was performed, during which the right and left flanks were cloned into the pYES vector (capture vector size 16.479 bp) (Figure 29). The assembled pYES-based capture vector was then successfully transformed into *E. coli* EPI300 and the accuracy of the cloning was tested by performing touchdown PCR and restriction digestion. The PCR results confirmed that the ligation of both vector flanks was successful, whereas the restriction assay with *Hin*dIII and *Bam*HI showed incomplete digestion in the gel electrophoresis (Figure 30 and Table 5). Nevertheless, the successful generation of the capture vector was confirmed by sequencing results.

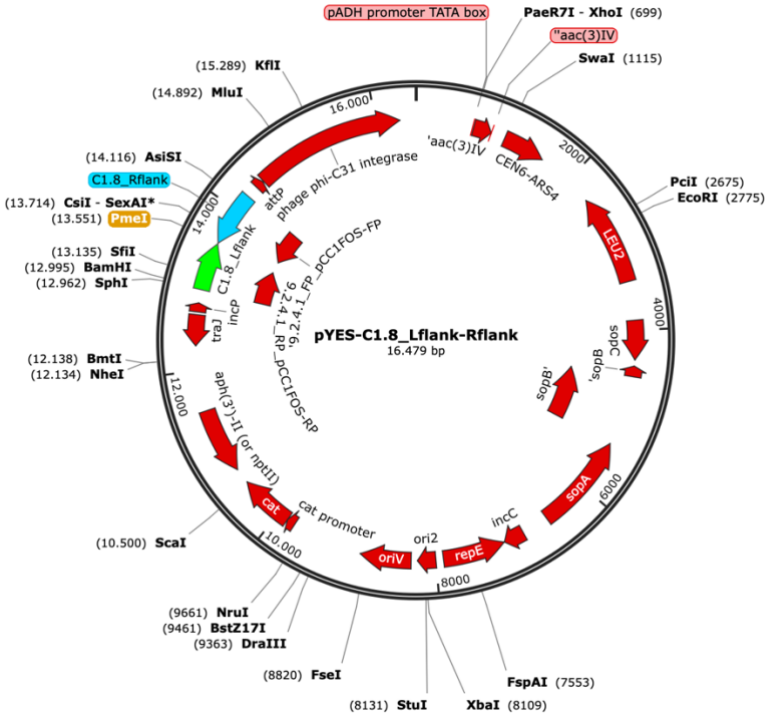

Figure 29. Map of the assembled pYES-based capture vector. The left flank is highlighted in green, while the right flank is marked in blue. *Pme*I, the restriction enzyme highlighted in yellow, linearizes the vector.

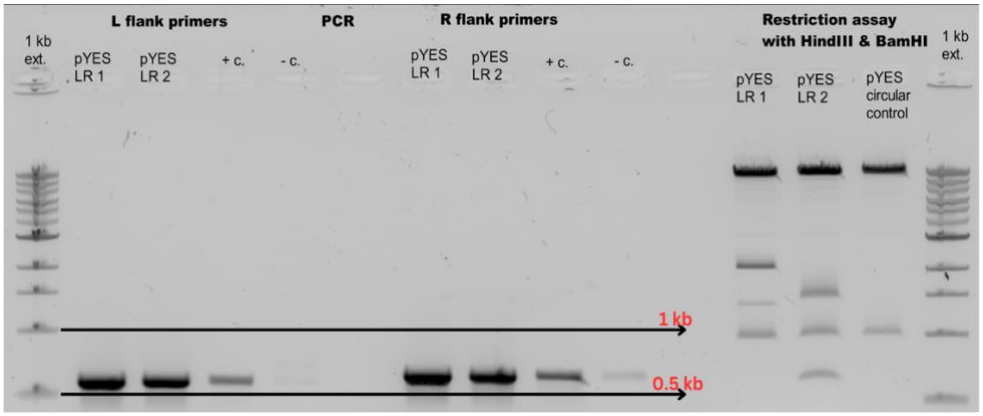

Figure 30. Gel electrophoresis results of the PCR (*left*) and restriction assay (*right*) analysis of the assembled, pYES-based capture vector. PCR results confirmed successful ligation by showing expected bands for the left flank (589 bp) and right flank (673 bp). However, restriction analysis with *Hin*dIII and *Bam*HI was supposed to deliver 13.4 kb, 2.06 kb, and 969 bp fragments of the capture vector, and 14.27 kb and 969 bp fragments of the circular pYES plasmid used as a control, but also showed some unexpected bands.

5.5. Assembly of the BGC 1.8 in yeast

To be able to clone the BGC 1.8 into the pYES-based capture vector, the latter construct was digested with the restriction enzyme *Pme*I, an endonuclease enzyme that has only one restriction site in the pYES-based capture vector (Table 5). This restriction site is between the two capture vector flanks, allowing the linearization of the construct. At the same time, the BGC1.8-containing fragment needed to be isolated from the pCC1FOS fosmid in order to be integrated into the capture vector. For this purpose, the fosmid-BGC construct was digested with two restriction enzymes, *Xba*I and *Spe*I, resulting in the BGC 1.8 fragment (38.820 bp) with ends complementary to the capture vector's flanks. (Figure 31and Table 5)

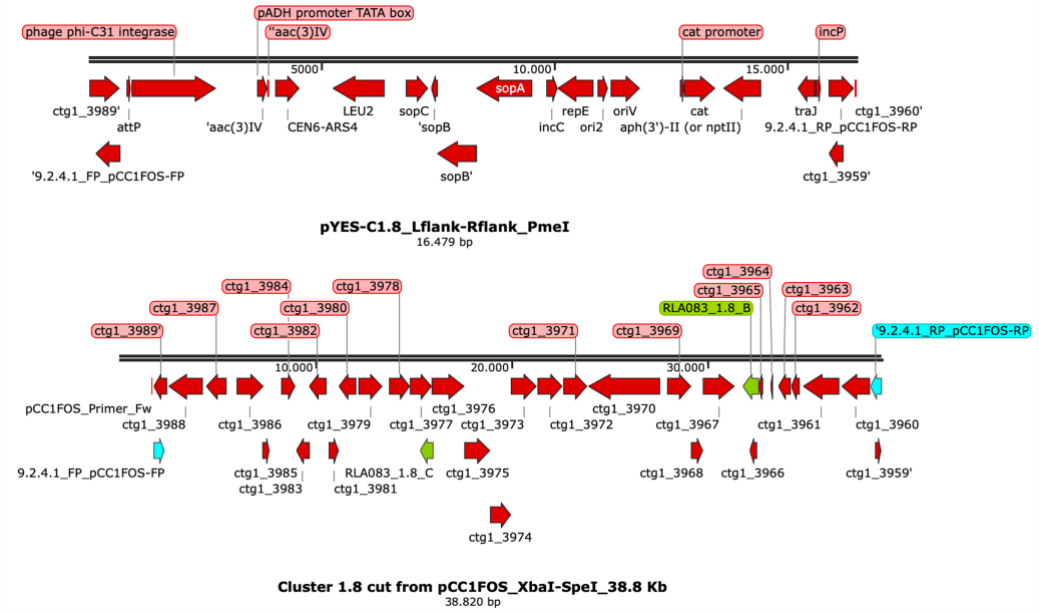

Figure 31. Linearized pYES-based capture vector after digestion reaction with *Pme*I (*above*) and BGC 1.8 fragment cut from the pCC1FOS fosmid with *Xba*I and *Spe*I (*below*).

The digested fragments were run in the agarose gel, and the expected fragment sizes were confirmed (Figure 32). The gel slices containing the digested capture vector and BGC 1.8 fragment were excised, and the fragments successfully isolated and purified from the gel (4.7.1.1.).

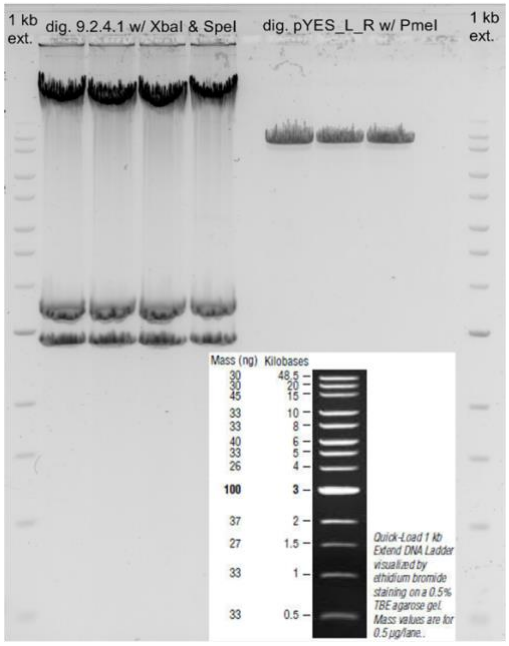

Figure 32. Gel electrophoresis results of the restriction reactions: 1. releasing of the BGC 1.8 fragment (9.2.4.1) from pCC1FOS fosmid with *Xba*I and *Spe*I, delivering three fragments (*left*), 2. linearization of the pYES-based capture vector with *Pme*I, resulting in one fragment of expected size (*right*). Quick-Load 1 kb Extend DNA Ladder was used as a control.

Using *in vivo* TAR (Transformation-Associated Recombination) (51) in *S. cerevisiae* BY4742, the capture vector and BGC 1.8 fragment were assembled into a construct (Figure 33). This yeast strain's leucine auxotrophy was used as a selection marker, so only the cells transformed with pYES-C1.8 construct were able to grow on a leucine-deficient medium.

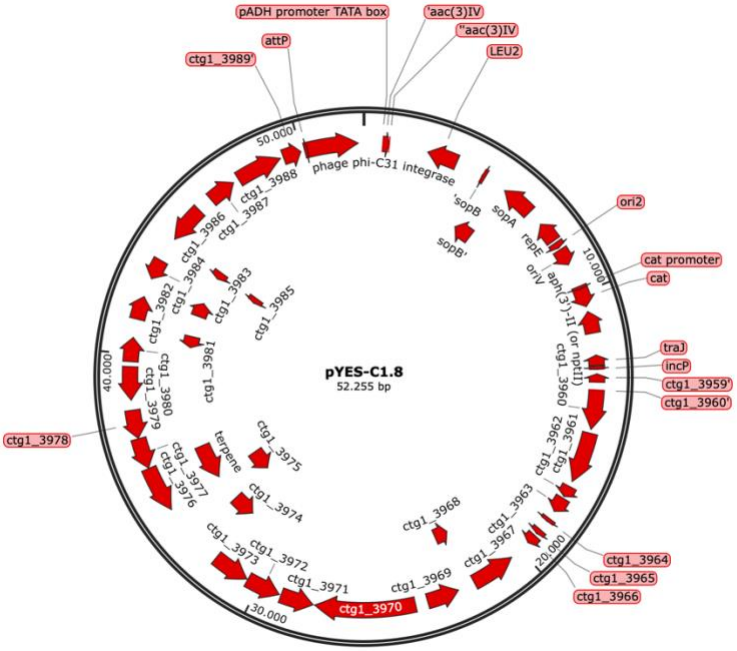

Figure 33. Map of the pYES-based capture vector with incorporated BGC 1.8 fragment (9.2.4.1) via homologous recombination in *S. cerevisiae* BY4742.

Colony PCR of the *S. cerevisiae* BY4742 transformants was performed to detect the clones with correctly assembled pYES-C1.8 construct. Nineteen yeast transformants were tested with primers specific for fragment B for the BGC 1.8 to verify whether the BGC was

incorporated into the capture vector. Only one clone showed amplification for fragment B and was additionally tested with primers for fragment C of the BGC 1.8. After showing an amplification for fragment C as well, the pYES-C1.8 was subsequently isolated and purified from this clone (4.7.4.).

5.6. Transformation of pYES-C1.8 in electrocompetent *E. coli* cells

The pYES-C1.8 construct was first transformed in electrocompetent *E. coli* EPI300 cells to obtain higher concentrations of the plasmid DNA for further project steps (see Methods). The construct was isolated and purified from liquid cultures of ten selected *E. coli* EPI300 clones and tested with primers specific for fragment C to confirm the presence of the BGC 1.8. The isolated plasmid DNA from the first and second clones was tested with 1.8B primers, and the first sample was additionally tested with primers for the left flank, therefore verifying the presence of the entire pYES-C1.8 construct (Figure 34). These gel electrophoresis results were also confirmed by restriction analysis (with *Nde*I, *Eco*RI, and *Nco*I) (Figure 35 and Table 5) and sequencing of the plasmid DNA from sample 1 (1.8pY1).

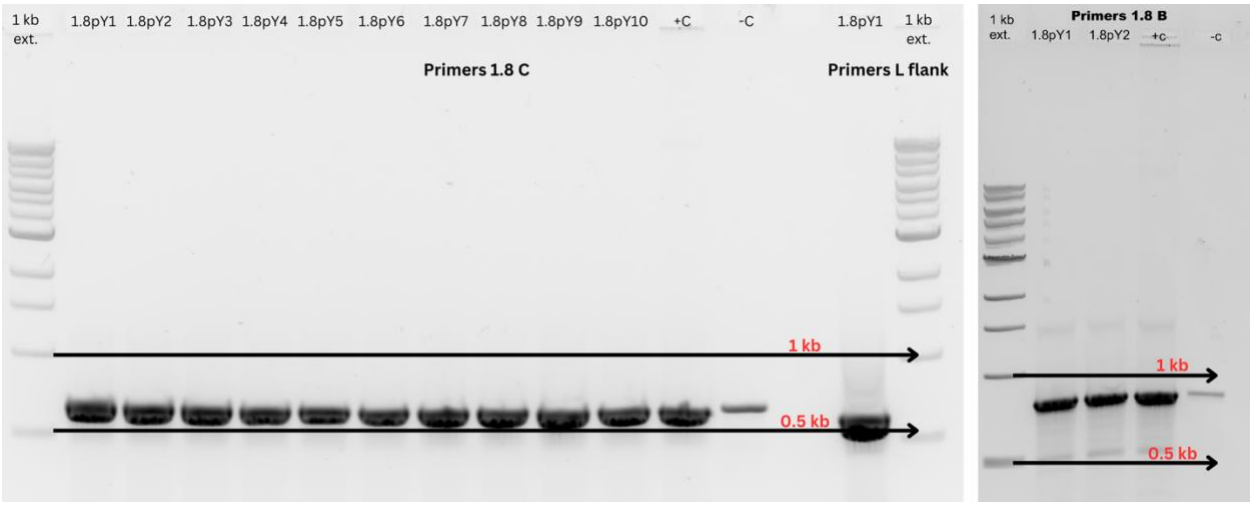

Figure 34. Gel electrophoresis results of ten pYES-C1.8 samples isolated from *E. coli* EPI300 cells tested with primers 1.8C, sample 1 (1.8pY1) tested with primers for the left flank, and sample 1 (1.8pY1) and sample 2 (1.8pY2) tested with primers 1.8B. Bands for fragment C (637 bp), left flank (589 bp), and fragment B (804 bp) confirmed the presence of the entire pYES-C1.8 construct in all tested samples.

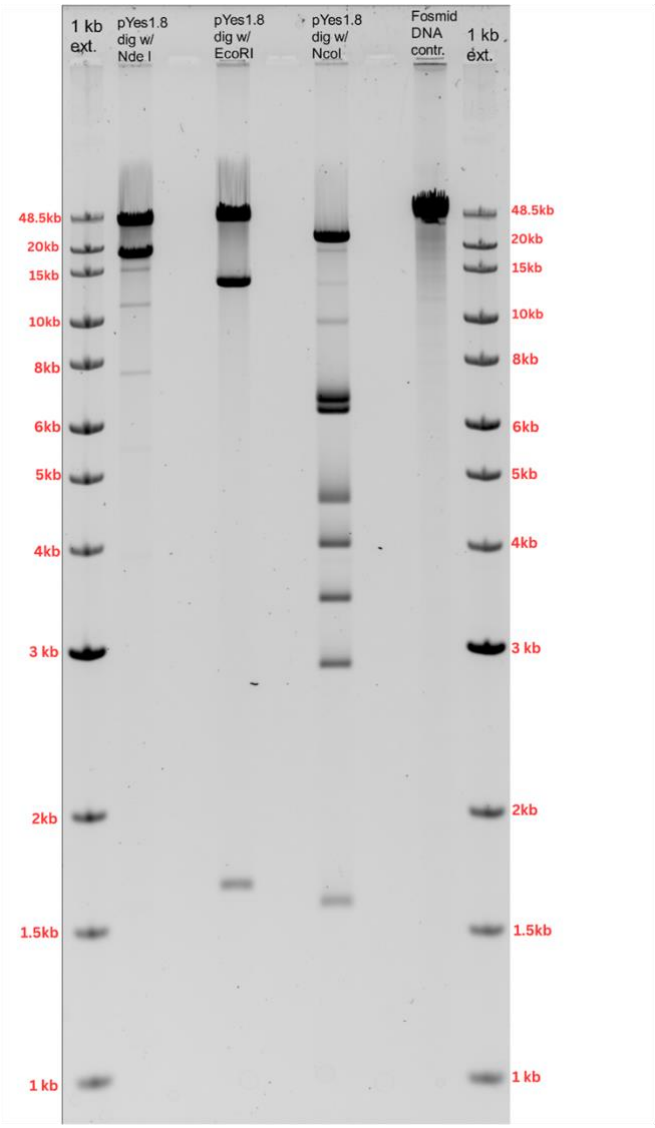

Figure 35. Restriction analysis verification of the pYES-C1.8 construct isolated from *E. coli* EPI300 clone 1. For its accuracy confirmation, fragment sizes of the digestion reaction with *Nde*I: 33.8 kb and 18.3 kb, *Eco*RI: 37.3 kb, 13.1 kb, and 1.6 kb, and *Nco*I: 22.3 kb, 6.5 kb, 6.2 kb, 4.5 kb, 4.0 kb, 3.4 kb, 2.8 kb, 1.6 kb, and 560 bp, were expected. The tested construct showed the appropriate restriction pattern.

Next, the pYES-C1.8BGC construct was transformed into electrocompetent *E. coli* ET12567 in order to conjugate unmethylated plasmid DNA into the actinomycete strains for heterologous expression. The pYES-C1.8BGC construct was successfully conjugated from *E. coli* ET12567 cells into three heterologous hosts: *Streptomyces coelicolor* M1154, *Streptomyces albus* J1074, and *Amycolatopsis japonicum* (4.9.).

5.7. Fermentation of recombinant actinomycete strains and extract analysis

The fermentation step was carried out with the aim of secondary metabolite production. First, the seeding cultures of the recombinant strains, *Streptomyces* and *Amycolatopsis*, were prepared (see Methods). After the seeding cultures had reached sufficient density, two media, NL111V and SM17, were used to start the fermentation cultures for each strain. As controls, actinomycete strains harbouring an empty pYES vector were utilized. For *A. japonicum*, the fermentation was stopped at both 3 and 8 days, and for *S. albus* and *S. coelicolor* it was stopped after 7 days. The cultures were then freeze-dried and extracted with methanol.

The methanolic extracts were analyzed by HPLC, LC-MS, and disc diffusion bioassays. Although the compound production was not detectable using HPLC because of the product's weak chromophore and low produced amounts, it was identified using LC-MS. Disc diffusion assays of methanol extracts' bioactivity with three test organisms (4.12.) showed no growth inhibition of the tested strains. The produced compound exhibited no antibacterial activity against *Staphylococcus carnosus* and *Pseudomonas putida* and no antifungal activity against *Saccharomyces cerevisiae*.

5.7.1. LC-MS analysis

LC-MS analysis, performed by Dr. Martin Zehl at the Mass Spectrometry Center, University of Vienna (4.11.2.), showed the production of several structurally similar compounds in the recombinant *Streptomyces* strains, *S. albus* and *S. coelicolor*, while no such compounds were detected in recombinant *A. japonicum*. The highest compound production was shown for *S. coelicolor* M1154/pYES-C1.8.2_3 grown in NL111V medium in comparison to *S. coelicolor* M1154/pYES-C1.8.2_1 and *S. coelicolor* M1154/pYES-C1.8.2_2 grown in the same medium (Figure 36). In SM17 medium, only traces of the compounds were produced.

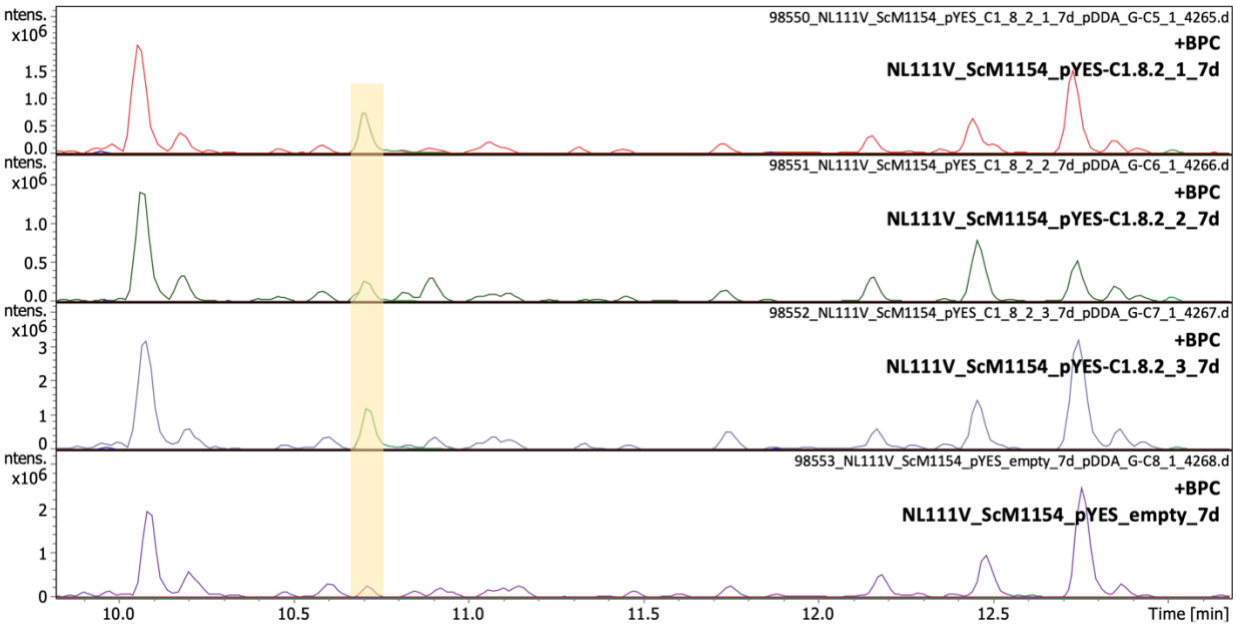

Figure 36. UV-Vis Chromatogram of the extracts from *S. coelicolor* M1154/pYES-C1.8.2_1, *S. coelicolor* M1154/pYES-C1.8.2_2, and *S. coelicolor* M1154/pYES-C1.8.2_3 grown in NL111V medium. *S. coelicolor* M1154 containing the empty pYES vector served as a control. Detection of the absorbance peaks occurred at around 10.7 minutes (marked in yellow).

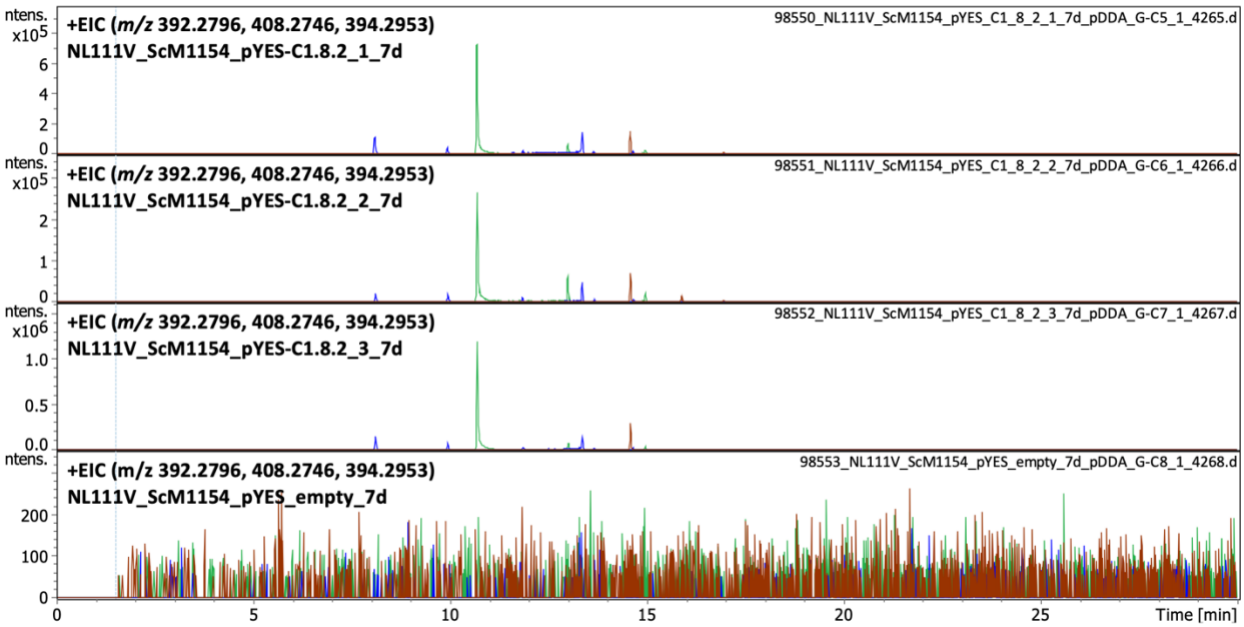

Figure 37. +EIC (m/z 392.2796, 408.2746, 394.2953) of the extracts from *S. coelicolor* M1154/pYES-C1.8.2_1, *S. coelicolor* M1154/pYES-C1.8.2_2, and *S. coelicolor* M1154/pYES-C1.8.2_3 grown in NL111V medium. *S. coelicolor* M1154 containing the empty pYES vector served as a control. The compound eluted at approximately 10.7 minutes.

Heterologous expression of the BGC 1.8 was predicted to result in the production of a glycosylated terpene. However, the final product appears to be the aglycone, with a proposed mass of 410.2901 (m/z). Additionally, the fragmentation results indicate that the main congener with a sum formula $C_{23}H_{39}NO_5$ is a terpene with multiple hydroxyl groups and one amino group. (Figure 38)

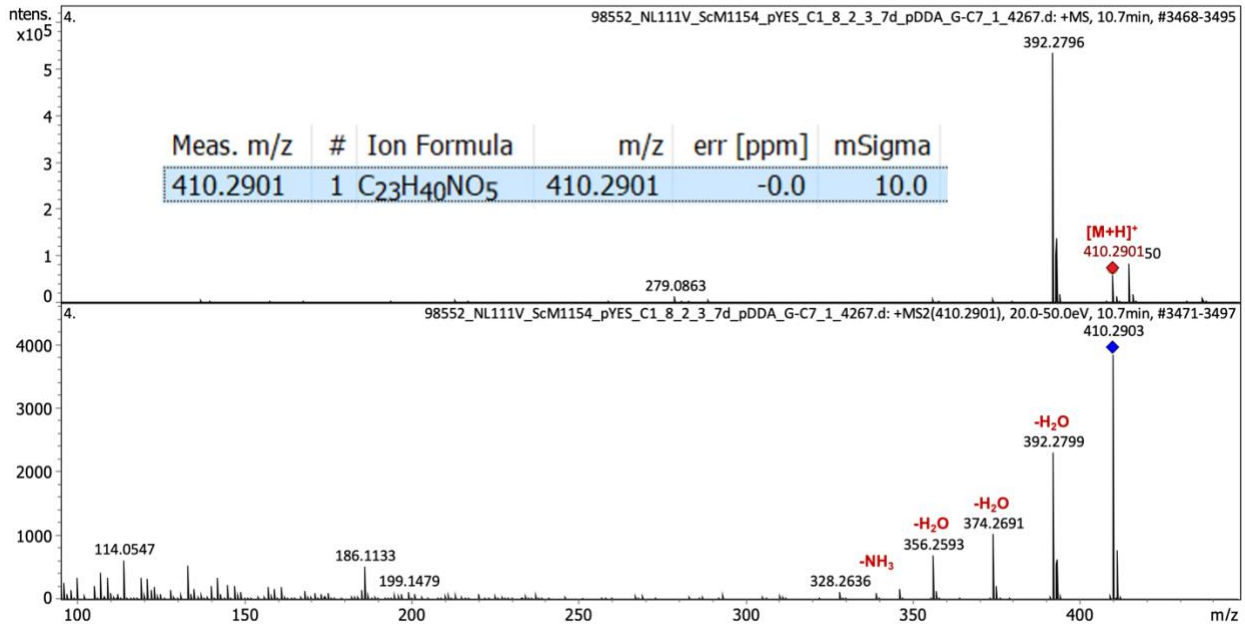

Figure 38. Tandem mass spectrometry (MS/MS) results for the compound produced by *S. coelicolor* M1154/pYES-C1.8.2_3 grown in NL111V medium. $[M+H]^+$ peak corresponds to the compound molecular weight with the mass of the added proton (*above*). Peaks corresponding to the loss of water molecules (H_2O) indicate the presence of three hydroxyl groups, while the detection of ammoniac (NH_3) suggests the presence of an amino group in the molecule (*below*). Based on the fragmentation results, the mass and ion formula of the compound were predicted.

6. Discussion

Microorganisms have long been known as a significant source of novel secondary metabolites with important pharmaceutical and industrial applications. *Micromonospora* spp. have repeatedly shown their potential for production of compounds with interesting bioactivities. Especially microorganisms isolated from underexplored environments may possess unique biosynthetic pathways for novel bioactive compounds. *Micromonospora* sp. RLA083 was isolated from *Leontopodium nivale* subsp. *alpinum* rhizosphere, which is considered to be an appealing environment for bioprospecting of Actinobacteria (14).

The antiSMASH analysis of the *Micromonospora* sp. RLA083 genome suggested a presence of 24 BGCs encoded within it. In this study, BGC 1.8 was chosen to be worked with due to its low similarity to the already described BGCs, making it more likely for its product to be a novel compound. The software also indicated that the BGC 1.8 encodes a glycosylated terpene, and terpenes have previously exhibited a wide range of biological activities, of which some are beneficial for human health. To achieve the production and characterise the secondary metabolite, the BGC 1.8 was first successfully identified within a single clone in the *Micromonospora* sp. RLA083 genomic library, then assembled in *Saccharomyces cerevisiae* BY4742, and finally heterologously expressed in *Streptomyces albus* J1074 and *Streptomyces coelicolor* M1154. The highest yields of production were detected in recombinant *S. coelicolor* M1154 clone harbouring BGC1.8 and grown in NL111V medium, suggesting that this medium is an optimal choice for future compound production.

Even though the compound production was not detectable by HPLC, likely due to the compound's weak chromophore and low amounts produced, LC-MS analysis detected its elution from the chromatographic column at approximately 10.7 minutes. The produced compound, tentatively named brasilicardin H, seems to be a terpene aglycone with multiple hydroxyl groups, one amino group and a sum formula $C_{23}H_{39}NO_5$. These data do not match the data of any already reported microbial metabolite, which supports the hypothesis that BGC 1.8 encodes a previously uncharacterized compound.

Nocardia terpenica IFM 0406 BGC encoding brasilicardin A was initially identified and described by Dairi and co-workers, who reported that the BGC consists of 11 genes (*bra1-bra11*) (61). However, a more recent study redefined the BGC's borders by providing evidence of two additional genes being present in the cluster, *bra0* and *bra12* (31). Comparative analysis of *Micromonospora* sp. RLA083 BGC 1.8 and *Nocardia terpenica* IFM 0406 BGC for brasilicardin A revealed that the biosynthetic pathway of the compound detected in this study appears to be similar to that of brasilicardin A aglycon (Figure 39). This analysis led to an observation that the homologs of *bra*0-6 genes are present in both clusters, possibly leading to the same compound backbone. However, BGC 1.8 seems to be missing the gene coding for O-methyltransferase, causing the absence of the methyl group in brasilicardin H. Additionally, BGC1.8 lacks *bra*7-11 genes, having other product-modifying genes instead. Moreover, BGC 1.8 appears to have three glycosyltransferases, whereas brasilicardin A BGC has only one, which likely adds different sugars to the core structure. (Table 15) However, the glycosyltransferases seem to be non-functional or not expressed in the brasilicardin H-producing recombinant *Streptomyces* strains. (Figure 40)

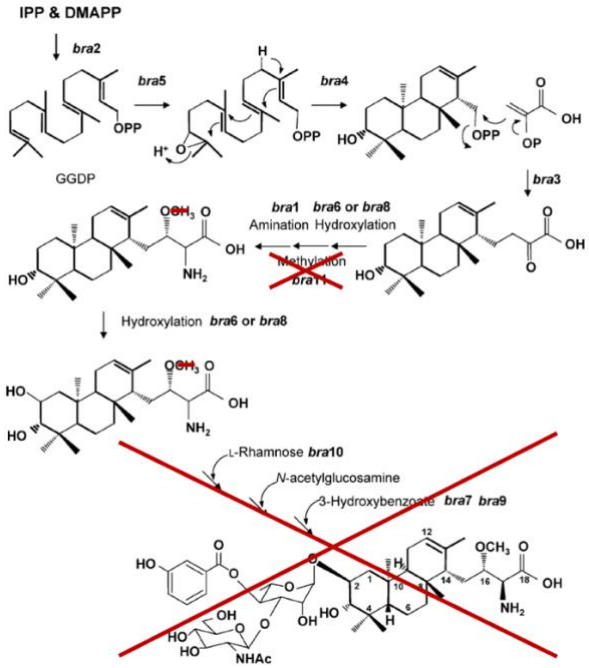

Figure 39. Suggested synthesis pathway of brasilicardin H produced by *S. coelicolor* M1154/pYES-C1.8.2_3 grown in NL111V medium based on the synthesis pathway of brasilicardin A. *modified figure (61)

Table 15. Comparative analysis of gene composition in *Micromonospora* sp. RLA083 BGC 1.8 and *Nocardia terpenica* IFM 0406 BGC brasilicardin A.

Gene category	***Micromonospora* sp. RLA083 BGC 1.8 gene(s)**	**Count**	***Nocardia terpenica* IFM 0406 BGC Brasilicardin A gene(s)**	**Count**
Tetrapyrrole methylase	*ctg1_3967*	1	-	0
Precorrin-8X methylmutase	*ctg1_3968*	1	-	0
Nitrite/Sulfite reductase ferredoxin-like half domain	*ctg1_3969*	1	-	0
Cobaltochelatase subunit CobN	*ctg1_3970*	1	-	0
Cytochrome P450	*ctg1_3971*	1	*bra 6*	1
Glycosyltransferases	*ctg1_3972* (MGT family), *ctg1_3973* (GT family 2), *ctg1_3979* (MGT family)	3	*bra 10* (Rhamnosyltransferase)	1
Aminotransferase	*ctg1_3974*	1	*bra1*	1
Epoxidase	*ctg1_3975* (Monooxygenase FAD-binding)	1	*bra5* (Geranylgeranyl diphosphate epoxidase)	1
Terpene Cyclase	*ctg1_3976*	1	*bra4* (Diterpene cyclase)	1
Prenyltransferase	*ctg1_3977* (UbiA)	1	*bra3*	1
Isoprenoid Synthases	*ctg1_3978* (Polyprenyl synthetase)	1	*bra2* (Geranylgeranyl diphosphate synthase)	1
Dioxygenase	*ctg1_3980* (TauD/TfdA family)	1	*bra0*	1
Unknown/ferredoxin domain-containing protein	*ctg1_3981*	1	-	0
Calcineurin-like phosphoesterase	*ctg1_3982*	1	-	0
4'-phosphopantetheinyl transferase	*ctg1_3983*	1	-	0
HAD-superfamily hydrolase subfamily IA	*ctg1_3984*	1	-	0
Regulators	*ctg1_3985* (Lsr2 family), *ctg1_3987* (LacI family)	2	*bra12* (AfsR)	1

Glycosyl hydrolase	*ctg1_3986*	1	-	0
ABC Transporter	*ctg1_3988*	1	-	0
2,3-dihydroxybenzoate-AMP ligase	-	0	*bra7*	1
FK506 oxidase	-	0	*bra8*	1
O-methyltransferase	-	0	*bra11*	1

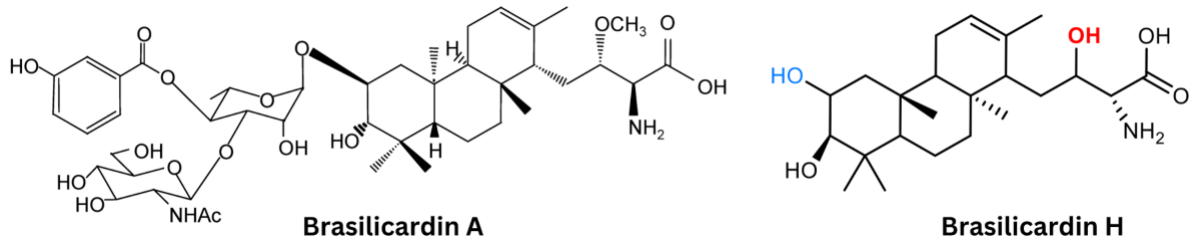

Figure 40. Structural brasilicardin A formula (61) and a proposed structural formula of brasilicardin H, with highlighted differences in their structure.

The lack of glycosylation of the final product could lead to a reduction in its bioactivity potential, solubility, or overall stability. The reason for this occurrence could be the incapability of the recombinant *Streptomyces* host strains to produce sugars, in which case the use of the sugar biosynthesis cassettes could be the solution. In an experiment performed by Schell et al., three biosynthetic gene cassettes for the deoxysugar production, D-mycaminose, D-angolosamine and D-desosamine, were constructed (62). *Saccharopolyspora erythraea*, which was primarily not able to produce the aglycone and sugars, was however able to express the constructed gene cassettes and transfer the synthesized deoxysugars to the respective exogenously added aglycone acceptors (62). Similarly, Fayed et al. also used D-mycaminose and D-angolosamine cassettes to glycosylate 6-deoxyerythronolide B and erythronolide B in *Streptomyces* host, which resulted in the expected compound production (63).
The absence of glycosylation of the produced compound may be a result of the presence of glycosyl hydrolase in the BGC. A possible strategy to prevent the loss of sugars might be the use of the CRISPR-Cas system to knock-out the gene *ctg1_3986* encoding glycosyl hydrolase. CRISPR-Cas, the genome editing method, allows the DNA modifications at specific sites, leading to either the loss of a target gene's activity or acquisition of new functional traits (64). The use of the TAR shuttle vector pYES for BGC 1.8 assembly was particularly important, since it supports the application of the CRISPR-Cas system in yeast.

Brasilicardin A shows immunosuppressive activity with a mode of action that completely differs from any currently marketed immunosuppressive drug (65). The resemblance in biosynthetic pathway between brasilicardin A and H could potentially result in the same or similar activity, which needs to be further characterized.
Performed bioassays with methanol extracts containing brasilicardin H exhibited neither antibacterial activity against *Staphylococcus carnosus* and *Pseudomonas putida*, nor antifungal activity against *Saccharomyces cerevisiae*. Despite the absence of antimicrobial activity against the tested strains, the possibility of antimicrobial activity cannot be ruled out until further tests with other strains are performed.

In conclusion, once optimized, the heterologous expression of BGC 1.8 could lead to production of a compound with potentially beneficial effects on human health. Further structural NMR, bioactivity, and glycosyltransferase characterization will be necessary to further investigate the potential of BGC1.8 for production of health-beneficial compound(s).

7. Conclusion and Outlook

In this Master's thesis project, the BGC 1.8 from the genome of *Micromonospora* sp. RLA083 was analyzed using bioinformatics, genetic, and molecular biology tools to assess its potential for novel bioactive compound production. The cluster was successfully identified in a single clone from the created genomic library, assembled in a vector, and heterologously expressed in three strains with higher production yields observed in *S. coelicolor* M1154.

The produced compound was not detectable by HPLC, possibly due to the compound's weak chromophore or low production yields, but was successfully identified by LC-MS. The analysis indicated that the compound is a terpene aglycon, whose biosynthetic pathway appears to be similar to that of brasilicardin A aglycon. However, the expected glycosylation of the final product was not observed, possibly because of the *Streptomyces* host strain's inability to produce sugars, or the glycosyl hydrolase's gene presence in the BGC. These limitations could potentially be solved by using sugar biosynthesis cassettes, or by knocking out a glycosyl hydrolase gene utilizing CRISPR-Cas system. Given the structural and biosynthetic pathway similarities to brasilicardin A, it is necessary to further examine the compound's potential as an immunosuppressive drug. Despite no detection of the compound's bioactivity against the tested strains in the bioassays, its biological activity potential should not be ruled out.

To fully evaluate the produced compound's potential as a novel drug, it will be necessary to optimize production yields, perform additional bioactivity tests, characterize glycosyltransferases, and further analyze the compound's structure using NMR.

8. References

1. Geller-McGrath D, Mara P, Taylor GT, Suter E, Edgcomb V, Pachiadaki M. Diverse secondary metabolites are expressed in particle-associated and free-living microorganisms of the permanently anoxic Cariaco Basin. Nat Commun. 2023 Feb 6;14(1):656.

2. Sengupta S, Chattopadhyay MK, Grossart HP. The multifaceted roles of antibiotics and antibiotic resistance in nature. Front Microbiol [Internet]. 2013 [cited 2025 Feb 19];4. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2013.00047/abstract

3. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk HP, et al. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiol Mol Biol Rev. 2016 Mar;80(1):1–43.

4. Hamedi J, Poorinmohammad N, Papiran R. Growth and Life Cycle of Actinobacteria. In: Wink J, Mohammadipanah F, Hamedi J, editors. Biology and Biotechnology of Actinobacteria [Internet]. Cham: Springer International Publishing; 2017 [cited 2025 Feb 21]. p. 29–50. Available from: http://link.springer.com/10.1007/978-3-319-60339-1_3

5. Ngamcharungchit C, Chaimusik N, Panbangred W, Euanorasetr J, Intra B. Bioactive Metabolites from Terrestrial and Marine Actinomycetes. Molecules. 2023 Aug 6;28(15):5915.

6. Hutchings MI, Truman AW, Wilkinson B. Antibiotics: past, present and future. Curr Opin Microbiol. 2019 Oct;51:72–80.

7. Abdel-Razek AS, El-Naggar ME, Allam A, Morsy OM, Othman SI. Microbial Natural Products in Drug Discovery. Processes. 2020 Apr 16;8(4):470.

8. MacNair CR, Tsai CN, Rutherford ST, Tan MW. Returning to Nature for the Next Generation of Antimicrobial Therapeutics. Antibiotics. 2023 Aug 1;12(8):1267.

9. Yan S, Zeng M, Wang H, Zhang H. *Micromonospora* : A Prolific Source of Bioactive Secondary Metabolites with Therapeutic Potential. J Med Chem. 2022 Jul 14;65(13):8735–71.

10. Hifnawy MS, Fouda MM, Sayed AM, Mohammed R, Hassan HM, AbouZid SF, et al. The genus *Micromonospora* as a model microorganism for bioactive natural product discovery. RSC Adv. 2020;10(35):20939–59.

11. Wei Z, Shi X, Lian R, Wang W, Hong W, Guo S. Exclusive Production of Gentamicin C1a from Micromonospora purpurea by Metabolic Engineering. Antibiotics. 2019 Dec 14;8(4):267.

12. Zhang F, Zhao M, Braun DR, Ericksen SS, Piotrowski JS, Nelson J, et al. A marine microbiome antifungal targets urgent-threat drug-resistant fungi. Science. 2020 Nov 20;370(6519):974–8.

13. Back CR, Stennett HL, Williams SE, Wang L, Ojeda Gomez J, Abdulle OM, et al. A New Micromonospora Strain with Antibiotic Activity Isolated from the Microbiome of a Mid-Atlantic Deep-Sea Sponge. Mar Drugs. 2021 Feb 11;19(2):105.

14. Oberhofer M, Hess J, Leutgeb M, Gössnitzer F, Rattei T, Wawrosch C, et al. Exploring Actinobacteria Associated With Rhizosphere and Endosphere of the Native Alpine Medicinal Plant Leontopodium nivale Subspecies alpinum. Front Microbiol. 2019 Nov 8;10:2531.

15. Crits-Christoph A, Bhattacharya N, Olm MR, Song YS, Banfield JF. Transporter genes in biosynthetic gene clusters predict metabolite characteristics and siderophore activity. Genome Res. 2021 Feb;31(2):239–50.

16. Kwon MJ, Steiniger C, Cairns TC, Wisecaver JH, Lind AL, Pohl C, et al. Beyond the Biosynthetic Gene Cluster Paradigm: Genome-Wide Coexpression Networks Connect Clustered and Unclustered Transcription Factors to Secondary Metabolic Pathways.

17. Zotchev SB. Antibiotics: Biosynthesis. In: Wiley Encyclopedia of Chemical Biology [Internet]. 1st ed. Wiley; 2008 [cited 2024 Oct 30]. p. 1–9. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9780470048672.wecb014

18. Akhter N, Liu Y, Auckloo BN, Shi Y, Wang K, Chen J, et al. Stress-Driven Discovery of New Angucycline-Type Antibiotics from a Marine Streptomyces pratensis NA-ZhouS1. Mar Drugs. 2018 Sep 12;16(9):331.

19. Jones JA, Wang X. Use of bacterial co-cultures for the efficient production of chemicals. Curr Opin Biotechnol. 2018 Oct;53:33–8.

20. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, et al. SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol. 2010 Sep;47(9):736–41.

21. Tarasova EV, Luchnikova NA, Grishko VV, Ivshina IB. Actinomycetes as Producers of Biologically Active Terpenoids: Current Trends and Patents. Pharmaceuticals. 2023 Jun 12;16(6):872.

22. Rinaldi MA, Ferraz CA, Scrutton NS. Alternative metabolic pathways and strategies to high-titre terpenoid production in *Escherichia coli*. Nat Prod Rep. 2022;39(1):90–118.

23. Andreu A, Ćorović M, Garcia-Sanz C, Santos AS, Milivojević A, Ortega-Nieto C, et al. Enzymatic Glycosylation Strategies in the Production of Bioactive Compounds. Catalysts. 2023 Oct 11;13(10):1359.

24. Singh S, Phillips Jr. GN, Thorson JS. The structural biology of enzymes involved in natural product glycosylation. Nat Prod Rep. 2012;29(10):1201.

25. Schmid J, Heider D, Wendel NJ, Sperl N, Sieber V. Bacterial Glycosyltransferases: Challenges and Opportunities of a Highly Diverse Enzyme Class Toward Tailoring Natural Products. Front Microbiol [Internet]. 2016 Feb 18 [cited 2025 Feb 28];7. Available from: http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00182/abstract

26. Elshahawi SI, Shaaban KA, Kharel MK, Thorson JS. A comprehensive review of glycosylated bacterial natural products. Chem Soc Rev. 2015;44(21):7591–697.

27. Richter TKS, Hughes CC, Moore BS. Sioxanthin, a novel glycosylated carotenoid, reveals an unusual subclustered biosynthetic pathway. Environ Microbiol. 2015 Jun;17(6):2158–71.

28. Dürr C, Schnell HJ, Luzhetskyy A, Murillo R, Weber M, Welzel K, et al. Biosynthesis of the Terpene Phenalinolactone in Streptomyces sp. Tü6071: Analysis of the Gene Cluster and Generation of Derivatives. Chem Biol. 2006 Apr;13(4):365–77.

29. Gebhardt K, Meyer SW, Schinko J, Bringmann G, Zeeck A, Fiedler HP. Phenalinolactones A–D, terpenoglycoside antibiotics from Streptomyces sp. Tü 6071. J Antibiot (Tokyo). 2011 Mar;64(3):229–32.

30. Usui T, Nagumo Y, Watanabe A, Kubota T, Komatsu K, Kobayashi J, et al. Brasilicardin A, a Natural Immunosuppressant, Targets Amino Acid Transport System L. Chem Biol. 2006 Nov;13(11):1153–60.

31. Botas A, Eitel M, Schwarz PN, Buchmann A, Costales P, Núñez LE, et al. Genetic Engineering in Combination with Semi-Synthesis Leads to a New Route for Gram-Scale Production of the Immunosuppressive Natural Product Brasilicardin A. Angew Chem Int Ed. 2021 Jun 7;60(24):13536–41.

32. Beattie AJ, Hay M, Magnusson B, De Nys R, Smeathers J, Vincent JFV. Ecology and bioprospecting. Austral Ecol. 2011 May;36(3):341–56.

33. Schneider YK. Bacterial Natural Product Drug Discovery for New Antibiotics: Strategies for Tackling the Problem of Antibiotic Resistance by Efficient Bioprospecting.

34. Sekurova ON, Schneider O, Zotchev SB. Novel bioactive natural products from bacteria via bioprospecting, genome mining and metabolic engineering. Microb Biotechnol. 2019 Sep;12(5):828–44.

35. Browne PD, Nielsen TK, Kot W, Aggerholm A, Gilbert MTP, Puetz L, et al. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. GigaScience. 2020 Feb 1;9(2):giaa008.

36. Xu M, Wright GD. Heterologous expression-facilitated natural products' discovery in actinomycetes. J Ind Microbiol Biotechnol. 2019 Mar 1;46(3–4):415–31.

37. Atanasov AG, Zotchev SB, Dirsch VM, the International Natural Product Sciences Taskforce, Orhan IE, Banach M, et al. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 2021 Mar;20(3):200–16.

38. Medema MH, Blin K, Cimermancic P, De Jager V, Zakrzewski P, Fischbach MA, et al. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 2011 Jul 1;39(suppl_2):W339–46.

39. Epstein SC, Charkoudian LK, Medema MH. A standardized workflow for submitting data to the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository: prospects for research-based educational experiences. Stand Genomic Sci. 2018 Dec;13(1):16.

40. Ongley SE, Bian X, Neilan BA, Müller R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. Nat Prod Rep. 2013;30(8):1121.

41. Singh TA, Passari AK, Jajoo A, Bhasin S, Gupta VK, Hashem A, et al. Tapping Into Actinobacterial Genomes for Natural Product Discovery. Front Microbiol. 2021 Jun 22;12:655620.

42. Siegl T, Luzhetskyy A. Actinomycetes genome engineering approaches. Antonie Van Leeuwenhoek. 2012 Oct;102(3):503–16.

43. Huo L, Hug JJ, Fu C, Bian X, Zhang Y, Müller R. Heterologous expression of bacterial natural product biosynthetic pathways. Nat Prod Rep. 2019;36(10):1412–36.

44. Florez-Cardona V, Khani J, McNutt E, Manta B, Berkmen M. Plasmid Library Construction From Genomic DNA. Curr Protoc. 2025 Jan;5(1):e70088.

45. Owen JG, Reddy BVB, Ternei MA, Charlop-Powers Z, Calle PY, Kim JH, et al. Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. Proc Natl Acad Sci. 2013 Jul 16;110(29):11797–802.

46. Zhang H, Boghigian BA, Armando J, Pfeifer BA. Methods and options for the heterologous production of complex natural products. Nat Prod Rep. 2011;28(1):125–51.

47. Vieira Gomes AM, Souza Carmo T, Silva Carvalho L, Mendonça Bahia F, Parachin NS. Comparison of Yeasts as Hosts for Recombinant Protein Production. Microorganisms. 2018 Apr 29;6(2):38.

48. Tippelt A, Nett M. Saccharomyces cerevisiae as host for the recombinant production of polyketides and nonribosomal peptides. Microb Cell Factories. 2021 Dec;20(1):161.

49. Gomez-Escribano JP, Bibb MJ. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. Microb Biotechnol. 2011 Mar;4(2):207–15.

50. Nah HJ, Pyeon HR, Kang SH, Choi SS, Kim ES. Cloning and Heterologous Expression of a Large-sized Natural Product Biosynthetic Gene Cluster in Streptomyces Species. Front Microbiol [Internet]. 2017 Mar 15 [cited 2025 Mar 3];8. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2017.00394/full

51. Kouprina N, Larionov V. Transformation-associated recombination (TAR) cloning and its applications for gene function; genome architecture and evolution; biotechnology and biomedicine. Oncotarget. 2023 Dec 22;14(1):1009–33.

52. Bilyk O, Sekurova ON, Zotchev SB, Luzhetskyy A. Cloning and Heterologous Expression of the Grecocycline Biosynthetic Gene Cluster. Virolle MJ, editor. PLOS ONE. 2016 Jul 13;11(7):e0158682.

53. Xu M, Wang W, Waglechner N, Culp EJ, Guitor AK, Wright GD. GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. Nat Commun. 2020 Oct 16;11(1):5232.

54. Xu M, Wang W, Waglechner N, Culp EJ, Guitor AK, Wright GD. GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. Nat Commun. 2020 Oct 16;11(1):5232.

55. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res. 2023 Jul 5;51(W1):W46–50.

56. BLAST: Basic Local Alignment Search Tool [Internet]. [cited 2025 Mar 8]. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi

57. Manual CopyControl-Fosmid-library-production-kits [Internet]. [cited 2025 Mar 9]. Available from: https://biosearchtech.a.bigcontent.io/v1/static/manual_CCIS125_CopyControl-Fosmid-library-production-kits

58. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 1979 Nov 24;7(6):1513–23.

59. Vaidya SM. Discovery of natural products from Actinobacteria via genome mining [Internet]. [Wien]; 2021. Available from: https://doi.org/10.25365/thesis.70454

60. Guerrero Garzón JF. Discovery of novel biologically active molecules via conventional and genome-guided bioprospecting of actinomycete bacteria [Internet]. [Wien]; 2019. Available from: https://doi.org/10.25365/thesis.59046

61. Hayashi Y, Matsuura N, Toshima H, Itoh N, Ishikawa J, Mikami Y, et al. Cloning of the Gene Cluster Responsible for the Biosynthesis of Brasilicardin A, a Unique Diterpenoid. J Antibiot (Tokyo). 2008 Mar;61(3):164–74.

62. Schell U, Haydock SF, Kaja AL, Carletti I, Lill RE, Read E, et al. Engineered biosynthesis of hybrid macrolide polyketides containing d-angolosamine and d-mycaminose moieties. Org Biomol Chem. 2008;6(18):3315.

63. Fayed B, Ashford DA, Hashem AM, Amin MA, El Gazayerly ON, Gregory MA, et al. Multiplexed Integrating Plasmids for Engineering of the Erythromycin Gene Cluster for Expression in Streptomyces spp. and Combinatorial Biosynthesis. Elliot MA, editor. Appl Environ Microbiol. 2015 Dec 15;81(24):8402–13.

64. Xu Y, Li Z. CRISPR-Cas systems: Overview, innovations and applications in human disease research and gene therapy. Comput Struct Biotechnol J. 2020;18:2401–15.

65. Botas A, Eitel M, Schwarz PN, Buchmann A, Costales P, Núñez LE, et al. Genetic Engineering in Combination with Semi-Synthesis Leads to a New Route for Gram-Scale Production of the Immunosuppressive Natural Product Brasilicardin A. Angew Chem Int Ed. 2021 Jun 7;60(24):13536–41.