universität wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Cloning and heterologous expression of cryptic biosynthetic gene clusters from actinomycete bacterium *Actinoalloteichus fjordicus*“

verfasst von / submitted by

Stefanie Schalko

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

Magistra der Pharmazie (Mag.pharm.)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 449
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Diplomstudium Pharmazie
Betreut von / Supervisor:	Univ.-Prof. Dr. Sergey B. Zotchev
Mitbetreut von / Co-Supervisor:	Dr. Olha Schneider, MSc

Inhaltsverzeichnis

1. List of abbreviations

ACP	Acyl carrier protein
AT	Acyl transferase
antiSMASH	Antibiotics and Secondary Metabolite Analysis Shell
BAC	Bacterial artificial chromosome
BGC	Biosynthetic Gene Cluster
BLAST	Basic Local Alignment Search Tool
bp	Base pair
CATCH	Cas9-Assisted Targeting of Chromosome
DEBS	Deoxyerythronolide B Synthases
DH	Dehydratase
eDNA	Environmental DNA
fDNA	Fosmid DNA
gDNA	Genomic DNA
HTH	Helix-turn-helix
IR	Integrase-mediated recombination
KS	Ketosynthase
LAL	Large ATP-binding regulators of the LuxR family
NCBI	National Center for Biotechnology Information
NP	Natural Product
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OSMAC	One Strain Many Compounds
pDNA	Plasmid DNA
PK	Polyketide
PKS	Polyketide synthase
PPT	4-phosphopantetheine transferase
SARS	*Streptomyces* Antibiotic Regulatory Protein
SM	Secondary Metabolite
SMURF	Secondary Metabolite Unknown Regions Finder
ssDNA	Single stranded DNA
TAR	Transformation-associated recombination
TE	Thioesterase
TetR	Tetracycline Repressor protein
TF	Transcription Factor

2. Zusammenfassung

Seit tausenden Jahren werden Sekundärmetabolite, welche von Pflanzen, Mikroorganismen oder Tieren produziert werden, als Heilmittel verwendet. Bis heute sind sie die wichtigste Quelle moderner Arzneimittelzubereitungen. Die Mehrheit an Bakterien- produzierten Medikamenten, insbesondere mit antibiotischen Eigenschaften, wird von der Familie der *Actinomyceten* produziert. Aufgrund der steigenden Häufigkeit an Infektionen mit multi-resistenten Mikroorganismen, steigt die Notwendigkeit der Entdeckung neuer Arzneistoffe umso mehr.

Aus diesem Grund, wurde in diesem Projekt das sogenannte „Genome mining“ mit der DNA des Organismus *Actinoalloteichus fjordicus* DSM 46856 angewendet, mit dem Ziel, neue bioaktive Substanzen mit antibiotischer Aktivität zu finden.

„Genome mining“ beruht auf der Analyse von bakteriallen Genomen, in Kombination mit dem Onlinetool antiSMASH, welches biosynthetische Gencluster indentifizieren kann, die für die Synthese neuer Arzneimittel verantworlich sind.

Ziel dieser Arbeit war es eine genomische Library von dem Bakterium *Actinoalloteichus fjordicus* zu konstruieren, wessen Genom unter anderem zwei vielversprechende biosynthetische Gencluster besitzt. Anschließend wurde mit Hilfe von pooled-PCR die fertige genomische Library gescreent, wodurch *E.coli* Klone mit den gewünschten biosynthetischen Genclustern gefunden wurden. Durch homologe Rekombination in Hefe wurden diese Fragmente mit dem „Shuttlevector“ pCLY10 eingebaut, wodurch die beiden Gencluster produziert wurden. Als der Zusammenbau erfolgreich war, wurde eine heterologe Expression in drei verschiedenen Wirtsorganismen durchgeführt. Dabei wurden *Streptomyces albus* B4*, Streptomyces albus* J1074 und *Streptomyces coelicolor* M1154 verwendet.

Die daraus gewonnenen Extrakte wurden mittels HPLC-Analyse und Bioaktivitätstest analysiert.

Resultate dieser Analysen haben im Vergleich zu den Kontrollextrakten keine Ergebnisse gezeigt.

Diese Ergebnisse deuten darauf hin, dass eine weitere Manipulation der klonierten BGCs erforderlich ist, um ihre funktionelle Expression zu erreichen, was in zukünftigen Arbeiten realisiert wird.

3. Abstract

Since many years natural products (NPs) of plant, microbial or animal origin have been used as therapeutic agents to treat numerous human diseases. Even until today they remain the most important source of modern drugs. The majority of bacteria-derived NPs, especially with antibiotic properties, were isolated from actinomycete bacteria. Increasing frequency of infection caused by multi-drug resistant microorganisms highlights the importance of discovery of novel NPs with antibiotic activity. For this reason, the focus of this project was "genome minig" of the actinobacterial strain *Actinoalloteichus fjordicus* DSM 46856 with the purpose of discovering new NPs potentially having antibiotic activity.

"Genome mining" relies on the analysis of bacterial genomes with the bioinformatic online tool, antiSMASH, that can identify biosynthetic gene clusters (BGCs) that may govern biosynthesis of novel NPs.

The aim of this work was the construction of a genome library of marine sponge-derived actinomycete bacterium *Actinoalloteichus fjordicus,* which genome harboured two BGCs potentially specifying novel bioactive NPs. Subsequently, the completed library was screened via pooled-PCR to detect library clones with the parts of the BGCs of interest. These parts were assembled into complete BGCs on the modified shuttle vector pCLY10 using transformation-associated recombination in yeast. The assembled BGCs were introduced into three different host organisms: *Streptomyces albus* B4*, Streptomyces albus* J1074 and *Streptomyces coelicolor* M1154 for heterologous expression. The extracts produced from the recombinant strains carrying assembled BGCs were analyzed with HPLC and disc diffusion assays. No differential HPLC peaks nor bioactivity against test organisms used in this work could be observed. These results suggest that further manipulation of the cloned BGCs is required to achieve their functional expression, which will be pursued in a follow-up work.

4. Introduction

4.1. History of new Natural products (NPs)

Since thousands of years, new natural products (NPs) were used as curative agents in traditional medicine. Until today they are the most important origin of modern drugs. A very prominent example of a natural product-derived drug is acetylsalicylic acid, known as aspirin. It has an anti-inflammatory effect and was first isolated from the bark of the plant *Salix alba.*
More than 2500 years ago, Chinese discovered therapeutic effects of moldy soybeans. It was known as the first use of antibiotics, even though the antibiotic compounds were not identified. In 1928, “the milestone in the development of antimicrobial drugs” was reached, as the Scottish microbiologist and pharmacologist Alexander Fleming detected by coincidence the antimicrobial effect of a compound produced by a fungus *Penicillium crysogenum*, named penicillin. The production of penicillin and its purification was substantially improved via intensive research and it had a very important role during the Second World War, saving lives of soldiers suffering from the bacteria-infected wounds.
A few years later the antibiotic chemotherapy drugs, streptomycin and tetracycline were discovered which are effective against many pathogenic bacteria. [1][2]

4.2. Natural products as drugs

Natural products (NPs) encompass a large diversity of chemical structures with a wide range of biological activities, and they are represented by the so-called secondary metabolites (SMs). Metabolism can basically be separated into two different branches, which are intimately linked: primary and secondary metabolism. Primary metabolites are important for the normal functioning of the host organism, such as growth and survival, while secondary metabolites are not essential for the producer, although they appear to give some advantages in its natural environment. [3]

Secondary metabolites (SM) with antibacterial, antiviral, anticancer, antidiabetic, antifungal and other activities are used in human and animal medicine and in agriculture. They derive from different organisms, such as plants, animals, fungi or bacteria. Fig. 1 shows an overview of groups of NPs deriving from different organisms. NPs with anticancer effect are important in the discovery of new active drugs against tumors. For the most part, anticancer drugs originate from plants compared to antibacterial agents which were mostly isolated from microorganisms. [3][2][4]

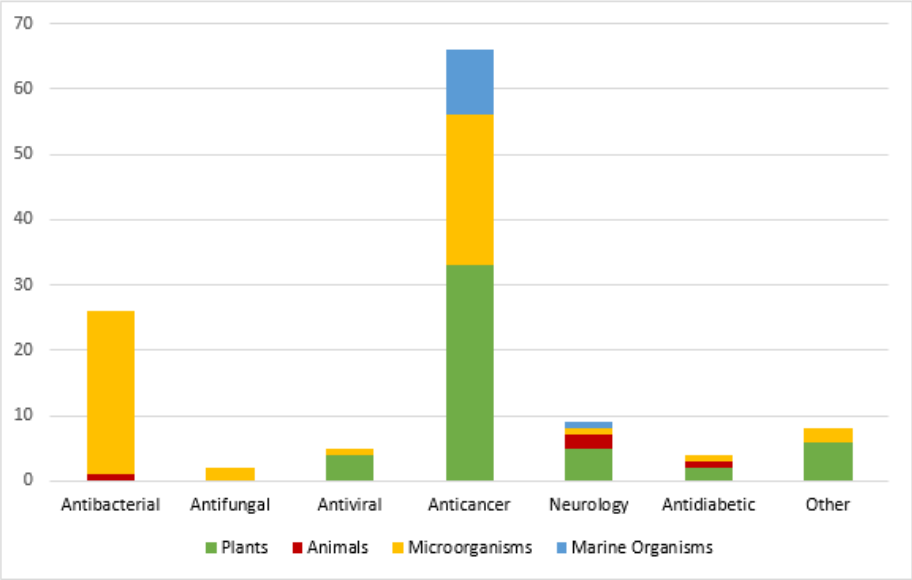


Figure 1: Medicinal groups of NPs, which are in clinical trials ans their origin. [3]

New NPs from plants have been used as medicines since thousands of years, and these days plants are the source of around 25 % of all drugs. Morphine, which was isolated from *Papaver somniferum* was the first natural plant-derived drug for the treating Parkinson´s disease. Another plant-derived SM is artemisinin, which was isolated from the plant *Artemisia annua* to treat malaria, and is now successfully produced by genetically engineered *Saccharomyces cerevisiae.* Anti-tumor agents, like vincristine, vinblastine and paclitaxel, which block mitosis by binding microtubules, are also derived from plants. Furthermore, the topoisomerase II inhibitor etoposide, the topoisomerase I inhibitor camptothecin and its semi-synthetic derivate topotecan, and the translation inhibitor omacetaxine mepesuccinate have plant origins. Additionally, canagliflozin is a type 2 antidiabetic agent and a semi-synthetic derivative of phlorizin, which can be found in many plants. But not only drugs for severe diseases can be obtained from plants, but also active ingredients for topically used creams, like capsaicin from chili peppers or food flavorings, like coumarin. [3][2]

Microorganisms-produced SMs can originate either from bacteria or fungi. The latter has first been reported in 1928 with the discovery of penicillin. More than 20 years later, the important antibacterial agent vancomycin was discovered, isolated from the actinomycete bacterium *Amycolatopsis orientalis.* Lovastatin is another prominent example of currently used drug, which was isolated from the fungus *Aspergillus terreus.* It belongs to the class of statins and became very important in the treatment of hypercholesteremia. [3][2]

4.3. Bacterial natural products

An important source for the production of new NPs and especially antibiotics are bacteria. The majority of NPs derived from bacteria are produced by the family of *Actinomycetaceae*, which are filamentous Gram-positive bacteria. SMs from microorganisms, in particular bacetria, are structurally diverse because of their different natural habitat and have a greater chemical diversity than NPs from plants and animals. Discovery of bacterial metabolites started with the detection of “Actinomycin”, an antibiotic isolated from *Actinomycetes* in 1940. In 1950-1970s many antibiotics were discovered by isolating and screening bacteria producing various secondary metabolites. Prominent examples of discovered antibiotics were vancomycin and erythromycin, which were antibacterial agents, nystatin and amphotericin B used as antifungal agents, as well as anticancer drugs like doxorubicin and bleomycin. Because of the high interest in developing drugs for the treatment of chronic diseases, pharmaceutical companies invested into developments such as genomics, metagenomics and chemo- and bioinformatics. New ways of drug discovery were established to enable the discovery of better NPs. Additionally, efforts were made in the development new methods for isolating rare actinomycetes and cultivating previously uncharacterized bacteria from different environment. [3][2][5]

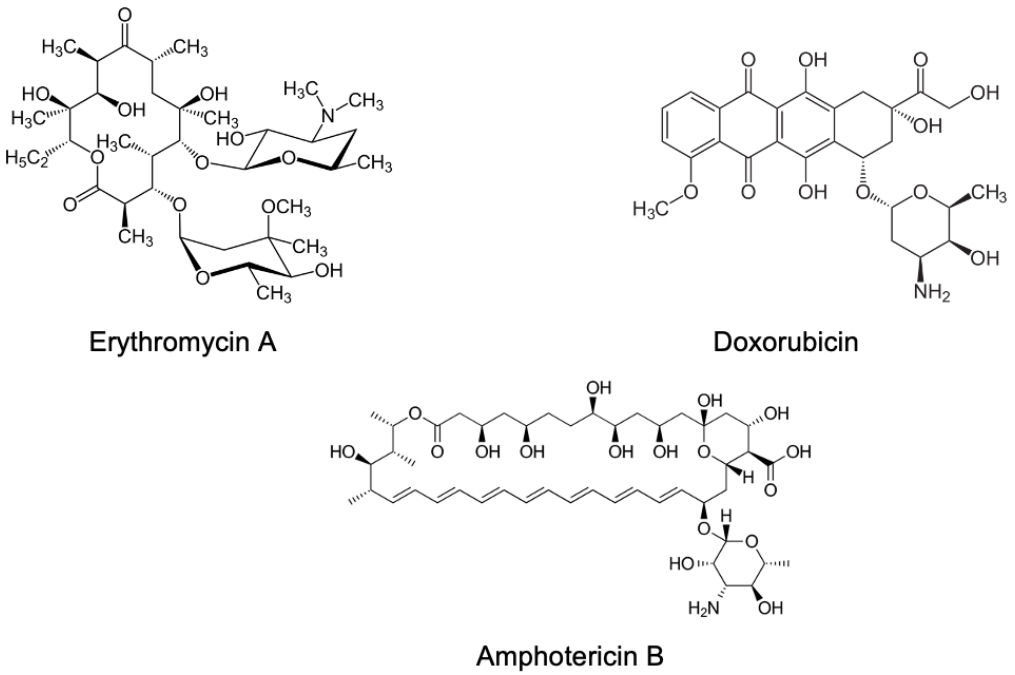


Figure 2: Chemical structures of antibiotic (Erythromycin A), anticancer (Doxorubicin) and antifungal (Amphotericin B) agents isolated from different *Streptomyces* strains. [2]

Even though actinobacteria remain an important source for the discovery of novel NPs, new approaches of investigating underexplored bacterial phyla showed interesting outcomes. By prospecting biodiversity of human microbiota two antibiotics: lugdunin and lactocillin could be detected. The latter one is a thiopeptide antibiotic, isolated from *Lactobacillus gasseri* and effective against vaginal pathogenic bacteria. Lugdunin is a cyclic peptide, isolated from *Staphylococcus lugdunensis* and effective against Gram-positive bacteria. [5]

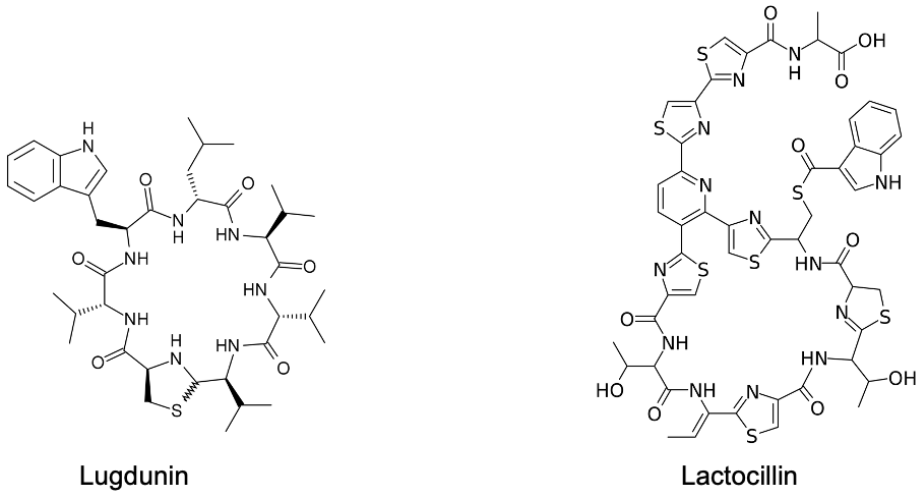


Figure 3: Chemical structures of the NPs with antibiotic properties: Lugdunin and Lactocillin. [5]

4.4. Biosynthetic gene clusters – biosynthesis of secondary metabolites

4.4.1. Definition of BGCs

NPs from microorganisms are produced through metabolic pathways encoded by genes which are co-localized in the genomes of producing strains, forming the so-called biosynthetic gene clusters, BGCs. Fig. 4 shows a schematic overview of a typical biosynthetic gene cluster.

The central part of a BGC is defined as a "*core*" which contains genes that encode enzymes necessary for the biosynthesis of the antibiotic scaffold, which is the core structure of the biosynthesized compound. Each NP group has its own typical scaffold. Furthermore, BGCs contain genes encoding transporters, resistance proteins for self-protection and regulatory proteins, that are necessary to control the transcription of the genes and thus for the regulation of the SM production. NRPS and PKS are the most famous examples of antibiotic biosynthesizing core enzymes. [6][7][8]

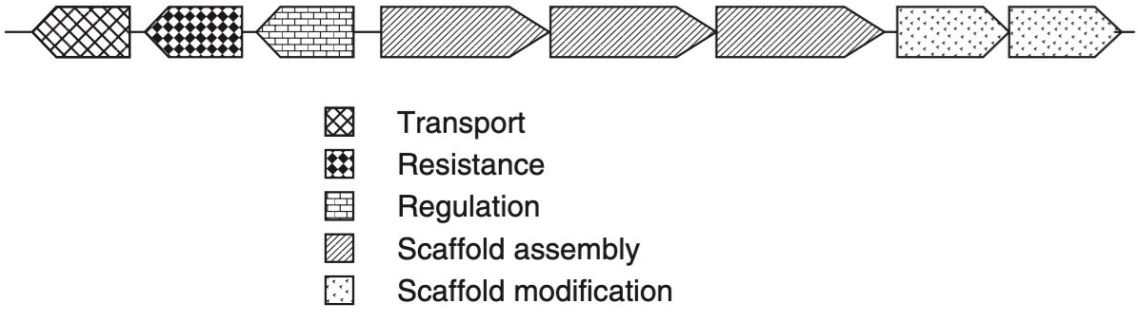


Figure 4: Schematic overview of the organization of a typical biosynthetic gene cluster. [7]

4.4.1. BGCs in single bacterial genomes and metagenomes

Non-ribosomal peptide synthetases (NRPSs)

Non-ribosomal peptide synthetases (NRPSs) are big modular enzymes that accomplish the synthesis of many biotechnologically relevant compounds, so called non-ribosomal peptides NRPs, without the ribosomal machinery and transfer RNAs (tRNA). Bacteria and fungi are the main producers of these small, but highly structurally diverse peptides, which can have antibiotic, antiviral, anticancer, anti-inflammatory and immunosuppressive activities. Some important NPs are for example daptomycin, surfactin, gramicidin S or tyrocidine A playing an important role in nowadays medicine. [9][10] Daptomycin is an antimicrobial compound acting bactericidal by disrupting the plasma membrane function. It is effective against multidrug-resistant, gram-positive bacteria and used for the treatment of blood stream infections, endocarditis and complicated skin and soft tissue infections. [11] Surfactin is known for inhibiting cancer progression by growth inhibition, cell cycle arrest, apoptosis and metastasis arrest. It shows cytotoxicity against some cancer types e.g. Ehrlich ascites, breast and colon cancers, as well as leukemia and hepatoma. [12] Gramicidin S is an antibiotic agent, but because of severe toxicity for human red blood cells, it is limited to topical applications, such as external ear infections, throat infections and root canal infections. [13] Tyrocidine A has a bactericidal effect by membrane disruption. It is more effective, when combining with linear gramicidins, resulting in tyrothricin. The latter is used e.g. topically to improve wound healing. [14]

The biosynthesis of NRPs is schematically presented in Fig. 5. **(A)** The depicted NRPS contains three modules: the initiation (blue), the elongation (green) and the termination (orange) module. **(B)** NRPS modules are activated by the addition of a 4´-phosphopantetheine (PPT) group to each T-domain. **(C)** The adenylation domain (A) selects and activates a certain amino acid by adenylation, and links it to the PPT group of the

downstream T-domain. **(D)** The linked substrate on the T-domain is then moved to the C-domain, which catalyses the peptide bond formation between the donor substrate (from the upstream module) and the acceptor domain (from the downstream module). **(E)** The assembled peptide is transferred from the upstream T-domain to the downstream T-domain. **(F)** In the last step, either another elongation step is following or, if the peptide is complete, the product is released and also often cyclized by the thioesterase domain (TE) of the termination module. [9][15]

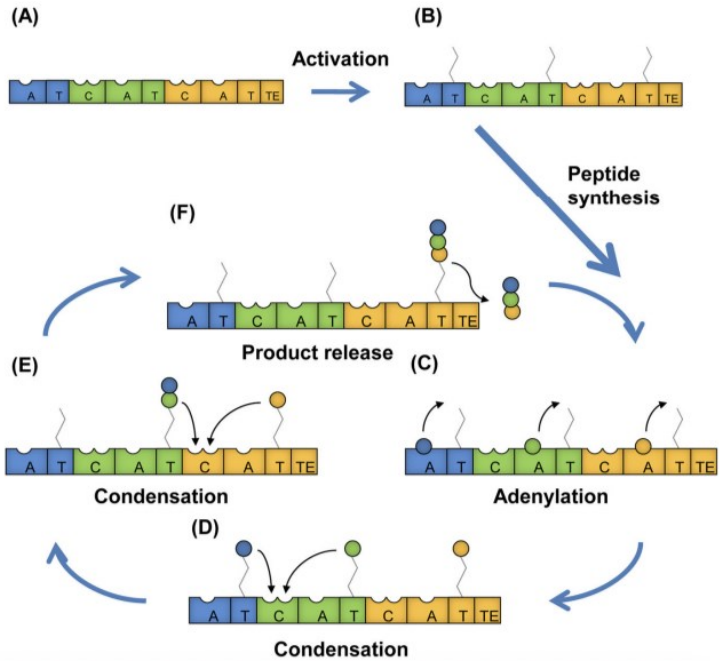


Figure 5: Schematic overview of the production of NRPs. [9]

Polyketide synthases (PKSs)

Polyketide synthases represent the second main group of large modular enzymes, which are producing polyketides (PKs). PKs are biosynthesized from acyl-CoA precursors (most often malonate and methylmalonate) and can be derived from bacteria, fungi or plants. PKs have a broad spectrum of bioactivities, acting as antibiotics, immunosuppressants, antitumorals, cholesterol-lowering agents, antiparasitics and insecticides. Prominent representatives are e.g. tetracycline, daunorubicin, erythromycin, rapamycin and lovastatin. [16][10]

The initiation module contains the following domains: an acyltransferase (AT) domain for substrate selection and the transmission to the acyl carrier protein (ACP) domain within the same module, for carrying the extending intermediates.

The elongation module consists of a ketosynthase (KS) domain, which catalyzes a decarboxylative Claisen condensation between the polyketide chain and the extender unit, as well as an AT domain and an ACP domain.

Before the newly produced intermediate was translocated onto the KS of the next module, it can be modified by additional domains, like ketoreductase (KR), dehydratase (DH), enoylreductase (ER) and methyltransferase (MT).
The termination module contains additionally a thioesterase (TE) domain being responsibe for the release of the complete polyketide by either hydrolysis or by macrocyclization.
PKSs are grouped into three main types: type I, II and III PKSs. [10][17][18][16][19]
Although all three types are structurally and mechanistically diverse, all of them biosynthesize polyketides by using the same process, namely by sequential decarboxylative condensation of the precursor acyl-CoAs catalyzed by the KS domains.

Type I PKSs are multifunctional enzymes organized into modules, each catalyzing one polyketide chain elongation and reductive modification cycle. The number of modules is corresponding to the acyl-CoA units incorporated into the final polyketide product.
The macrolide antibiotic, erythromycin A is produced by PKS type I. Its biosynthesis pathway can be divided into two steps. The first step is creating the polyketide backbone, 6-deoxyerythronolide, catalyzed by PKSs (comprised of three different deoxyerythronolide B synthases [DEBS 1-3]) through condensation of one unit propionyl CoA and six units methylmalonyl CoA. The second step is about the tailoring of 6-deoxyerythronolide with specific enzymes, such as hydroxylases, glycosyl transferases and methyl transferases. [16][20][18][21][22]

Type II PKSs harbor similar core catalytic domains like type I PKSs, except containing two KS domains: KS_{α} (responsible for the condensation) and KS_{β} (controls the polyketide length). Enzymatic activities ar present on individual proteins, working iteratively as a complex. PKSs are responsible for the biosynthesis of polycyclic aromatic PKs like tetracenomycin C. [16][20]

The last group, type III PKSs belong to the chalcone synthase (CHS) and stilbene synthase (STS) family of PKSs. They work as homodimers while functioning iteratively. Type III PKSs use free CoA thioesters as substrates to create polyketides by condensation, working independently from the ACP domain. Type III PKSs, such as RppA synthases, synthesize typically aromatic monocyclic or bicyclic polyketides, e.g. flavolin. [16][20][23]

Special cases are existing where the mix of type I PKSs and NRPSs create hybrid enzymes, called PKS-NRPS, yielding the production of unusual SMs, such as aspyridone, cyclopiazonic acid, flavipucine and pseurotin A. [10]

4.5. Genome mining – definition and approaches

Although actinomacete bacteria harbor up to 50 BGCs in their genomes, most of those are "*cryptic*" or "*silent*" and not expressed under laboratory conditions or their related compounds are produced in very small amounts. Hence, new approaches for NP discovery have been established, for example the so-called "genome mining", a combination of bioinformatics-based identification of BGCs and the type of NPs they specify, and genetic engineering aimed at activation of these BGCs. [24][25][26][5]

Fig. 6 represents three different approaches to genome mining. The phenotype-based approach is used, where the isolated microbial cultures are screened and bioactivity of extracts is analysed via bio-assays. Once the interesting bio-activity is observed and having the genome sequencing data on the hand, it is possible to predict which BGC might encode the active compound. To proof the BGC-compound connection, a "core" gene for the biosynthesis could be deactivated and the extract's bioactivity can be tested again. Activity will be lost, if the deactivated BGC is responsible for bio-active metabolite production. Another approach used the target-based strategy, where the choice of a specific chemical class of NP determined upon analysis of BGC with software antiSMASH [27] is the starting point. In this case, the genome sequencing data will be specifically scanned in order to find the chemical class of interest. When the BGC is identified, then this can be expressed in in a well-known or engineered heterologous host. The metabolite profile of a recombinant strain will be analyzed for example via HPLC and if the expression was successful and the produced compound could be detected via analytical method, the isolation and characterization of a target molecule can be processed. The last approach which is represented in Fig. 6 is based upon the uniqueness of the bacterial isolate that may be a promising source of novel metabolites. Here, the genome of an unique isolate will be sequenced and the data will be analyzed in order to identify putative new BGCs. The expression of chosen BGC could be activated or up-regulated by genetic manipulation of transcriptional regulators (described below) and by comparison of metabolomic profile of extracts from the recombinant strain and the wild type. [25][26][5]

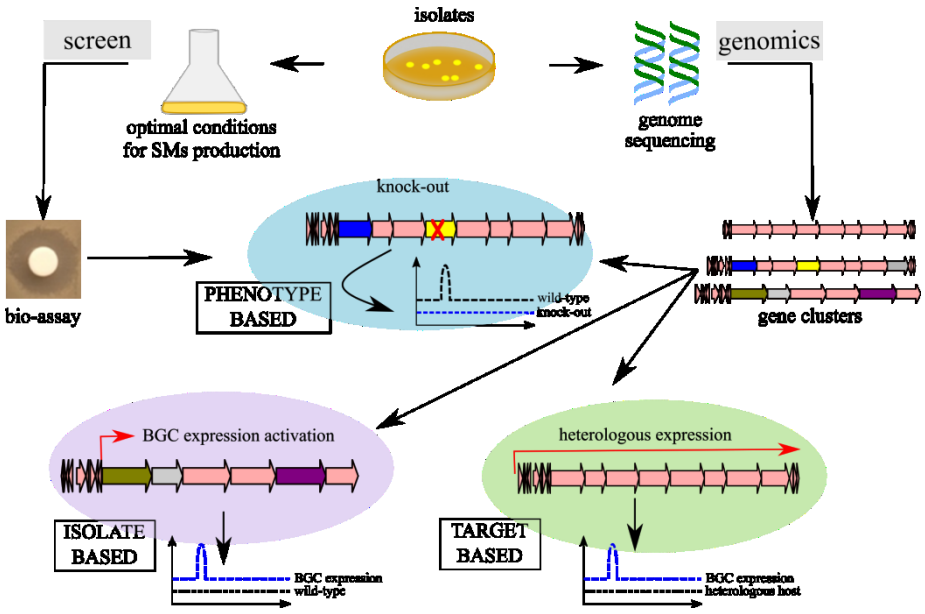


Figure 6: Three different approaches used in bacterial genome mining for the discovery of new NPs. [5]

The most genetically characterized strain from actinomycete species is *Streptomyces coelicolor.* More than twenty years ago its genome was completely sequensed. With the help of many genetic tools, it is possible to genetically manipulate this bacterium in order to heterologously produce polyketides and non-ribosomal peptides. [26]

4.5.1. Bioinformatic tools for the identification of BGCs

The importance of predicting BGCs in the genomes with following goal to activate the production of their cognate SMs cannot be overestimated for the discovery of new compounds. Increased availability of high-quality genome sequencing, in addition to the development of online software has advanced the discovery of novel metabolic pathways and NPs. [28]

Some examples of bioinformatic online tools are: NP.searcher, antiSMASH and SMURF. [29]

NP.searcher is a program designed to analyse microbial genomes to identify NRPS, PKS or hybrid NRPS-PKS clusters. [29] It is specialized for the prediction of putative chemical structures for the compounds specified by different gene clusters. [27] BLAST (Basic Local Alignment Search Tool) is used for this online tool, to search for similar regions between biological sequences and compares protein or nucleotide sequences to sequences saved in a database, NCBI (National Center for Biotechnology Information). [29]

antiSMASH is an online tool which identifies regions in the microbial genomes that encode enzymes of specialized metabolites such as PKs, NRPs, terpenes, aminoglycosides, aminocoumarins, beta-lactams, siderophores etc. [29] It predicts, analyzes and detects

BGCs in bacterial genome sequences by comparing identified regions from the gene cluster to their closest matches from a database, MIBiG, which contains all previously characterized BGCs. [27][5]
SMURF is used to analyse the genomes of fungi. The tool predicts clustered genes for SM biosynthesis based on their genomic context and domain content. [29] It is capable of creating a broader list of SM biosynthesis gene clusters, although its limited in further detailed analysis. [27]

4.5.2. OSMAC

In the last few decades the discovery of bioactive secondary metabolites declined because of the limitations of conventional cultivation and isolation methods. To overcome this problem, new approaches had to be developed. One of such approaches is defined as OSMAC – One Strain Many Compounds. The underlying idea behind the OSMAC is to activate silent BGCs in microorganisms in order to increase the chemical diversity and thus enable identification of new natural products. Most of the BGCs are silent under laboratory conditions, but can be activated by, for example, changing the culture medium. Basically, carbon and nitrogen sources are important for producing biomass, essential proteins and nucleic acids. The change of cultivation conditions is also very important for activating BGCs. Cultivating the same strain either in liquid or solid medium under static or dynamic condition leads to differently produced bioactive compounds. Performing co-cultivation with two or more different strains may also have a positive impact on production of already known or still not detected compounds. Additionally, important and successful examples of using OSMAC are exemplified by the addition of epigenetic modifiers or biosynthetic precursors, which can modify microbial physiology by inhibiting some active enzymes or by supporting the development of alternative metabolic pathways to yield the production of new NPs. [30]

4.5.3. Biosynthetic gene cluster activation in native host

As already mentioned, most of the BGCs are silent under laboratory conditions and have to be activated in order to yield production of NPs. Another approach to discover bioactive compounds from cryptic BGCs is the manipulation of transcription factor-encoding genes that control SM production pathway-specifically. This offers a good possibility of control and predictability when information from bioinformatics analyses is available, although the throughput is not very high.

Usually, pathway-specific regulators are located together with genes for resistance, biosynthesis and secretion in BGCs.

The expression of many BGCs is powered by overexpression of activator genes or deletion of repressor genes, which is described in the following chapters. [31][32][33]

When transcription factors act as repressors, they can prevent the expression of a specific gene or sets of genes in different ways. For example they compete with an activator for binding at the promotor region or they bind directly at the promotor region and prevent access of the RNA polymerase, which leads to an inhibition of transcription initiation. They can also bind downstream of the promotor and thus prevent the elongation of transcription. [34]

A prominent example for a pathway-specific repressor in *Streptomyces* is the tetracycline repressor family (TetR). A few TetR proteins contain a DNA-binding domain at their N-terminus and a ligand-binding domain at the C-terminus, which can undergo a conformational change upon interaction with a ligand. [33]

The classical, *Escherichia coli* TetR is the regulator of the expression of the tetracycline resistance gene encoded by *tetA*. In absence of tetracycline two TetR dimers bind to palindromic sequences between *tetA* and *tetR* and represses the expression of *tetA*, which encodes an efflux pump to export tetracycline out of the cell before it can inhibit the protein synthesis in bacteria. Binding of tetracycline to TetR induces conformational change that prevents TetR from binding the DNA. This protein then no longer inhibits the expression of *tetA* gene and tetracycline is pumped out of the cell. [35][36]

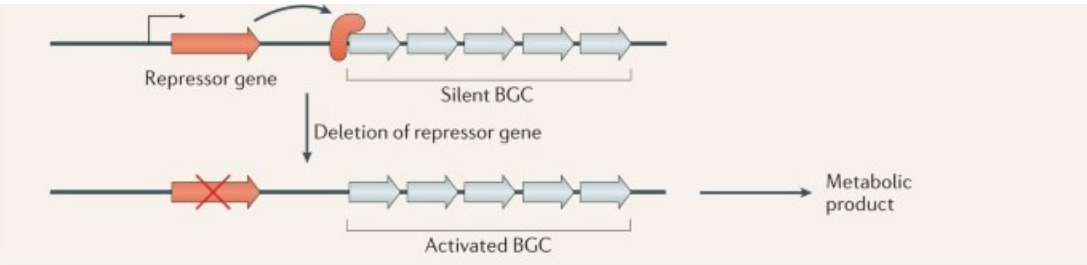


Figure 7: Schematic overview of the inactivation of a repressor gene which was the reason for a silent BGC. (modified from [31])

Transcription factors, which are able to act as activators of silent BGCs stabilize the initial polymerase-promotor complex or accelerate the transition to an open complex after the RNA polymerase bound the promotor. [34]

After the identification of a gene which encodes an activator, it is amplified via PCR and subsequently cloned into a vector under control of a strong constitutive promotor. The

resulting recombinant plasmid is then integrated into the genome of the host strain in order to overexpress the regulatory gene and thus turn on SMs production.

There are two major families of pathway-specific positive regulators, which are found in *Streptomyces:* the SARP- and the LAL-family. The latter was first discovered in proteobacteria and contains at the N-terminus an ATP-binding domain and at the C-terminus a helix-turn-helix domain for binding the DNA. SARPs are only found in actinomycetes. Their N-terminus contains a helix-turn-helix domain to bind the promotor region of the gene of interest and on their C-terminus a bacterial transcriptional activator domain is located. Most SARP-family members act at the end of signal transduction cascade and lead to an up-regulation of biosynthetic genes in the specific cluster.

The SARP-encoding genes are present in different BGCs, like type I and type II PKS clusters, nonribosobal peptide synthetase clusters, ß-lactam and lantibiotic clusters.

Famous examples of the SARP-family are ActII-ORF4 and DnrI which activate the production of actinorhodin in *S.coelicolor* and daunorubicin in *S.peucetius*, respectively. [33]

Another approach to activation of silent BGCs is to insert a constitutive promotor like *ermE*p* upstream of a pathway-specific positive regulator in order to activate the expression of a silent BGC. [37][38]

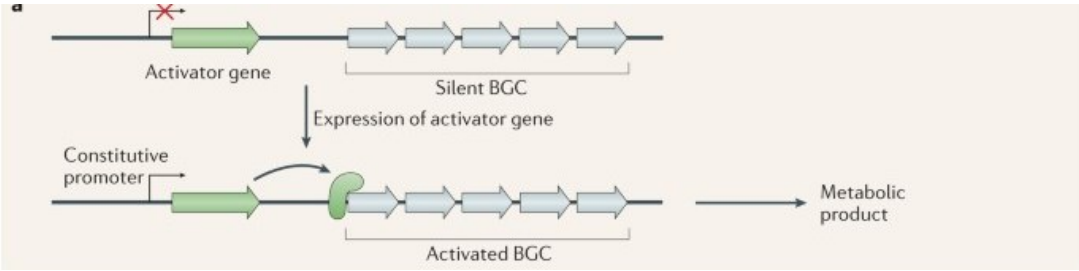


Figure 8: Schematic overview of the overexpression of an activator gene. (modified from [31])

4.5.4. Heterologous expression

Heterologous expression is a BGC-specific approach of genome mining to access, activate and express silent BGCs of Actinomycetes in engineered host strains, through which the discovery of new bioactive compounds can be achieved. This approach may not only lead to a higher yield in production levels but can also accelerate the process by using organisms which are rapidly growing and easy to modify compared to the original producer strain. [8][39][25][40]

The workflow of heterologous expression is as following: 1) isolation of the DNA from native producer and cloning into suitable vectors (usually cosmids, fosmids, or bacterial artificial chromosomes (BAC)) to generate genomic libraries. 2) screening of the generated genomic library to identify clones containing BGCs of interest. 3) reconstruction of the BGCs through homologous recombination in yeast. 4) introduction of the entire BGCs into a heterologous host, coupled with genetic manipulation for a greater expression. 5) optimization of the compound production of recombinant strains by optimizing e.g. growth conditions or fermentation media. [31][41]

Some key elements that are important for a successful heterologous expression of BGCs include a suitable stable vector carrying the BGC of interest, a suitable heterologous host, to enable the BGC´s expression, as well as the suitability of genetic manipulation. Even though appropriate host organisms and vectors were used, this must not guarantee a successful expression of the BGC. [41]

For capturing large BGCs (up to 200 kb in size) BAC vectors are used. Smaller DNA inserts with around 45 kb can be carried by cosmids and 100-300 kb DNA fragments are harbored by P1-derived artificial chromosome (PAC) vectors. [25] These vectors may thus carry an entire BGC. For successfull and sustainable heterologous expression in the host strain, a vector shall preferentially encode an integrase for stable integration into the host genome. [42]

For the selection of a suitable host for a heterologous expression, it is obligatory to know about genetic and physiological properties of native and host strain producers. Furthermore, a better expression of the BGC can be reached when heterologous organism and native producer are closely related. If this is not the case, an exchange of promotors can provide for increased expression and thus higher compound yields. [41]

4.5.5. Combined approaches of heterologous expression coupled with overexpression of activators or replacement of promotors

A limitation of bacterial producers of new NPs stems from some of them being hardly cultivalable and not suitable for industrial setting, which requires fast growth and accumulation of enough biomass to sustain high yield of the product. Hence, heterologous production in some bacterial host is the only alternative. *E.coli* and *Streptomyces* can be such suitable hosts. The former is simple to cultivate and its toolbox for gene manipulation is

well-developed, although some SMs cannot be produced in this bacterium because of the lack of certain precursors, sigma factors for the recognition of foreign promotors, protein folding problems or toxicity of the final compounds. To overcome these hurdles, bacteria of the genus *Streptomyces* were developed as heterologous hosts. Their capacity of producing precursors, performing post-translational modification and availability of resistance genes to counteract toxic metabolites is a big advantage compared to *E.coli*.
Because of the fact that *Streptomyces* synthesize often their own SMs, which can lead to competition with heterologous pathways and complications with the purification of NPs, such endogenous pathways have to be eliminated via deletion of the corresponding BGCs. In the well-known host *Streptomyces coelicolor* M145, four BGCs were deleted in order to enhance the precursors' pool and to create a simplified metabolic profile. [5]

The combination of several methods of genome mining is becoming more common, as exemplified for the strain *Streptomyces albus* J1074, which is often used for heterologous expression. Five BGCs in this strain's genome were manipulated by either placing a strong constitutive promotor in front of the biosynthetic genes, overexpressing pathway-specific regulators or via gene deletion. This combined approach has resulted in an isolation of new polycyclic tetramate macrolactams, an activation of candicidin and antimycin biosynthesis and identification of paulomycins BGC. [5]

4.6. Biosynthetic gene cluster cloning techniques

4.6.1. In vitro

When it comes to heterologous expression, cloning of large BGCs often spanning over 100 kb represents a technical challenge. *CATCH* - Cas9-assisted targeting of chromosome - is an *in vitro* cloning technique for large DNA fragments up to 100 kb in a single step. This method is using RNA-guided Cas9 nuclease, which cuts bacterial chromosomes at specific target sequences in low melting temperature agarose gel. The resulting digested fragments can then be ligated into a cloning vector, which has a 30 bp overlap at both ends with the target DNA by *Gibson assembly*. After the recombinant plasmid is generated, it is electroporated into a cloning host. [42][43] Gibson assembly is a single-pot, isothermal strategy to assemble overlapping DNA fragments by using a 5´exonuclease, a DNA polymerase and a DNA ligase. [44] Examples using this system was cloning the 36 kb jadomycin BGC from *Streptomyces venezuelae* and the 32 kb chlortetracycline BGC from *Streptomyces aureofaciens*. [42][43]

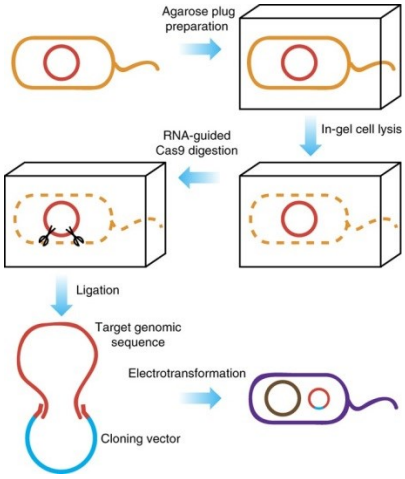


Figure 9: Overview of the *in vitro* cloning technique CATCH. [43]

4.6.2. In vivo

The TAR (Transformation Assoicated Recombination) system which utilizes natural homologous recombination of DNA fragments in *Saccharomyces cerevisiae* upon transformation is an example for *in vivo* cloning technique. [45] It is ideal for direct cloning of BGCs with a size of up to at least 150 kb. [42]

DNA representing flanking fragments of the BGC of interest can be amplified via PCR, generating a “capture” vector. Transformation of such vector into yeast cells, together with overlapping DNA fragments representing a BGC result in assembly of complete BGC in the vector. The resulting construct can then be introduced into an engineered host strain for heterologous expression. [39]

Fig. 11 shows an overview of the TAR cloning system.

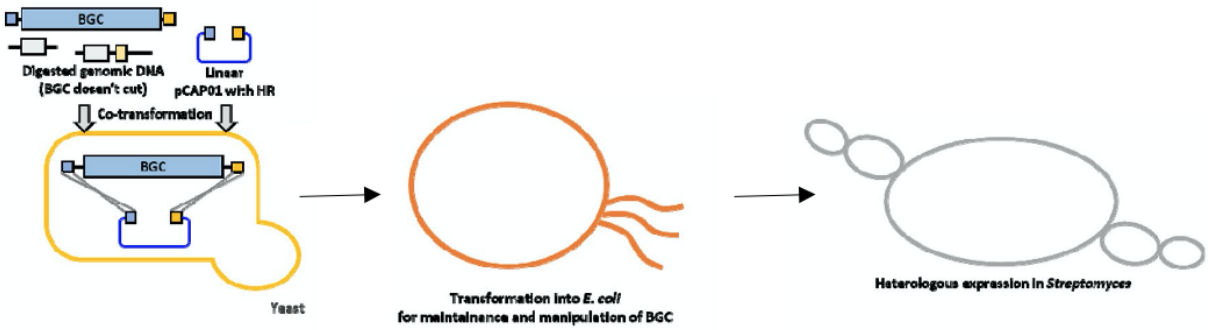


Figure 10: Schematic overview of the cloning system: TAR. (modified from [45])

The DNA isolated from environmental samples can also be used for the construction of genomic libraries capturing natural product gene clusters. Cosmid vectors, which can carry between 31 kb and 44 kb of foreign DNA, are used for the creation of such libraries.

However, many BGCs can reach over 100 kb in size, and cosmid vectors are not capable to accommodate such large fragments. To overcome such obstacles, alternative vectors capable of harbouring bigger DNA fragments, such as Bacterial Artificial Chromosomes (BACs) and P1 artificial chromosomes (PACs) can be used. BACs are able to carry inserts with a size of 60 kb to 150 kb, while PACs are capable of carrying fragments up to 700 kb. However, it is also possible to reassemble large NP gene clusters by using smaller overlapping clones. [39][46]

Another *in vivo* cloning technique for BGCs is a phage-mediated homologous recombination based on the *bacteriophage λ Red system*. This system can catalyze recombination between short homology regions and works even without RecA, an important protein for *E. coli* homologous recombination. λ bacteriophage encodes its own recombination functions and contains genes, *exo* and *beta*. *Exo* degrades dsDNA from 5´end and exposes 3´overhanging bases which can be bound by *beta.* When the complimentary strand is present, *beta* is able to anneal the two strands in order to generate recombinants. Another gene, *gam*, ensures full recombination potential and inactivates RecBCD enzyme, which can degrade incoming DNA. [47]

5. Aim of the work

The focus of this work was on the rare actinobacterial strain *Actinoalloteichus fjordicus* DSM 46856 which was isolated from a marine sponge *Geodia barretti* collected at the Trondheim fjord in Norway.
During this project two interesting biosynthetic gene clusters, BGC3 and BGC4 of this strain were chosen for heterologous expression, because according to the *in silico* analysis of genome sequence, those two BGCs may potentially encode new NPs.

The aim of this work was the construction of a genome library of *A. fjordicus* in *E. coli*, and to identify the *E.coli* clones with the BGCs of interest.

In terms of getting access to the putative bioactive compounds, both BGCs had to be assembled into the shuttle vector pCLY10 in yeast using TAR. Subsequently, heterologous expression was planned in three different engineered host strains: *Streptomyces albus* B4*, Streptomyces albus* J1074 and *Streptomyces coelicolor* M1154.

The *Streptomyces* strains containing BGCs 3 and 4 shall be fermented in different media and extract produced for further analysis by high-performance liquid chromatogphy (HPLC) and LC-MS/MS. Additionally, the extracts will be tested for bioactivity by disc diffusion assay.

6. Material

6.1. Media

6.1.1. Solid cultivation media

L- Agar (LA)

Bacto Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g
Agar	20 g
ddH_2O	up to 1000 ml

YPD

Bacto Peptone	20 g
Bacto Yeast Extract	10 g
Glucose	20 g
Agar	15 g
ddH_2O	up to 1000 ml

6.1.2. Liquid cultivation media

L- Broth medium (LB)

Bacto Tryptone	10 g
Bacto Yeast Extract	5 g
NaCl	5 g
dd H_2O	up to 1000 ml

6.1.3. Media for Yeast Transformation

YPD medium

Bacto Peptone	20 g
Bacto Yeast Extract	10 g
Glucose	20 g
ddH_2O	up to 1000 ml

Y1376- Agar

Yeast Synthetic Drop- out Medium Supplement	1.62 g
Yeast Nitrogen base without amino acids	6.7 g
Agar	20 g
ddH_2O	add to 960 ml
50% Glucose	40 ml

6.1.4. Media for conjugation

SFM- Agar

Soy Flour	20 g
Mannitol	20 g
Agar	20 g
Tap Water	up to 1000 ml

After autoclaving, 10 ml/L 1 M $MgCl_2$ solution was added to SFM medium if it is used for conjugation.

2x YT

Bacto Tryptone	16 g
Bacto Yeast	10 g
NaCl	5 g
H_2O	up to 1000 ml

6.1.5. Media for pre-cultures

Tryptic Soy Broth medium (TSB)

Tryptic Soy Broth	30 g
ddH_2O	up to 1000 ml

YEME

Oxoid Yeast Extract	3 g
Difco Bacto- Peptone	5 g
Roth Malt Extract	3 g

Glucose	10 g
Sucrose	340 g
ddH_2O	up to 1000 ml
After Autoclaving add:	
$MgCl_2 \times 6H_2O$ (2.5 M)	2 ml/L (5 mM final)

6.1.6. Media for Fermentation

MP1

Glucose	40 g
Yeast Extract	1.5 g
NH_4NO_3	2.5 g
$MgSO_4 \times 7H_2O$	0.5 g
KH_2PO_4	0.5 g
$CaCO_3$	3 g
ddH_2O	up to 1000 ml

SM17

Glucose	2 g
Glycerol	40 g
Soluble Starch	2 g
Soy Flour	5 g
Peptone	5 g
Yeast Extract	5 g
NaCl	5 g
$CaCO_3$	2 g
Tap Water	up to 1000 ml

R5 Medium

Sucrose	103 g
Glucose	10 g
Yeast Extract	5 g
Difco Casamino Acids	0.1 g
$MgCl_2 \times 6H_2O$	10.12 ml
K_2SO_4	0.25 g
TES	5.73 g

ddH_2O	up to 1000 ml	
Autoclave and add:		
$CaCl_2$ x $2H_2O$ (5 M)	0.4 ml	
KH_2PO_4 (0.5 %)	1 ml	
L-Proline (20 %)	1.5 ml	
NaOH (1 M)	0.7 ml	
Trace elements as for R2YE	2 ml	

MYM

Maltose	4 g	pH 7.3
Yeast Extract	4 g	
Malt Extract	10 g	
Tap Water	500 ml	
ddH_2O	up to 500 ml	
Autoclave and add:		
Trace elements as for R2YE	2 ml	

6.2. Stock Solutions

6.2.1. Antibiotic stock solutions

Table 1: Concentration and solvents used for the preparation of Antibiotic stocks.

Antibiotic	**Stock concentration (mg/ml)**	**work concentration *E.coli* (µl/ml)**	**work concentration *Streptomyces* (µl/ml)**	**Solvent**
Ampicillin (Amp)	100	100	-	ddH_2O
Apramycin (Am)	100	100	50	ddH_2O
Chloramphenicol (Cml)	25	25	-	Ethanol abs.
Kanamycin (Kan)	25	25	-	ddH_2O
Nalidixic acid (Nal)	30	-	30	0.1 M NaOH
Thiostrepton (Thio)	30	30	30	DMSO

6.2.2. Different stock solutions

Arabinose 1 M

Arabinose	1.5 g
ddH_2O	up to 10 ml

Arabinose was dissolved in ddH_2O, afterwards sterile filtered and stored as aliquots at -20°C.

20% Glycerol

Glycerol	20 ml
ddH_2O	up to 100 ml

1 M $MgCl_2$

$MgCl_2$ x $6H_2O$	101.65 g
ddH_2O	up to 500 ml

50 mM $CaCl_2$

$CaCl_2$ x $2H_2O$	7.35 g
ddH_2O	up to 1000 ml

Sol I

50 mM Tris/HCl	0.61 g	pH 8.0
10 mM EDTA	0.29 g	
50 mM Glucose	0.9 g	
ddH_2O	up to 100 ml	

Sol II

0.2 N NaOH	0.8 g
1% (w/v) SDS	1 g
ddH_2O	up to 100 ml

Sol III

3 M KOAc	29.44 g	pH 4.8
ddH_2O	up to 100 ml	

10 mM Tris/ HCl (pH 8)

Tris/ HCl 1000 mM	100 µl
ddH_2O	up to 10 ml

0.1 M Lithium Acetate

Lithium acetate 1 M	10 ml
ddH_2O	up to 100 ml

Trace elements as for R2YE

$ZnCl_2$	40 mg
$FeCl_3 \times 6H_2O$	200 mg
$CuCl_2 \times 2H_2O$	10 mg
$MnCl_2 \times 4H_2O$	10 mg
$Na_2B_4O_7 \times 10H_2O$	10 mg
$(NH_4)_6Mo_7O_{24} \times 4H_2O$	10 mg
ddH_2O	up to 1000 ml

0.8% Agarose Gel + GelRed

Agarose	3.2 g
TBE buffer 1x	up to 400 ml
GelRed	20 µl

Agarose was dissolved in TBE buffer by heating in the microwave for a couple of minutes. After the solution became totally clear, GelRed was added. The agarose gel was stored at 60°C until used for gel electrophoresis.

6.3. Strains, Plasmids and Primers

6.3.1. Strains

Table 2: Strains used during this work.

Strain	Characteristics	Source
Escherichia coli XL1-Blue MR	cloning host strain	Lee et al., 2006 [48]
Escherichia coli ET12567	methylation-deficient *E.coli* strain containing pUZ8002 used during conjugations	Choi et al., 2004 [49]
Escherichia coli EPI300	Electrocompetent *E.coli* strain, with deleted mutant *trfA* gene to increase the replication rate	Lucigen Corporation
Escherichia coli EPI300-T1^{R}	phage resistant strain and a stable host for fosmids, used for library screening	Nicolaou et al., 2012 [50]
Streptomyces albus B4	contains four directly oriented *attB* sites	Myronovskyi et al., 2018 [51]
Streptomyces albus J1074	Engineered host strain	
Streptomyces coelicolor M1154	*S.coelicolor* M154 with the deletion of four gene clusters	Gomez-Escribano et al., 2013 [26]
Saccharomyces cerevisiae BY4742	His3Δ1, leu2Δ0 and lys2Δ0 are deleted, which are commonly used genes as selectable markers for recombinant plasmids	Brachmann et al., 1998 [52]

6.3.2. Plasmids

Table 3: Plasmid vectors and their characteristics used during this work.

Plasmid vector	Characteristics	Source
pCLY10	shuttle vector for *E.coli,* Yeast and *Streptomyces*, Am^{R}	Bilyk et al., 2016 [39]
pCLY10_C3_LR	shuttle vector pCLY10 with flanking fragments of BGC3 for Yeast Assembly (Am^{R})	This work
pSS_C3	shuttle vector pCLY10 carrying the whole BGC3 from *Actinoalloteichus fjordicus*	This work
pCLY10_C4_LR	shuttle vector pCLY10 with flanking fragments of BGC4 for Yeast Assembly (Am^{R})	This work
pSS_C4	shuttle vector pCLY10 carrying the whole BGC4 from *Actinoalloteichus fjordicus*	This work

6.3.3. Primers

Table 4: primers listed with a grey backround were designed for pooled-PCR library screening of BGC3 and BGC4 from *Actinoalloteichus fjordicus*. SS_LR3_fwd/rev were used for the production of sticky ends on C3_LR_flank and 5H3_fwd/rev & 1H_fwd/rev were designed to produce the missing pieces of BGC4 via PCR.

Name	Primer sequence	Size of product
AF_3_fw_1	CCGGACGAGATGTTCCTCCAA	949 bp
AF_3_rev_1	GCAACTGTTCGGTCTGACTG	
AF_3_fw_2	ACGAGCCTGATCCTCGATGAGAC	896 bp
AF_3_rev_2	GATGACACCGAGTCCGGTGATG	
AF_3_fw_3	TGTACCTCGTCGCCGATGAGTG	633 bp
AF_3_rev_3	TCGCAGGCATCACCGAATCA	
SS_LR3_fwd	CGCGTAAGCTTTACCGGCATACAGC	1050 bp
SS_LR3_rev	AATATGCGGCCGCGGAACATGCCGTAGGCGAAC	
AF_4_fw_1	ATCGCACCTGGCAACGACTGT	795 bp
AF_4_rev_1	TGATCGGTCGGTCGACGTCAGAG	
AF_4_fw_2	CTGGTGAGATCCACGTCCGATT	697 bp
AF_4_rev_2	CGTCCAGCAGGATCAGGTACT	
AF_4_fw_3	CGAGGTTGGAGGCGTTGTTCA	584 bp
AF_4_rev_3	CACTCAGGCAGATGTTGGATTC	
5H3_fwd	GCGGATCGCTTCACTCATGG	4098 bp
5H3_rev	ACGCAGATCAGCTTCTCCAG	
1H1_fwd	TTCAGGGTGGCTTGGTACTC	5267 bp
1H1_rev	GCGACGACTCAAGGAGTCAC	

6.4. Kits

Table 5: used kits during this work and their functions.

Kits	Functions	Produced by
DNA Clean & Concentrator™	Purification of PCR-products	Zymo Research
Zymoclean™ Gel DNA Recovery	Recovery of DNA from gel	Zymo Research
Monarch Plasmid Miniprep	Isolation of plasmid DNA	New England BioLabs
Wizard® Plus SV Minipreps DNA Purification System Kit	Isolation of plasmid DNA	PROMEGA
Q5® High-Fidelity 2X Master Mix	PCR	New England BioLabs

7. Methods

7.1. Preparation of media and antibiotic stock solutions

7.1.1. Media

All media were prepared by weighting the needed ingredients listed under chapter 6.1. After filling up the volume with the necessary amount of water, the liquid was mixed and autoclaved at 121°C for 20 minutes to ensure proper sterilization. If needed, antibiotics were added. Media with antibiotics were stored at 4°C, and media without antibiotic were stored at room temperature (RT).

7.1.2. Antibiotic stock solutions

After dissolving antibiotics in their respective solvents by using the right concentration (Table 1, chapter 6.2.1), antibiotic stock solutions were sterile filtered and stored as aliquots at -20 °C until used for the selection of genetic modified cells. Chloramphenicol and thiostreptone were dissolved and stored directly in the freezer without filtrating.

7.2. Cultivation of *E. coli, Streptomyces* sp. and *S.cerevisiae*

E.coli, Streptomyces sp. and *S.cerevisiae* can grow in both liquid and solid media. For the cultivation of the named microorganisms, different media were used containing the required antibiotic stock solutions. (LB medium for *E.coli*, SFM agar for *Streptomyces* and YPD medium for *S.cerevisiae*. When growing on solid media 100 µl of the glycerol stock was pipetted on the agar and spread evenly with a spatula. For the liquid media cultivation single colonies were picked with sterile toothpicks and transferred into cultivation tubes. *E.coli* grow at 37 °C and *Sreptomyces* and *S.cerevisiae* need an incubation temperature of 28 °C.

7.3. Preparation of competent *E. coli* cells

7.3.1. By using the calcium chloride method

For the preparation of competent cells, which were used for transformation, *E.coli* XL1-Blue grew in 3 ml of LB medium over night at 37°C. 500 µl of the overnight culture was added to 50 ml of fresh LB in a 200-ml flask. The shaking was continued at 37 °C until an OD_{600} of 0.4-0.5 is reached. After the culture was chilled on ice for 20 minutes it was transferred to a sterile centrifugation tube. For the collection of the cells, it is necessary to centrifuge at 6000 rpm for 8 minutes at 4 °C. The supernatant was discarded, and the cells were resuspended in 20 ml of ice-cold 50 mM $CaCl_2$. After 20 minutes of incubation the cells were again

collected by centrifugation at 6000 rpm for 8 minutes at 4 °C. The supernatant was discarded. To storage the cells for a long period as a frozen stock, they were resuspended in 2.5 ml of ice-cold 50 mM $CaCl_2$ mixed with 10% glycerol. Competent *E.coli* cells were divided in aliquots of 100 µl into 1.5 ml Eppendorf tubes and frozen in liquid nitrogen for later use.

7.3.2. Electrocompetent *E.coli* EPI300 cells

For the transformation of big DNA constructs, competent cells were prepared as the following protocol:

E.coli EPI300 cells were grown overnight in 10 ml LB mixed with the appropriate antibiotic while shaking at 37 °C. 1 ml of the overnight culture was inoculated in 100 ml fresh LB containing the required antibiotic. For getting a good efficiency the cells were incubated by shaking for 3-4 hours at 30 °C until an OD_{600} of 0.4-0.6 is reached. After the appropriate OD was reached the cells were recovered by centrifugation at 4000 rpm for 5 minutes at 4 °C in a 50 ml falcon tube. The supernatant was discarded and the pellet was resuspended by mixing in 40 ml ice-cold sterile 10 % glycerol. The cells were again centrifuged at 4000 rpm for 5 minutes at 4 °C before the pellet was resuspended in 30 ml ice-cold 10 % glycerol. After another centrifugation step the liquid was decanted and the cell pellet was finally resuspended in the remaining 500-1000 µl of 10 % glycerol. Aliquots of 50 µl were prepared and stored at -80 °C.

7.4. Preparation of glycerol stocks

7.4.1. *E.coli* glycerol stocks

A single colony of *E.coli* was inoculated in 3 ml LB and incubated overnight at 37 °C. The following day 1.5 ml was pipetted in a sterile Eppendorf tube and centrifuged at maximum speed for 10 minutes. After the supernatant was discarded, the pellet was resuspended with 1.5 ml of 20 % glycerol. Finally, the liquid was transferred to a well inscribed CryoTube™ and stored at -80 °C.

7.4.2. *Streptomyces* spore suspensions

Streptomyces were grown on appropriate media for approximately one week at 28 °C until they sporulated. *Streptomyces albus* B4, wt *and Streptomyces coelicolor* M1154 need to be incubated on SFM media mixed with the required amount of $MgCl_2$. After the strains start to sporulate, 5 ml of 20 % Glycerol were pipetted on the petri dish and the surface was scratched gently with the tip of the pipette. The soaked-up suspension was filtered through a

sterile syringe filled with cotton into a sterile tube, transferred to a CryoTube™ and stored at -80 °C.

7.5. *In silico* methods

7.5.1. AntiSMASH and BLAST

AntiSMASH (antibiotics & Secondary Metabolite Analysis Shell) is a bioinformatic online tool to predict, analyze and detect secondary metabolite gene clusters in bacterial genome sequences. Once uploaded the biosynthetic gene cluster of interest, each gene can be connected with BLAST, which identifies their individual functions by comparing similarities of the sequence to NCBI (National Center for Biotechnology Information), a freely available database.

7.5.2. Primer design

For performing DNA amplification via Polymerase Chain Reaction (PCR) the program Clone Manager was used. Therefore, forward and reverse primers were designed *in silico* and afterwards delivered by Eurofins Genomics.

All primers were designed by containing 19-33 base pairs. The melting temperature was adjusted to a range of 60 °C and 75 °C and the difference between primer pairs was set to a maximum of 5 °C.

7.6. Polymerase chain reaction (PCR)

For the PCR-reactions all the substances regarding the necessary PCR-setup listed in Table 6, 7 and 8 were added together by keeping the reaction tubes on ice following the program in Table 9 with a Master Cycler Nexus X2 from Eppendorf®.

Table 6: PCR-setup for Hot Start *Taq* DNA Polymerase.

Component	**Volume**
10X Standard *Taq* Reaction Buffer	2.5 µl
10 mM dNTPs	0.5 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
DMSO	1 µl
Hot Start *Taq* DNA Polymerase	0.125 µl
Template DNA	1 µl
Nuclease-free water	Up to 25 µl

Table 7: PCR-setup for Q5 High-Fidelity 2X Master Mix.

Component	Volume
Q5 High-Fidelity 2X Master Mix	25 µl
10 µM Forward Primer	2.5 µl
10 µM Reverse Primer	2.5 µl
Template DNA	1 µl
Nuclease-free water	Up to 50 µl

Table 8: PCR-setup for Q5 High GC Enhancer.

Component	Volume
5X Q5 Reaction Buffer	5 µl
10 mM dNTPs	0.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
5X Q5 High GC Enhancer	5 µl
Q5 DNA Polymerase	0.25 µl
Template DNA	1 µl
Nuclease-free water	Up to 25 µl

Table 9: Conditions for PCR-reactions using Hot Start *Taq* DNA Polymerase, Q5 High-Fidelity 2X Master Mix.

Step	Temperature			Time		
	Taq	**Q5 MM**	**Q5 Enhancer**	**Taq**	**Q5 MM**	**Q5 Enhancer**
Initial Denaturation	95 °C	98 °C	98 °C	3 min	1 min	3 min
25 Cycles: Q5	95 °C	98 °C	98 °C	30 sec	10 sec	10 sec
30 Cycles: Taq	58-62 °C	56 °C	1H1: 59 °C 5H3: 61/62 °C	1 min	30 sec	45 sec
	68 °C	72 °C	72 °C	1 min	50 sec	1H1: 6 min 5h3: 4 min
Final Extension	68 °C	72 °C	72 °C	7 min	7 min	7 min
Hold	10 °C	4-10 °C	4-10 °C			

7.7. Library production

The aim of the production of a genomic library is to construct a collection of random gDNA fragments that make up the full-length genome of *Actinoalloteichus fjordicus.*

7.7.1. Shearing and End-repair of the gDNA

High-molecular-weight gDNA was isolated in TE buffer at a concentration of 0.5 µg/µl. For the ligation with CopyControl pCC1FOS vector a DNA size of approximately 40 kb-fragments was necessary. Therefore, the DNA was divided into several tubes to test how many times it is necessary to pipet for reaching the needed size of DNA fragments. A gel electrophoresis

was performed and DNA ladder was used as a reference. After shearing the DNA for 160 times the end-repair of the insert DNA was performed to remove 3´-overhangs and to create blunt ends. For the end-repair a master mix was created by placing the compounds listed in table 10 on ice and mixing them thoroughly together.

Table 10: Master Mix for the End-Repair of the Insert DNA.

Compounds	**Volume**
Sterile Water	32 µl
10X End-Repair Buffer	8 µl
2.5 mM dNTP Mix	8 µl
10 mM ATP	8 µl
sheared insert DNA	20 µl
End-Repair Enzyme Mix	4 µl
Total Reaction Volume	**80** µl

The master mix was centrifuged and incubated at room temperature for 45 minutes. After the inactivation of the End-Repair Enzyme Mix at 70 °C for 10 minutes gel electrophoresis was performed on a 0.5% agarose gel, overnight at 35 V. The following day the gel was cut and the GeneRuler High Range DNA Ladder and one pocket of the DNA was stained with GelRed by shaking gently at room temperature for 30 minutes. The gel was assembled together and cut with UV-light at approximately 48 kb of the DNA sample of *Actinoalloteichus fjordicus*. The gel slices were mashed with a scalpel, placed into 2 ml Eppendorf tubes and the weight was determined.

7.7.2. Recovery of DNA and Ligation with CopyControl pCC1FOS Vector

- **Recovery of DNA from 0.5% agarose gel**

Based on the weight of the gel slice the same volume of solvents were used. One volume of buffer-equilibrated phenol was added to the tube by vortexing for 10 seconds. After the solution was cloudy the tube was frozen at -70 °C for 20 minutes. When the time was up the Eppendorf tube was centrifuged at room temperature for 15 minutes at 13600 rpm. The upper phase containing the DNA was collected and placed into a new tube. One volume of phenol/chloroform and 0.1 volume of 3M sodium acetate were added, mixed gently and centrifuged at room temperature for 15 minutes at 13600 rpm. The upper phase was collected in a clean 1.5 ml tube on which one volume of ice-cold isopropanol was added. At this step a smear appeared and the tube was gently inverted until the liquid was clear. After 15 minutes of centrifugation at room temperature at 13600 rpm the supernatant was removed. The pellet was washed with 0.7 volume of ice-cold 70% ethanol before spinning

the tube for another 15 minutes at room temperature at 13600 rpm. The supernatant was discarded and the pellet was dried on air for around 10 minutes. When the ethanol was evaporated, 40 µl of rehydration buffer from the Wizard Genomic DNA Purification Kit were added. The dissolving of the DNA was carried out overnight at 4 °C. To ensure the success of this procedure another gel electrophoresis was performed with 2 µl of fDNA on 0.5% agarose gel.

- **Ligation with CopyControl pCC1FOS Vector**

For the ligation reaction the following reagents listed in table 11 were combined and mixed thoroughly after each addition.

Table 11: Compounds for the ligation reaction of the end-repaired DNA with CopyControl pCC1FOS vector.

Compound	Volume
Sterile Water	0 µl
10X Fast-Link Ligation Buffer	1 µl
10 mM ATP	1 µl
CopyControl pCC1FOS Vector	1 µl
concentrated insert DNA	6 µl
Fast-Link DNA Ligase	1 µl

The tube was incubated at room temperature for 4 hours, before it was placed into a heating block at 70 °C for 10 minutes to inactivate the Fast-Link DNA Ligase. The tube was frozen and stored at -20 °C.

E.coli EPI300-T1R cells were spread on a LA-agar plate and grown overnight at 37 °C. The following day a single colony of EPI300-T1R cells was inoculated in 50 ml LB + 10 mM $MgSO_4$ + 0.2% Maltose and incubated overnight at 37 °C while shaking at 200 rpm.

7.7.3. Packaging of the ligated DNA and transformation into *E.coli* EPI300-T1^R

- **Packaging of the ligated DNA**

On the day of the packaging reactions, 50 ml LB + 10 mM $MgSO_4$ + 0.2% Maltose were inoculated with 0.5 ml of the EPI300-T1R overnight culture. The flask was shaken at 37 °C to an OD_{600} of 0.8-1.0. After the appropriate OD was reached the cells were stored at 4 °C until needed. One tube of the MaxPlax Lambda Packaging Extracts was thawed on ice and immediately 25 µl were transferred to another 1.5 ml Eppendorf tube and placed on ice. 10 µl of the ligation reaction was added to the thawed extracts while being held on ice. The solution was mixed by pipetting several times without the introduction of air bubbles. After centrifuging the tubes shortly, the packaging reactions were incubated at 30 °C for 2 hours. When the time was up 25 µl of MaxPlax Lambda Packaging Extract were added to the

packaging reaction. The tubes were incubated for an additional 2 hours at 30 °C. At the end of the second incubation, 940 µl Phage Dilution Buffer (PDB) were added and combined. After the addition of 25 µl chloroform, the tubes were inverted gently and stored at 4 °C.

- **Transformation into *E.coli* EPI300-T1R**

Several dilutions of the 1 ml of packaged phage particles were created: $1:10^2$ and $1:10^3$. 10 µl of each of the dilutions and 10 µl of undiluted phages, individually, were added to 100 µl of the previously prepared EPI300-T1R cells and the tubes were incubated at 37 °C for one hour. The infected EPI300-T1R cells were spread on a LA-agar plate with 12.5 µg/ml chloramphenicol and incubated overnight at 37 °C to select packaged CopyControl Fosmid clones. The following day the grown colonies were counted and based on the titer of the packaged CopyControl Fosmid clones and the estimated number of clones that are required, the volume that was needed to prepare the genome library, was calculated. 3000 µl EPI300-T1R were mixed with 300 µl of phage particles. The tubes were incubated at 37 °C for one hour. The infected bacteria cells were spread on a LA-agar plate with 12.5 µg/ml chloramphenicol and grown overnight at 37 °C.

7.7.4. Preparation of 96-well plates

The plates were labeled properly and a mix of LB-medium, 1 mM Arabinose and 12.5 µg/ml chloramphenicol was created. 200 µl of this mix was pipetted in each of the 96 holes for all of the 18 plates. Single overnight grown colonies were picked with previously autoclaved toothpicks and placed into the holes by rotating in the liquid. The plates were incubated overnight at 37 °C. The next day 100 µl of 50% glycerol was added, mixed and finally the completed packaged fosmid library was stored at- 80 °C.

7.8. Library screening

The whole genome of *Actinoalloteichus fjordicus* was harbored in 18 96-well-plates. The genome library screening was performed according to the following protocol: The stamp was flamed two times with 96% ethanol. After the liquid was evaporated, the stamp was dipped into the left side (48 holes) of the well-plate and immediately stamped on a LA + 12.5 µg/ml Cml petri dish. This step was repeated while stamping onto a second petri dish, which was needed as a backup. After flaming the stamp again with 96% ethanol the procedure was continued with the right side of the microtiter plate. The petri dishes were incubated over night at 37 °C. The next day the colonies were isolated following the protocol "Isolation of fosmid DNA of *E.coli* from petri dishes" to receive fDNA. Backup plates were stored at 4 °C.

7.9. Isolation of fosmid DNA of *E.coli* from solid medium for pooled PRC screening

The isolation process starts with colonies of each petri dish. After performing PCR-reactions and making gel electrophoresis, only *E.coli* colonies, which showed a band on the gel were isolated according to the following protocol:

2 ml of LB medium was pipetted onto each Petri dish and the colonies were carefully washed down with the tip of a blue pipette or an inoculation loop. By holding the plate at an angle, the liquid was drawn up and transferred to 2 ml Eppendorf tubes. After centrifuging the tubes for 1 min at 13200 rpm, the supernatant was tipped off. 200 µl of cold Sol I were mixed with RNase (10 µl RNase (stock: 10 mg/ml)/ 10 ml Sol I), added to the Eppendorf tube and vortexed. After the addition of 400 µl Sol II, the tubes were inverted 5 times and the solution became clear. 300 µl of cold Sol III was pipetted and inverted 5 times, while a white precipitate was formed. After each tube was centrifuged at 13200 rpm for 5 min, the clear supernatant was transferred to a fresh 2 ml Eppendorf tube. 1 ml of ice-cold isopropanol was added and inverted until no streaks were visible and the liquid was completely clear. After centrifugation at 13200 rpm for 10 minutes, the supernatant was discarded. The pellet was washed with 200 µl of ice-cold 70% ethanol (pipetting only on top) before centrifuging the tubes for another 5 minutes at 13200 rpm. The supernatant was tipped off and the Eppendorf tubes were dried upside down on a paper for 10 minutes. To ensure complete removal of the ethanol, the tubes were placed in the heating block at 65 °C for another 10 minutes. Finally, 100 µl of sterile water were added to the pellet and stored at -20 °C until use.

7.10. Isolation of fosmid DNA of *E-coli* from liquid medium for single colony screening

A large colony was inoculated in 20 ml LB medium containing chloramphenicol (0.5 µl/ml) and L-arabinose (1 µl/ml). The incubation was carried out overnight at 37°C while shaking at 200 rpm.

The next day 20 ml liquid culture was centrifuged at 13000 rpm for 1 minute at 4°C. After the addition of 800 µl of ice-cold Sol I, the pellet was resuspended by vortexing. The liquid was divided equally among 4 Eppendorf tubes (2 ml). Immediately 400 µl of Sol II were added by inverting the tubes 10 times. Quickly 300 µl of ice-cold Sol III were added and mixed by inverting the tubes 5 times. The tubes were spun at full speed in a microcentrifuge for 5 minutes at 4°C. After the clear supernatant was transferred to a new Eppendorf tube, 300 µl of Protein Precipitation solution were added, mixed by inverting and incubated on ice for 10

minutes. The tubes were spun at full speed in a microcentrifuge for 5 minutes at 4 °C. The clear supernatant was transferred to a fresh Eppendorf tube and 2µl RNase (stock 10mg/ml) were added to each tube. After incubating for 1 hour at 37°C, the tubes were cooled down on ice for 5 minutes. 0.8 Vol. 2-propanol were added and mixed by inverting until no streaks were visible. The tubes were left on ice for 10 minutes. They were spun at full speed for 15 minutes at 4 °C and the pellet was washed with 500 µl of 70% ethanol. The tubes were spun again at full speed for 15 minutes at 4 °C, before the liquid was drained away and left at room temperature for 10 minutes to let the pellet dry. The pellet was resuspended in 100 µl 10 mM Tris/HCl (pH 8) and stored 4 °C.

7.11. Restriction digestion

Restriction digestion was a necessary step for the preparation of linear fragments for the assembly reaction in yeast cells. After the isolated fDNA was sequenced, a restriction digestion with specific endonucleases was performed. All restriction reactions were mixed after the same concept: DNA of interest, buffer, ddH_2O and the required enzymes were mixed together by vortexing before they were centrifuged. The appropriate temperature and duration of the restriction is depending on the different endonucleases. During this project, fosmids listed in table 12 were used. The program Clone manager allowed a prediction of restriction enzymes, which cut the fosmids on the right positions for performing yeast assembly.

Table 12: List of fosmids of BGC3 and BGC4, enzymes, their required buffers and the temperature at which the enzymes cut the fosmids.

Fosmids	Enzymes	Buffer	Temperature
BGC3			
11A7	BstZ17I PciI NdeI	3.1	37 °C
13H6	BstZ17I NheI	4.1	37 °C
BGC4			
5H3	HindIII FseI	4.1	37 °C
1H1	AvrII PciI	4.1 3.1	37 °C

7.12. Gel electrophoresis and recovery of the DNA

7.12.1. Gel electrophoresis

Gel electrophoresis was performed to characterize and identify DNA fragments, by separating it through size. There are many purposes for which gel electrophoresis is important. In this work, this technique was used to purify amplified DNA products before cloning, to validate the sizes of new DNA-constructs and for the purification after restriction digestion.

7.12.2. Recovery of pDNA from 0.8 % agarose gel

A gel electrophoresis on a 0.8% agarose gel was performed after the restriction of vectors or inserts with appropiate endonucleases. To recover the plasmid DNA, the piece of the gel containing the DNA was excised and placed into an Eppendorf tube. The process was continued following the protocol of Zymoclean™ Gel DNA Recovery Kit (Zymo Research).

7.13. Ligation

Ligation was performed for the assembly of a vector and an insert DNA. Prior to the beginning of the ligation both pieces were digested with the same restriction endonucleases. For the ligation reaction T4 DNA Ligase Buffer (10X) and T4 DNA Ligase were used, mixed with the other ingredients listed in Table 13, centrifuged and incubated for one hour at room temperature.

Table 13: Components of a ligation reaction.

Components	**Volume**
Vector DNA	1 µl
Insert DNA	5 µl
T4 DNA Ligase Buffer (10X)	2 µl
T4 DNA Ligase	1 µl
ddH2O	up to 20 µl

7.14. Transformation in *E.coli*

Transformation is a process where foreign DNA gets introduced into a bacterial cell. In this project two different methods of transformation from the DNA of interest into the E. coli strains were applied: chemical transformation and the transformation via electoporation. Electroporation will be described in Methods 7.16.4. For the chemical transformation the competent cells (Methods 7.3) were thawed on ice. After around half an hour either 1 µl of vector DNA or 20 µl of ligation reaction was pipetted to 100 µl of competent cells. The

mixture was kept on ice for another 30 minutes. The tube was transferred to a 42 °C heating block for 45 seconds. After the heat shock was performed, the tube was quickly placed onto the ice for 5 minutes. 900 µl of fresh LB medium was added and incubated for 1 hour while shaking at 37 °C at 200 rpm. Cells were spread on LA petri dishes containing the appropriate antibiotics. The plates were incubated overnight at 37 °C.

7.15. Isolation of plasmid DNA from *E.coli*

Single *E.coli* colonies from the grown plates were picked with sterile toothpicks and placed into tubes containing 3 ml LB mixed with appropriate antibiotics. The Incubation was carried out overnight in a shaker at 37 °C at 200 rpm. The next day the plasmid DNA was isolated following the protocol of the Monarch® Plasmid Miniprep Kit and finally stored at -20 °C.

7.16. Yeast Assembly technique

7.16.1. Yeast Transformation

Saccharomyces cerevisiae BY4742 was plated on YPD agar plates and incubated overnight at 28 °C. The next day a single colony was picked and inoculated in 30 ml of YPD liquid medium and grown overnight at 28 °C in a shaker. 0.4 ml of the overnight culture was pipetted in 70 ml fresh YPD medium and incubated by shaking at 200 rpm at 30 °C for approximately 4-5 h until an OD_{600} of 1.0 was reached. In the meantime, 350 µl of single stranded salmon sperm DNA (SS DNA) was pipetted in an Eppendorf tube and placed in a 100 °C heating block for 10 minutes to linearize the DNA. Once the SS DNA was cooled down on ice for 10 minutes a master mix was pipetted together using the components listed in table 14. The volume of the ingredients for the “Master Mix” was calculated for one reaction and has to be adjusted for the reactions needed for the yeast transformation. The mixture was vortexed and placed on ice.

Table 14: Ingredients and volume for Master mix for 1 reaction.

Components	**Volume**
50% PEG (MW 3350)	240 µl
1 M Lithium acetate	60 µl
SS DNA (boiled, cooled on ice)	50 µl

Separately a DNA mix with a volume of 36 µl was produced containing 2 µl linear plasmid and 10 µl fosmid DNA filled up with sterile water. To be sure that the yeast assembly was successful a positive (untreated vector) and negative (restricted vector) control were created. After an OD_{600} of 1.0 was reached, aliquots of 10 ml yeast culture were separated in 15 ml falcon tubes and centrifuged at 9000 rpm. The supernatant was decanted, the pellets were

resuspended with 500 µl sterile water and transferred to a fresh Eppendorf tube. After spinning the tubes 1 minute at 6000 rpm the water was discarded and the pellets were gently resuspended with 400 µl sterile 100 mM lithium acetate solution. Another centrifugation step was performed at 6000 rpm for 1 minute before the supernatant was removed with a pipette tip. To each Falcon tube 350 µl of the "Master mix" and also 36 µl DNA mix were added on top of the pellet. Yeast pellet, "Master mix" and DNA mix were vortexed for 10-20 seconds until the solution was cloudy. After the tubes were incubated in a water bath of 42 °C for 45 minutes they were centrifuged for 1 minute at 6000 rpm. The PEG solution was pipetted off and 500 µl synthetic complete drop-out medium (Y1376) were added. With the tip of a blue pipette the yeast pellet was gently resuspended and finally 500 µl of the resuspended mixture was plated on big Y1376 agar plates which were incubated at 28 °C for 2 days.

7.16.2. Yeast colony PCR

To ensure the success of yeast assembly, colony PCR was performed. Single yeast colonies were picked randomly with a sterile toothpick and transferred to a new Y1376 agar plate divided into 40 sections and incubated overnight at 30 °C. The following day 40 x 1.5 ml sterile Eppendorf tubes were prepared containing aliquots of 20 µl NaOH (20 mM). Grown colonies were picked with a yellow pipette tip and resuspended in NaOH. The tubes were incubated at 95 °C for 45 minutes and afterwards centrifuged at full speed for 10 minutes. For the performing of the PCR 2 µl of the supernatant was used as template DNA and the tubes were then stored at -20 °C. The first PCR was performed with one primer pair using the Hot Start Taq DNA Polymerase. Those clones which showed a band on the gel were tested also with the other two primer pairs to see if the Yeast Assembly was successful.

7.16.3. Isolation of plasmid DNA from Yeast

The success of DNA assembly in yeast was confirmed via analytical PCR with all three primersets. The positive clones of *S.cerevisiae* were picked with a sterile toothpick and inoculated in 5 ml Yeast Synthetic Drop- out medium. The incubation of the colonies was carried out overnight at 30 °C shaking at 200 rpm and the isolation was continued according to the following protocol:

2.5 ml of the overnight culture were centrifuged at 6000 rpm and the yeast pellet was resuspended in 200 µl ddH_2O. 50 µl of previously flicked Zymolyase® were added to the tube before the content was gently mixed. The cell suspension was incubated at 37 °C for 1 hour. After the time was up the cell suspension was centrifuged at 10000 rpm for 5 minutes and the supernatant was discarded. At this step components of the Wizard® Plus SV Minipreps

DNA Purification System Kit (PROMEGA) were used. 250 µl Cell Resuspension Solution were added to the pellet and resuspended. After the addition of 500 µl Cell Lysis Solution for resuspension, the tubes were inverted 4 times and incubated for 5 minutes at room temperature. 10 µl of Alkaline Protease Solution were added and the tubes were inverted again 4 times before incubating for another 5 minutes at room temperature. After adding 700 µl of Neutralization Solution the tubes were immediately inverted 4 times and centrifuged at 13600 rpm for 10 minutes at room temperature. The clear lysate was transferred to a spin column by decanting and a centrifugation was performed at full speed for 1 minute. The spin column was washed two times with 750 µl Column Wash Solution by spinning the tubes for 2 minutes at maximum speed. Finally, the plasmid DNA was eluted in 60 µl Nuclease-free water, which was previously preheated at 50 °C, by incubating at 37 °C for 15 minutes. The tubes were centrifuged for 2 minutes and the isolated plasmid DNA was stored at -20 °C.

7.16.4. Transformation of pDNA isolated from yeast per electroporation into *E.coli*

After the yeast assembly the resulting plasmids, namely pSS_3 and pSS_4 were available in low copy, whereby for further work the high quantity of pDNA was required. Therefore it has to be transformed in EPI 300 *E.coli* to reproduce the plasmid. The transformation was performed according to the following protocol:

Before starting the electroporation the following compounds should be kept on ice: competent EPI300 *E.coli* cells, plasmid DNA, sterile LB medium filled up in a 50 ml falcon and 0.2 cm cuvettes. For getting a better efficiency all steps have to be carried out very quickly. 50 µl of competent cells were placed in a cuvette. After adding 2 µl of pDNA the content was mixed by gently pipetting up and down. The electroporation was carried out using a GenePulser II (BioRad) which was set to 200 Ω, 25 µF and 2.5 kV to receive an expected time constant of 4.5-4.9 ms. Immediately 900 µl ice-cold LB medium were added to the shocked cells, mixed and transferred in a 2 ml Eppendorf tube. The incubation was carried out at 37 °C for 1 hour while shaking at 200 rpm. After the time was up the tubes were centrifuged and around 750 µl were removed with a pipette. Finally, the rest was resuspended and spread on LA plates containing apramycin for the selection. The plates were incubated overnight at 37 °C.

7.17. Conjugative transfer of pDNA into the *Streptomyces* strains

Conjugation is one of the three mechanisms of gene transfer in bacteria. In this work, the mutant strains *Streptomyces albus* B4*, Streptomyces albus* wt and *Streptomyces coelicolor* M1154 were used. Before the conjugation, a transformation of the plasmid DNA into ET12567 *E.coli* cells was performed, which was then conjugated in the previously named *Streptomyces* strains. Four clones of *E.coli* ET12567/pUZ8002 transformants were plated densely on a LA petri dish containing Kanamycin, Chloramphenicol and Apramycin as selective markers. The plates were incubated overnight at 37 °C. The following day the best grown colonies were swiped from the plate with a sterile loop and put into 500 µl 2x YT. The mixture was gently resuspended with a blue pipette. 50 µl of the required *Streptomyces* spore suspensions was mixed with 350 µl of 2x YT and incubated in a heating block with 50 °C for 7 minutes. After the tubes were cooled down at room temperature, 100 µl of *E.coli* ET12567/pUZ8002 cell suspension were added to the tube of the previously heat shocked spores. The tubes were inverted gently one time and centrifuged for 1 minute at 5000 rpm. After removing 300 µl of the supernatant, the pellet was resuspended in the remaining liquid and spread on SFM agar plates containing 1 M $MgCl_2$. The incubation of the plates was carried out overnight at 28 °C. After 18-20 hours of growth the plates were coated evenly with 1 ml of the following mixture: 1 ml H_2O, 30 µl Nalidixic acid (NaI) and 15 µl Apramycin (Am). NaI was necessary to reduce the growth of *E.coli* and Am was added to select the correct transconjugants. The plates were dried for around 40 minutes and then again incubated at 28 °C for approximately 2 days. When the growth of the clones was visible, transconjugants were picked with a sterile toothpick and transferred onto SFM agar plates containing 30 µl NaI and 15 µl Am. After an incubation of 2-3 days or when colonies were sufficiently grown, transconjugants were picked with a sterile loop and plated densely onto a fresh SFM agar plate containing 15 µl Am. Finally, a spore suspension of the *Streptomyces* strains was made, as soon as they sporulated.

7.18. Fermentation, extraction and analytical determination of secondary metabolites

7.18.1. Fermentation

Before the fermentation process starts, precultures were produced for the recombinant *Streptomyces* strains and also for their wild types, which were used as a control. 15 ml TSB-medium mixed with 7.5 µl Am were used for *S.albus* and 15 ml YEME-medium mixed with

7.5 µl Am were used for *S.coelicolor*. As an additional control, flasks with only media and antibiotics and without any bacteria were prepared. The incubation of the flasks was carried out overnight at 28 °C shaking at 200 rpm. 2.5 ml of *Streptomyces* overnight culture were inoculated in 50 ml fermentation media listed in materials 6.1.6. The flasks were incubated by shaking at 28 °C for 7 days. Finally, the flasks were freeze-dried and extracted.

7.18.2. Extraction

50 ml of 100 % methanol were added to the freeze-dried culture from 7.18.1 and the methanol/pellet mixture was mixed for 2 h at room temperature at ca. 100 rpm. Methanol extracts were filtrated through paper filter and concentrated in rotary evaporator. Resulting pellet was solved in 5 mL 100 % methanol. The methanol extracts were stored at -20 °C and tested via HPLC and disk diffusion test.

7.18.3. Bioassays

Methanolic extracts were tested for antimicrobial activity via disk diffusion test. Nine microorganisms listed in table 15 were tested in order to verify whether the extracts contain potentional antimicrobial compounds. 15 µl of methanolic extract were applied onto a paper disk (6 mm Whatman filter paper discs) and dried for 20 min at RT. In the meantime, 150 µl cell suspension of different strains were plated on an appropriate agar plate and spread with a sterile spatula. When the plates and the paper disks were dry, the disks were carefully placed on the agar plates. Depending on the ideal growth temperature, the plates were placed into an incubator overnight and next day the growth was observed.

Table 15: Test organisms and their optimal growth temperature and medium used for bioactivity testing.

Strain	**Type of Microorganism**	**Temperature**	**Medium**
Escherichia coli DH5αF	Gram-negative bacteria	37 °C	LA
Pseudomonas putida KT 2440	Gram-negative bacteria	28 °C	LA
Erwinia persicina DSMZ 19328	Gram-negative bacteria	28 °C	LA
Bacillus subtilis DSMZ 10	Gram-positive bacteria	28 °C	LA
Kocuria rhizophila DSMZ 348	Gram-positive bacteria	28 °C	LA
Micrococcus luteus DSMZ 1790	Gram-positive bacteria	28 °C	LA
Enterococcus mundtii DSMZ 4840	Gram-positive bacteria	37 °C	LA
Staphylococcus carnosus DSMZ 20501	Gram-positive bacteria	37 °C	LA
Saccharomyces cerevisiae	Fungi	28 °C	YPD

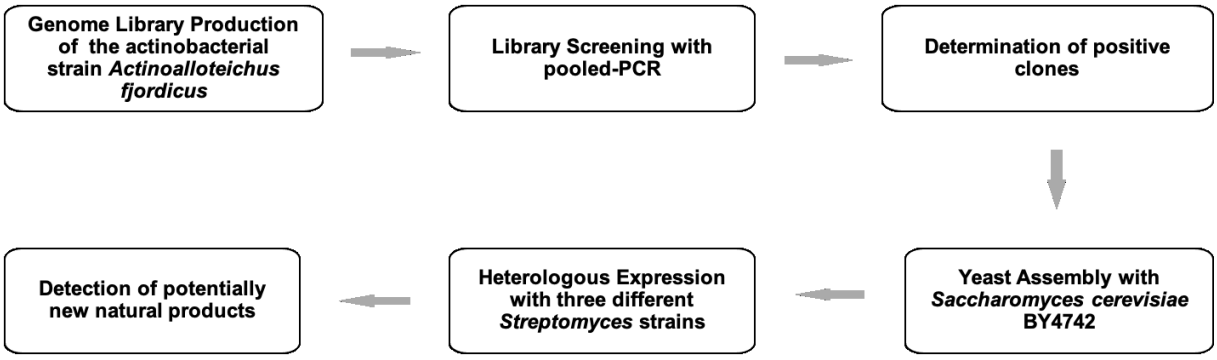


Figure 11: Schematic overview of the workflow of this work. Starting with the production of a genome library of the rare actinobacterial strain *Actinoalloteichus fjordicus,* followed by a pooled-PCR screening of the library to identify positive *E.coli* clones, which carry BGC3 and BGC4. After yeast assembly in *Saccharomyces cerevisiae* BY4742, heterologous expression in three different *Streptomyces* strains was attempted. Aim of this project was to produce potentially new NPs.

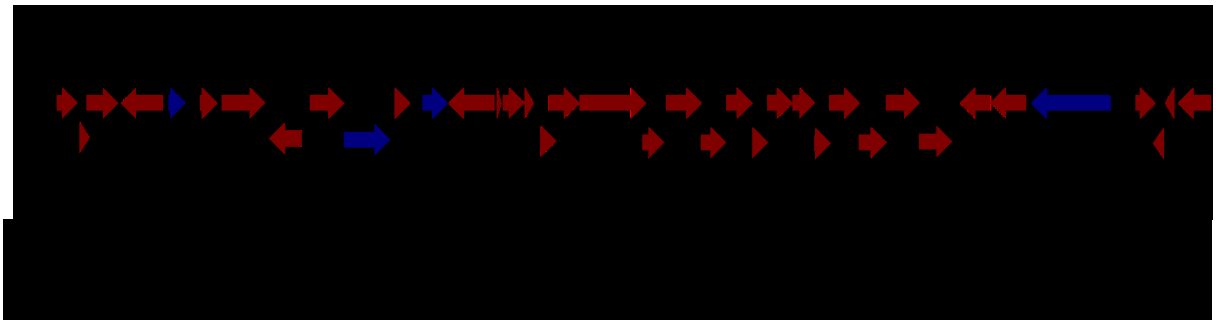
8. Results

8.1. Genome mining for BGC3 and BGC4

In silico analysis of genome sequencing data showed that *Actinoalloteichus fjordicus* DSM 46856 carries 24 BGCs. Two BGCs, namely BGC3 and BGC4, which might potentially specify biosynthesis of new natural products were chosen as target for this project. With bioinformatic online tools antiSMASH and BLAST the prediction and analyzes of these BGCs was performed.

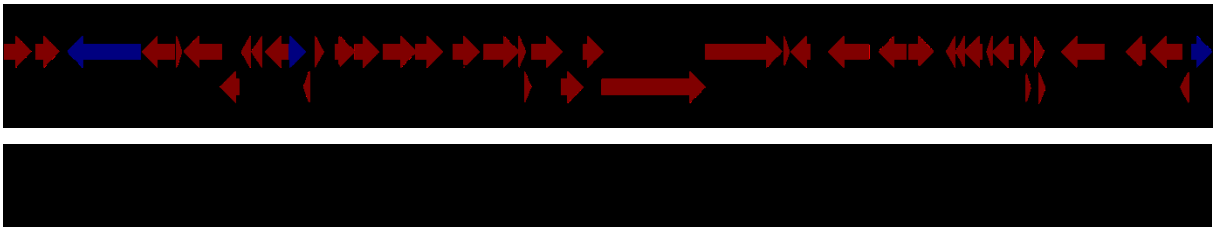
BGC3 was predicted to encode for ladderane, which is an unique lipid structure in bacterials membrane [53], while BGC4 should specify a NRP-T1PK hybrid compound. For BGC3 and BGC4 there are no clusters with significant degree of similarity, which suggests that BGC4 might specify novel NPs.

Fig. 13 and Fig. 14 show the genetic compositions of BGC3 and BGC4. BGC3 has a size of 42328 bp, encompassing 38 genes including four regulatory genes. BGC4 (49899 bp) contains 45 genes of which three encode transcriptional regulators.



BGC3 (42328 bp)

Figure 13: Schematic overview of the gene composition of BGC3 from *Actinoalloteichus fjordicus*. 38 genes of BGC3 are shown as arrows. Four genes (5,10,12,34) are colored in blue and represent regulatory genes. The predicted functions of the genes and their homology are listed in Table 16.



BGC4 (49899 bp)

Figure 14: Schematic overview of the gene composition of BGC4 from *Actinoalloteichus fjordicus*. 45 genes of BGC4 are shown as arrows. Three genes (3,11,45) are colored in blue and represent regulatory genes. The predicted functions of the genes and their homology are listed in Table 17.

Tables 16 and 17 show the genes of BGC3 and BGC4 detected via antiSMASH. One gene (*orf* 19) from BGC3 encodes beta-ketoacyl synthase family protein, which is typical for ladderane class of NPs. For BGC4 seven genes (*orf* 19,20,21,22,23,24,25) encode the predicted NRPS and one gene encodes type I PKS (*orf* 26).

Table 16: List of 38 genes of BGC3 from *Actinoalloteichus fjordicus* detected via antiSMASH. Gene encoding beta-ketoacyl synthase is marked in grey.

	Gene number	**Homolog protein**	**Protein source**	**Percent identity**
orf 1	MP12924_1173	YbaB/Ebfc family nucleotid-associated protein	*Actinoalloteichus hoggarensis*	70.64%
orf 2	MP12924_1174	WXG100 family type VII secretion target	*Actinoalloteichus hoggarensis*	86.24%
orf 3	MP12924_1175	Hypothetical protein	*Actinoalloteichus hoggarensis*	91.18%
orf 4	MP12924_1176	Hypothetical protein	*Actinoalloteichus hoggarensis*	83.44%
orf 5	MP12924_1177	TetR family transcriptional regulator	*Actinoalloteichus hoggarensis*	77.78%
orf 6	MP12924_1178	DTDP-4-dehydrorhamnose 3,5-epimerase family protein	*Actinoalloteichus hoggarensis*	85.07%
orf 7	MP12924_1179	NDP-hexose 2,3-degydratase family protein	*Actinoalloteichus hoggarensis*	85.11%
orf 8	MP12924_1180	DegT/DnrJ/EryC1/StrS family aminotransferase	*Actinoalloteichus hoggarensis*	94.04%
orf 9	MP12924_1181	Glycosyl transferase	*Actinoalloteichus hymeniacidonis*	89.17%
orf 10	MP12924_1182	AfsR/SARP family transcriptional regulator	*Actinoalloteichus hymeniacidonis*	84.16%
orf 11	MP12924_1183	NUDIX hydrolase	*Nocardia sp. NRRL S-836*	59.26%
orf 12	MP12924_1184	5´-deoxynucleotidase YfbR-like HD superfamily hydrolase	*Saccharopolyspora spinosa*	90.38%
orf 13	MP12924_1185	FAD-dependent monooxygenase	*Actinoalloteichus hoggarensis*	91.06%
orf 14	MP12924_1186	Hypothetical protein UA74_05705	*Actinoalloteichus fjordicus*	100.00%
orf 15	MP12924_1187	3-oxoacyl-ACP reductase	*Actinoalloteichus*	97.59%

		FabG	*hymeniacidonis*	
orf 16	MP12924_1188	Acyl carrier protein	*Actinoalloteichus hoggarensis*	98.00%
orf 17	MP12924_1189	3-hydroxyl-ACP dehydratase FabZ	*Actinoalloteichus hymeniacidonis*	91.21%
orf 18	MP12924_1190	Glycine cleavage system protein	*Actinoalloteichus hoggarensis*	94.07%
orf 19	MP12924_1191	Beta-ketoacyl synthase family protein	*Actinoalloteichus hoggarensis*	91.67%
orf 20	MP12924_1192	Nitroreductase family protein	*Actinoalloteichus hoggarensis*	95.87%
orf 21	MP12924_1193	NAD(P)/FAD-dependent oxidoreductase	*Actinoalloteichus hymeniacidonis*	95.16%
orf 22	MP12924_1194	Alpha/beta hydrolase	*Actinoalloteichus hoggarensis*	93.38%
orf 23	MP12924_1195	ACP S-malonyltransferase	*Actinoalloteichus hymeniacidonis*	84.33%
orf 24	MP12924_1196	Hypothetical protein	*Actinoalloteichus hymeniacidonis*	70.72%
orf 25	MP12924_1197	Ester cyclase	*Actinoalloteichus hoggarensis*	87.06%
orf 26	MP12924_1198	Alpha/beta hydrolase	*Actinoalloteichus hoggarensis*	92.31%
orf 27	MP12924_1199	4´-phosphopantetheinyl transferase superfamily protein	*Actinoalloteichus hoggarensis*	71.88%
orf 28	MP12924_1200	Right-handed parallel beta-helix repeat-containing protein	*Actinoalloteichus hoggarensis*	93.08%
orf 29	MP12924_1201	ABC transporter ATP-binding protein	*Actinoalloteichus hoggarensis*	97.44%
orf 30	MP12924_1202	ABC transporter permease	*Actinoalloteichus hoggarensis*	97.19%
orf 31	MP12924_1203	ABC transporter permease	*Actinoalloteichus hoggarensis*	90.55%
orf 32	MP12924_1204	NAD(P)-dependent oxidoreductase	*Actinoalloteichus hoggarensis*	79.40%
orf 33	MP12924_1205	Class I SAM-dependent methyltransferase	*Actinoalloteichus hoggarensis*	91.18%
orf 34	MP12924_1206	LuxR family transcriptional regulator	*Actinoalloteichus hoggarensis*	86.26%
orf 35	MP12924_1207	16s rRNA (adenine(1408)-N(1))-Methyltransferase KamB	*Actinoalloteichus hymeniacidonis*	87.96%
orf 36	MP12924_1208	Nucleotidyltransferase domain-containing protein	*Actinophtocola sp.*	68.97%
orf 37	MP12924_1209	Nucleotidyltransferase domain-containing protein	*Phycicoccus sp.*	72.50%
orf 38	MP12924_1210	DUF371 domain-containing protein	*Actinoalloteichus hoggarensis*	83.20%

Table 17: List of 45 genes of BGC4 from *Actinoalloteichus fjordicus* detected via antiSMASH, including seven genes encoding NRPS (marked in grey) and one gene encoding T1PKS (marked in darkgrey).

	Gene number	**Homolog protein**	**Protein source**	**Percent identity**
orf 1	MP12924_1318	Putative methyltransferase YcgJ	*Actinoalloteichus hoggarensis*	88.95%
orf 2	MP12924_1319	Acetyltransferase-like isoleucine patch superfamily enzyme	*Actinoalloteichus hymeniacidonis*	80.97%
orf 3	MP12924_1320	Tetratricopeptide repeat protein	*Actinoalloteichus hoggarensis*	87.41%
orf 4	MP12924_1321	Glycosyltransferase family 4 protein	*Actinoalloteichus hoggarensis*	86.76%
orf 5	MP12924_1322	Hypothetical protein	*Actinoalloteichus hymeniacidonis*	64.18%
orf 6	MP12924_1323	1,4-alpha-glucan branching enzyme	*Actinoalloteichus hoggarensis*	90.89%
orf 7	MP12924_1324	Class I SAM-dependent methyltransferase	*Actinoalloteichus hoggarensis*	86.80%
orf 8	MP12924_1325	ParA family protein	*Herbidospora cretacea*	55.00%
orf 9	MP12924_1326	PPOX class F420-dependent oxidoreductase	*Actinoalloteichus hoggarensis*	76.64%
orf 10	MP12924_1327	Oxidoreductase	*Actinoalloteichus hoggarensis*	70.69%
orf 11	MP12924_1328	TetR/AcrR family transcriptional regulator	*Actinoalloteichus hoggarensis*	87.27%
orf 12	MP12924_1329	Hypothetical protein UA74_96425	*Actinoalloteichus fjordicus*	100.00%
orf 13	MP12924_1330	Hypothetical protein	*Actinoalloteichus hoggarensis*	95.50%
orf 14	MP12924_1331	Electron transfer flavoprotein subunit beta/FixA family protein	*Actinoalloteichus hoggarensis*	96.15%
orf 15	MP12924_1332	Electron transfer flavoprotein subunit alpha/FixB family protein	*Actinoalloteichus hoggarensis*	90.91%
orf 16	MP12924_1333	Cysteine desulfurase	*Actinoalloteichus hoggarensis*	83.73%
orf 17	MP12924_1334	TRNA-specific 2-thiouridylase MnmA	*Actinoalloteichus hoggarensis*	95.89%
orf 18	MP12924_1335	SAM-dependent methyltransferase	*Actinoalloteichus hoggarensis*	85.84%
orf 19	MP12924_1336	AMP-binding protein	*Actinoalloteichus hoggarensis*	89.41%
orf 20	MP12924_1337	Acyl carrier protein	*Actinoalloteichus hoggarensis*	82.76%
orf 21	MP12924_1338	Hypothetical protein AHOG_06240	*Actinoalloteichus hoggarensis*	87.21%
orf 22	MP12924_1339	Beta-ketoacyl-synthase family protein	*Actinoalloteichus hoggarensis*	77.80%
orf 23	MP12924_1340	Ketosynthase	*Actinoalloteichus hoggarensis*	68.44%
orf 24	MP12924_1341	Hypothetical protein AHOG_06255	*Actinoalloteichus hoggarensis*	67.31%

orf 25	MP12924_1342	Non-ribosomal peptide synthetase	*Actinoalloteichus hoggarensis*	69.43%
orf 26	MP12924_1343	Type I polyketide synthase	*Actinoalloteichus hoggarensis*	83.21%
orf 27	MP12924_1344	MbtH family NRPS accessory protein	*Actinoalloteichus hoggarensis*	86.89%
orf 28	MP12924_1345	Linear gramicidin dehydrogenase LgrE	*Actinoalloteichus hoggarensis*	91.02%
orf 29	MP12924_1346	Substrate-binding domain-containing protein	*Actinoalloteichus hymeniacidonis*	75.10%
orf 30	MP12924_1347	Putative oxidoreductase CzcO	*Actinoalloteichus hoggarensis*	84.07%
orf 31	MP12924_1348	Phosphotransferase	*Actinoalloteichus hymeniacidonis*	79.50%
orf 32	MP12924_1349	Hypothetical protein	*Actinoalloteichus hoggarensis*	54.46%
orf 33	MP12924_1350	Hypothetical protein	*Actinoalloteichus hymeniacidonis*	64.86%
orf 34	MP12924_1351	Hypothetical protein	*Actinoalloteichus hoggarensis*	53.72%
orf 35	MP12924_1352	DUF397 domain-containing protein	*Actinoalloteichus hymeniacidonis*	67.74%
orf 36	MP12924_1353	Helix-turn-helix domain-containing protein	*Streptoalloteichus hindustanus*	60.64%
orf 37	MP12924_1354	Hypothetical protein TL08_15110	*Actinoalloteichus hymeniacidonis*	54.72%
orf 38	MP12924_1355	MULTISPECIES: Hypothetical protein	*Actinoalloteichus*	100.00%
orf 39	MP12924_1356	Hypothetical protein	*Actinoalloteichus hoggarensis*	61.11%
orf 40	MP12924_1357	Hypothetical protein	*Actinoalloteichus hoggarensis*	73.08%
orf 41	MP12924_1358	Putative ABC transporter ATP-binding protein YjjK	*Actinoalloteichus hoggarensis*	88.89%
orf 42	MP12924_1359	Glucose 1-dehydrogenase	*Herbihabitans rhizosphaerae*	72.73%
orf 43	MP12924_1360	FAD-dependent oxidoreductase	*Actinoalloteichus hoggarensis*	86.85%
orf 44	MP12924_1361	Bifunctional 3-phenylpropionate/cinnamic acid dioxygenase ferredoxin subunit	*Actinoalloteichus hoggarensis*	95.19%
orf 45	MP12924_1362	IclR family transcriptional regulator	*Actinoalloteichus hoggarensis*	97.40%

8.2. Genome library construction

During this project a genome library of *Actinoalloteichus fjordicus* was created to cover the full length of the actinobacterial genome. For the production of a complete and unbiased fosmid library, DNA-fragments of approximately 40 kb were produced by shearing the genomic DNA (gDNA) 160 times using procedure described in Methods. The concentration of the DNA was measured with Nanodrop in order to know how much water had to be added

for the needed concentration for further working steps. The successful shearing was confirmed with gel electrophoresis, which produced a typical “smear”, as shown in Figure 15A.

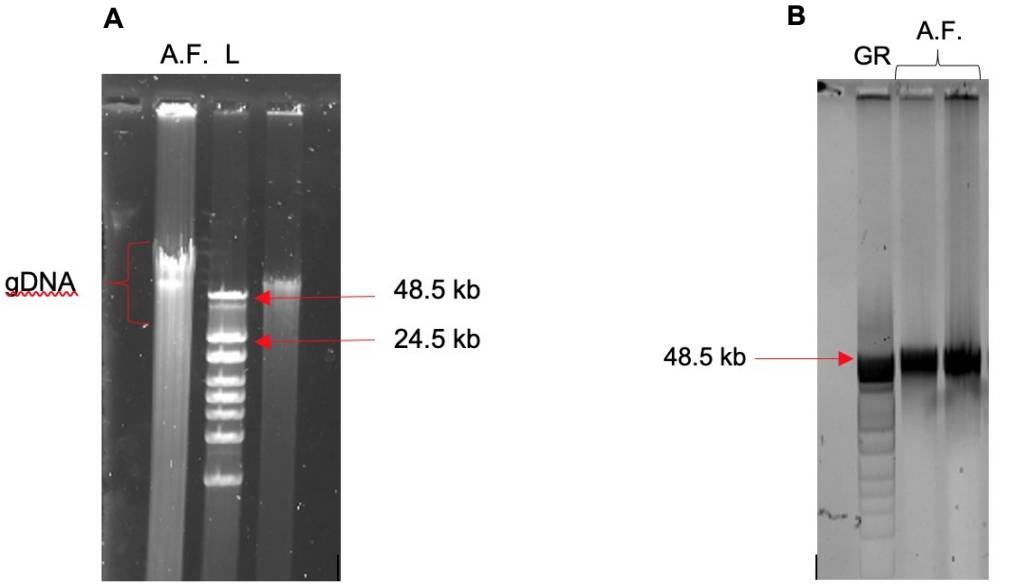


Figure 15: (A) Gel image of the end-repaired insert DNA (A.F.), which was cut out of the gel at a height of around 40 kb. (B) Gel image after the DNA was cut out and recovered to prove the presence of the DNA. 1 kb Ladder (L) and GeneRuler High Range DNA Ladder (GR) were used as marker.

The DNA was cut out of the gel, recovered and tested with gel electrophoresis, which can be seen in Figure 15B. After the ligation with CopyControl pCC1FOS vector, the ligation mixture was packaged in phage lambda particles, followed by infection of *E.coli* EPI300-T1R. The infection worked well, as CopyControl fosmid-containing clones grew overnight. Single colonies were counted and the titer of the packaged phage particles calculated. The working process of the genome library production is demonstrated as schematic overview in Fig. 16.
To obtain 10x coverage of the full-length genome of *Actinoalloteichus fjordicus*, at least 1700 clones were needed. Therefore, 18 96-well-plates were labeled and prepared which harbor a total of 1728 clones.

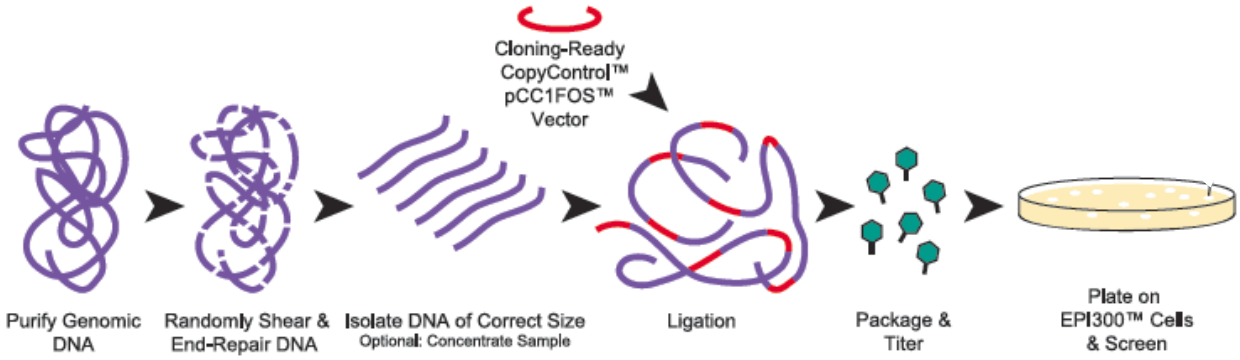


Figure 16: Schematic overview of the construction process of a CopyControl™ Fosmid Library. [54]

8.3. Primer design and library screening

The construction of the genome library provided the possibility to discover new natural products. In terms of finding *E.coli* clones which carry parts of the BGCs of interest, BGC3 and BGC4, the genome library of *Actinoalloteichus fjordicus* was screened via pooled-PCR. First of all, the functionality of the ordered primer pairs was tested with gel electrophoresis as shown in Fig. 17, to prove the expected size of the PCR products. The native gDNA of *Actinoalloteichus fjordicus* was used as template DNA. All PCR reactions of the library screening were performed using Hot Start *Taq* DNA Polymerase.

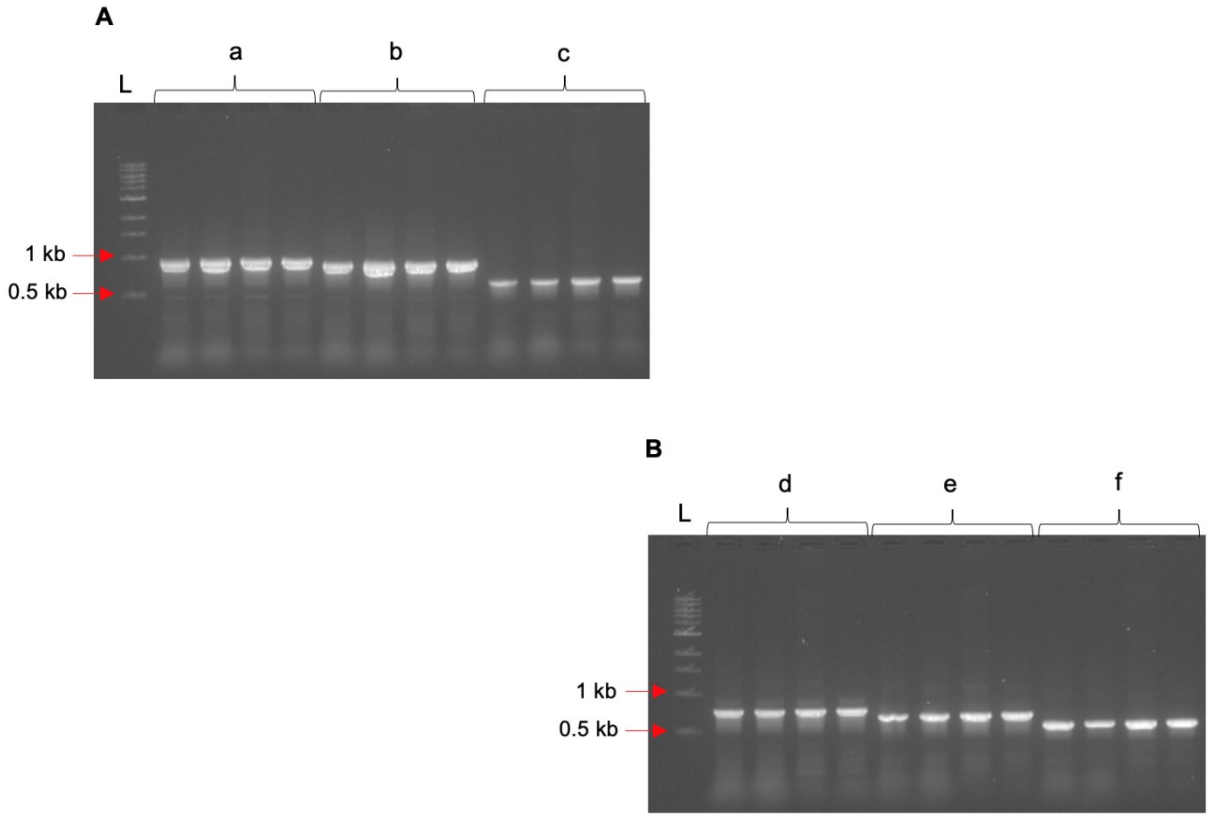


Figure 17: (A) Gel image to prove the functionality of the ordered primer pairs a, b and c for BGC3. PCR products showed the expected signals like following: primerset a: 949 bp, primerset b: 896 bp and primerset c: 633 bp. (B) In case of BGC4 primer pairs d, e and f were tested to verify the expected size of PCR products, which showed the signals like following: primerset d: 795 bp, primerset e: 697 bp and primerset e: 584 bp. 1 kb Ladder (L) was used as a marker.

To find *E.coli* clones containing BGC3 and BGC4 (or parts thereof), all 1728 clones of the 18 96-well-plates were screened. At first, the plates were divided into two halves, resulting in the left (L) and right (R) part, each containing 48 *E.coli* clones. A L- and R-mixes of all 18 plates was prepared.

In terms of BGC3, plate 11R showed a band with primerset a and b and plate 13L showed one with primersets b and c. The plate was separated into smaller pools until the *E.coli* clones were tested individually. Finally, clones 11A7 and 13H6 showed a positive band on

the gel. Positive clones of pooled-PCR were verified via sequencing. In case of BGC3, results demonstrated that the whole cluster was splitted between those two *E.coli* clones.
Regarding BGC4, the plates 5L and 1L yielded positive PCR results. Subsequently, after continuing with testing the different *E.coli* clones individually, 5H3 showed a band on the gel with primersets d and e, and 1H1 had a positive signal with primerset f. As the sequencing results demonstrated, those two different *E.coli* clones did not make up the complete cluster, therefore the missing piece had to be produced via PCR.
Figure 18 presents a 96-well-plate, which was separated into part L and part R, showing the position of the positive *E.coli* clones for BGC3 and BGC4.

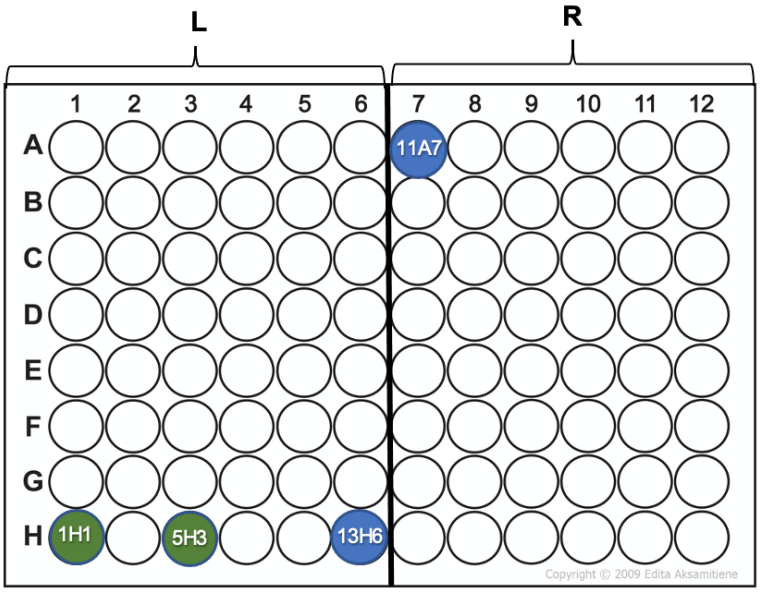


Figure 18: 96-well-plate separated in part L and R. Two fosmids with the blue background belong to BGC3, while the other two *E.coli* clones with the green background carry genes of BGC4.

Table 18: Overview of fosmids from BGC3 and BGC4, which were used for yeast assembly.

BGCs	Fosmid	Plate	Clone	Primerset
BGC3	11A7_pCC1FOS	11R	11A7	a and b
	13H6_pCC1FOS	13L	13H6	b and c
BGC4	pFOS1_5H3	5L	5H3	d and e
	pFOS1_1H1	1L	1H1	f

The fDNA of the *E.coli* clones shown in Table 18 was isolated following the protocol in Methods 7.10. Subsequently, the isolated DNA was sequenced by Eurofins Genomics to identify whether all genes of the BGCs of interest were present in the respective clones, or some genes were still missing and have to be amplified via PCR.

8.4. Preparation of fragments from BGC3 and BGC4 for Yeast Assembly

Sequencing results provided information about the fosmid sizes and the gene composition of BGC3 and BGC4, as well as about the location of endonucleases restriction sites, which could be used for cuting out the parts of fDNA of interest.

8.4.1. Fragments for BGC 3

The complete BGC3 with its 38 genes was carried by two different *E.coli* clones, which overlapped in genes 15 and 16 and contain some other genes, which are out of the cluster. After the fDNA was isolated, a restriction digestion of the fosmids was performed.

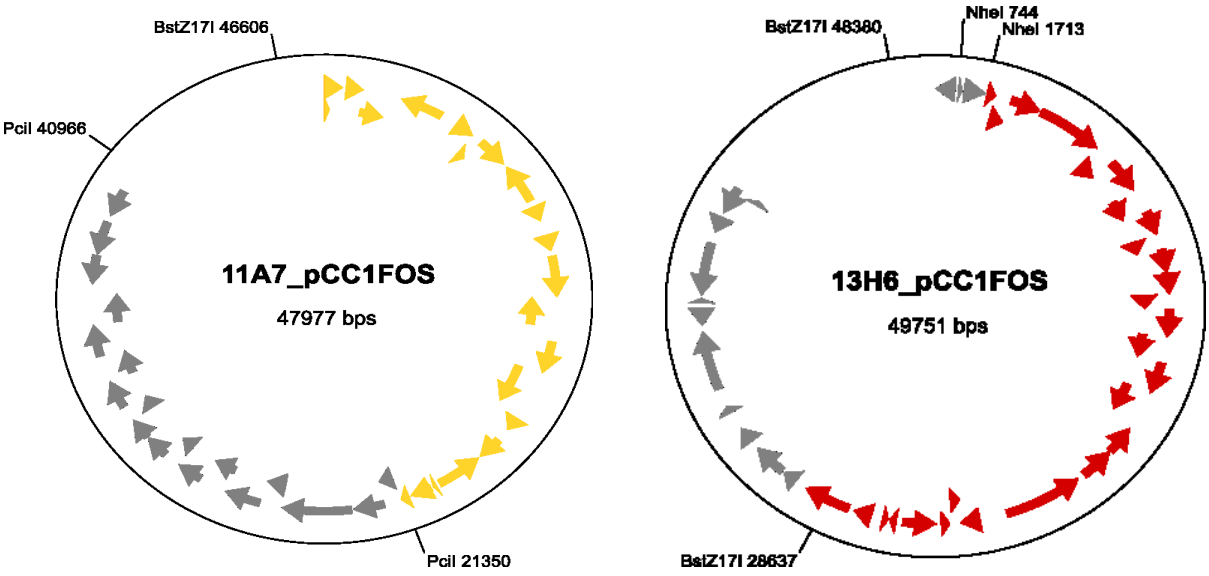


Figure 19: Schematic overview of the genes of 11A7_pCC1FOS and 13H6_pCC1FOS, which will be assembled after restriction digestion in yeast. Yellow and red colored arrows result together in the complete cluster 3, while the grey arrows are out of BGC3. The fosmid on the left side was cut with *Bst*Z17I, *Pci*I and *Nde*I and resulted in six fragments: 17471 bp, 1396 bp, 749 bp, 2479 bp, 3163 bp and 22719 bp. The latter fragment was the necessary one, which was used later for yeast assembly. The second fosmid was restricted with *Bst*Z17I and *Nhe*I and was predicted to show 4 fragments: 969 bp, 2113 bp, 19743 bp and 26926 bp. The fragment of 13H6 with a size of 26926 bp was also used for yeast assembly.

Clone 11A7 was cut with the enzymes *Bst*Z17I, *Pci*I and *Nde*I and according to the prediction of Clone Manager six fragments should result with the sizes of 17471 bp, 1396 bp, 749 bp, 2479 bp, 3163 bp and 22719 bp. The restriction of 13H6 was carried out with *Bst*Z17I and *Nhe*I and 4 bands with the sizes of 969 bp, 2113 bp, 19743 bp and 26926 bp were predicted to appear on the gel. Subsequently, the cut fosmids were applied on the agarose gel as showed in Fig. 20. Regarding clone 11A7 one band did not appear, while clone 13H6 showed one additional band that was not expected. This could suggest that one restriction enzyme did not work properly. In case of clone 11A7, the fragment with the size of 22719 bp

and from the second fosmid 13H6 the fragment with the size of 26926 bp were used for yeast assembly.

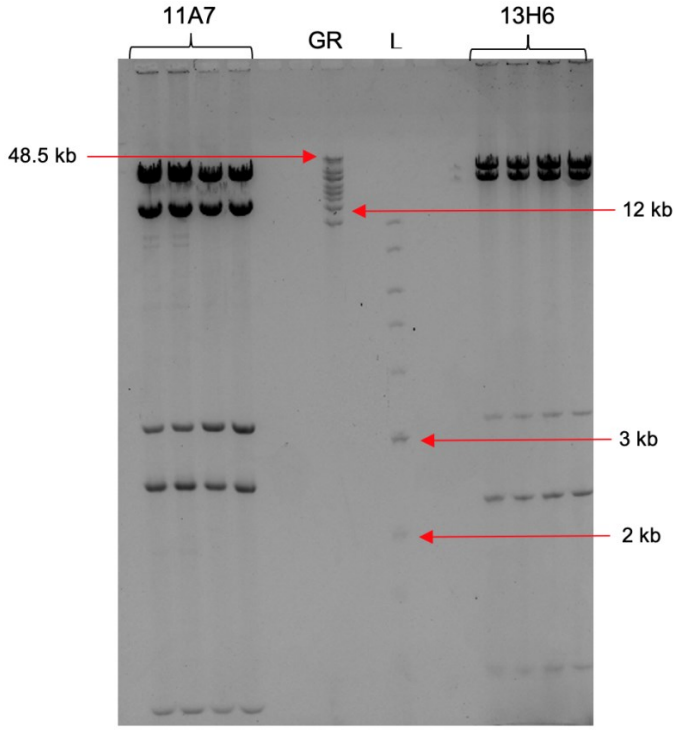


Figure 20: Gel image of clone 11A7 and 13H6 after restriction digestion with the required enzymes listed in Table 12. The needed fragment from 11A7 has a size of 23 kb, while the needed fragment from 13H6 has a size of 27 kb. Those two bands were cut out, recovered and later used for yeast assembly.

8.4.2. Fragments for BGC4

For BGC4 two different *E.coli* clones, which showed a positive band on the gel were picked for sequencing. As sequencing results showed, those two clones, 5H3 and 1H1 did not contain all the genes of BGC4. Therefore, missing pieces had to be amplified via PCR with ordered primers. As shown in Fig. 21 and Fig. 22, gene 1 and 2 (5H3_PCR) and gene 37, 38, 39, and 40 (1H1_PCR) were amplified and then used in the yeast assembly.

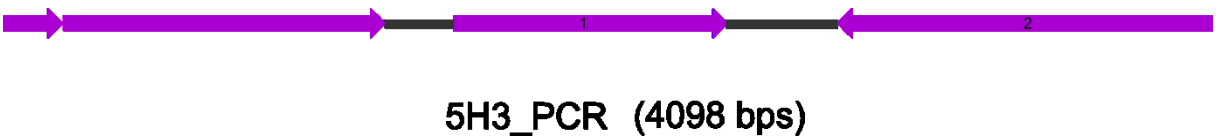


Figure 21: PCR product of missing pieces of cluster 4 with primerset 5H3_fwd/rev, which will be used in yeast assembly.

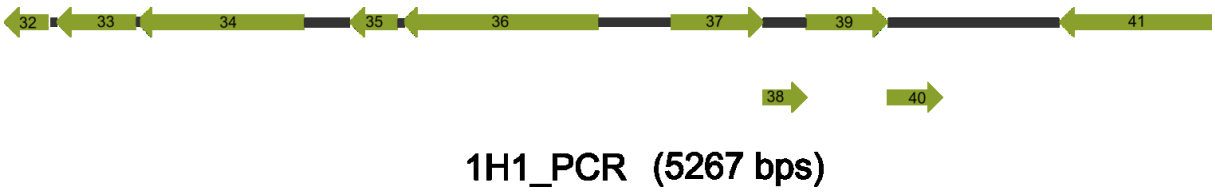


Figure 22: PCR product of missing pieces of cluster 4 with primerset 1H1_fwd/rev, which will be used in yeast assembly.

The fDNA of the sequenced clones was isolated and then digested with respective endonucleases. pFOS1_5H3 was cut with the enzymes *Hind*III and *Fse*I. The software Clone Manager predicted 3 fragments with the sizes of 37110 bp, 7410 bp and 3705 bp, while pFOS_1H1 was digested with *Pci*I and *Avr*II, yielding two fragments with the sizes of 11706 bp and 39278 bp.

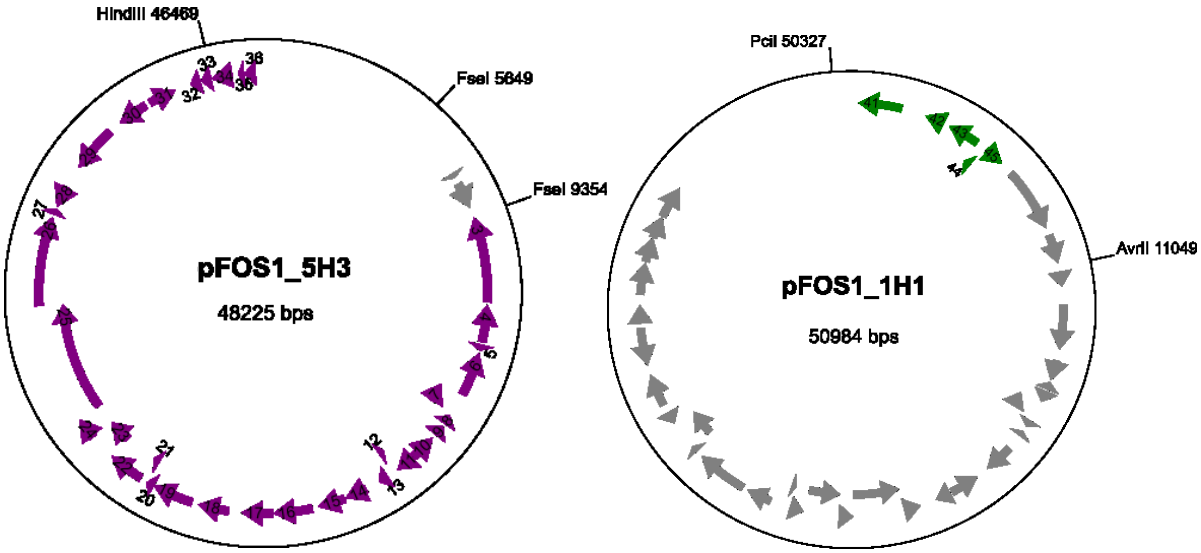


Figure 23: Schematic overview of the genes of pFOS1_5H3 and pFOS1_1H1, which will be assembled in yeast after performing a restriction digestion. Purple and green colored arrows represent genes in cluster 4, while the grey arrows are out of BGC4. The fosmid on the left side was cut with *Hind*III and *Fse*I and resulted in three fragments of 37110 bp, 7410 bp and 3705 bp. The biggest fragment with 37 kb was used later for yeast assembly. The second fosmid was restricted with *PciI* and *Avr*II and resulted in two products of 11706 bp and 39278 bp. The smaller one was used for yeast assembly.

As shown in Fig. 24A, the digested fosmid 5H3 was applied on the agarose gel and the largest band (37 kb) was isolated from the gel and used in yeast assembly. The restriction digestion of 1H1 needed more steps before it could be used for the assembly, because the enzymes did not cut with the same buffer. First, the fosmid was cut with *Avr*II, subsequently it was applied on the gel. After isolation, the DNA was cut with *Pci*I and again applied on the gel as shown in Fig. 24B. Finally, the smaller fragment (12 kb) was recovered and used as DNA-fragment for the assembly in yeast.

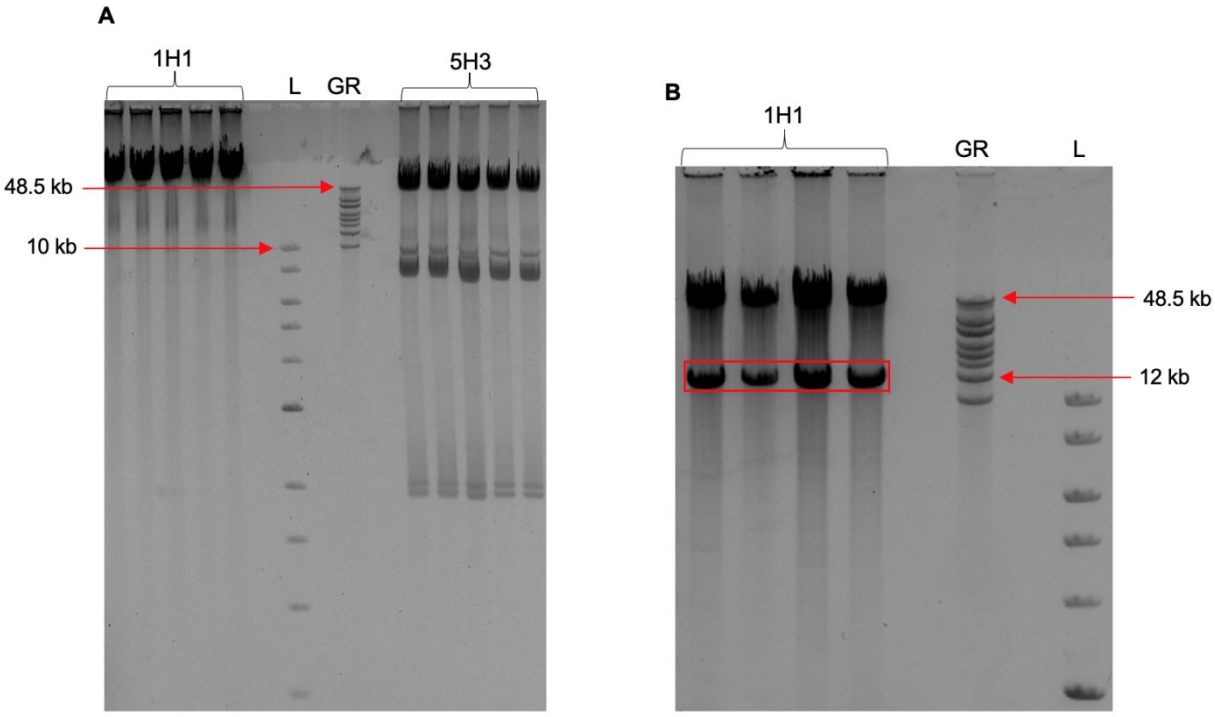


Figure 24: (A) Gel image of clone 1H1 and 5H3 after restriction digestion with the required enzymes listed in Table 12. In case of 5H3 the band at the top with a size of 37 kb was cut out of the gel, recovered and ready for the transformation in yeast. Regarding 1H1 a few more steps were needed before it was used for the assembly, because the enzymes did not cut with the same buffer. Clone 1H1 was first cut with *Avr*II, then applied on the gel. (B) After isolating the DNA, a restriction digestion with *Pci*I was performed and again applied on the gel. The resulting needed DNA-fragment with a size of 12 kb was cut out, recovered and ready for yeast assembly. 1 kb Ladder (L) and GeneRuler High Range DNA Ladder (GR) were used as reference.

8.5. Modification of the vector pCLY10

As soon as the fragments were ready for being assembled in yeast, the shuttle vector pCLY10 was chosen for cloning BGC3 and BGC4 using *S.cerevisiae* as assembly host. This vector had to be modified to carry “capture arms” that are homologous to the cluster’s flanking regions. The resulting “capture vectors” had to be linearized with an endonuclease to enable homologous recombination with the DNA-fragments 11A7_pCC1FOS and 13H6_pCC1FOS of BGC3 and pFOS1_5H3 and pFOS1_1H1 of BGC4. The process of yeast-based DNA assembly is described in chapter 4.6.2.

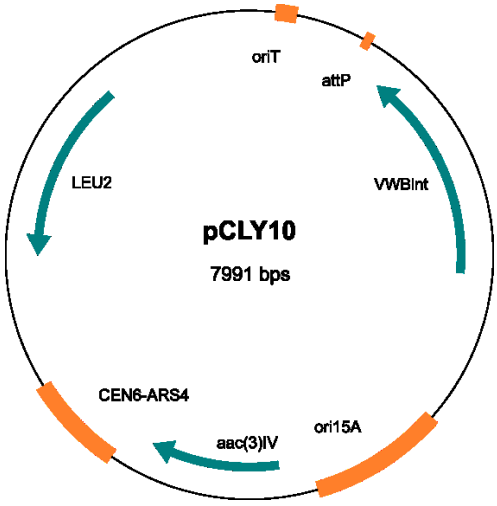


Figure 25: Genetic map of the vector pCLY10, containing the 7 elements: *oriT,* an origin of transfer; *attP,* the site for intergration into the *Streptomyces*-genome; *VWBint,* phage integrase; *ori15A,* origin of plasmid replication in *E. coli; aac(3)IV,* apramycin resistance gene; *CEN6-ARS4,* origin of replication in yeast and *LEU2,* yeast marker.

In Fig. 25, the genetic map of the shuttle vector pCLY10 is shown. It has a size of 7991 bp and contains the following seven elements. *LEU2* is a yeast selectable marker, to select recombinant clones in yeast. The element *oriT*, origin of transfer is responsible for the plasmid conjugation into an acceptor cell. To enable the replication in yeast, *CEN6-ARS4*, a yeast centromere/replication origin is necessary. *aac(3)IV* is an apramycin resistance marker for selection in *E.coli* and *Streptomyces.* For replication of a vector in *E.coli* the element, *ori15A* is needed. *VWBint*, encoding an integrase and *attP*, phage attachment site, are important for the integration of a vector into *Streptomyces* genome. [39]

8.5.1. Modification of pCLY10 for assembly of BGC3

The insert, C3_LR_flank which was later ligated into the shuttle vector pCLY10 was consisting a part of *orf 1* (“LF 537 bp) and a part of *orf 38* (RF” 529 bp). This fragment was synthesized by BioCat. To facilitate the ligation with pCLY10, the product was amplified via PCR in order to add the restriction sites for enzymes *Hind*III and *Not*I. The PCR was performed with Q5 High Fidelty 2X Master Mix with the primerset listed in Table 7, Materials 6.3.3. A gel electrophoresis of C3_LR_flank was carried out, as shown in Fig. 26A and subsequently the DNA was recovered from the gel.

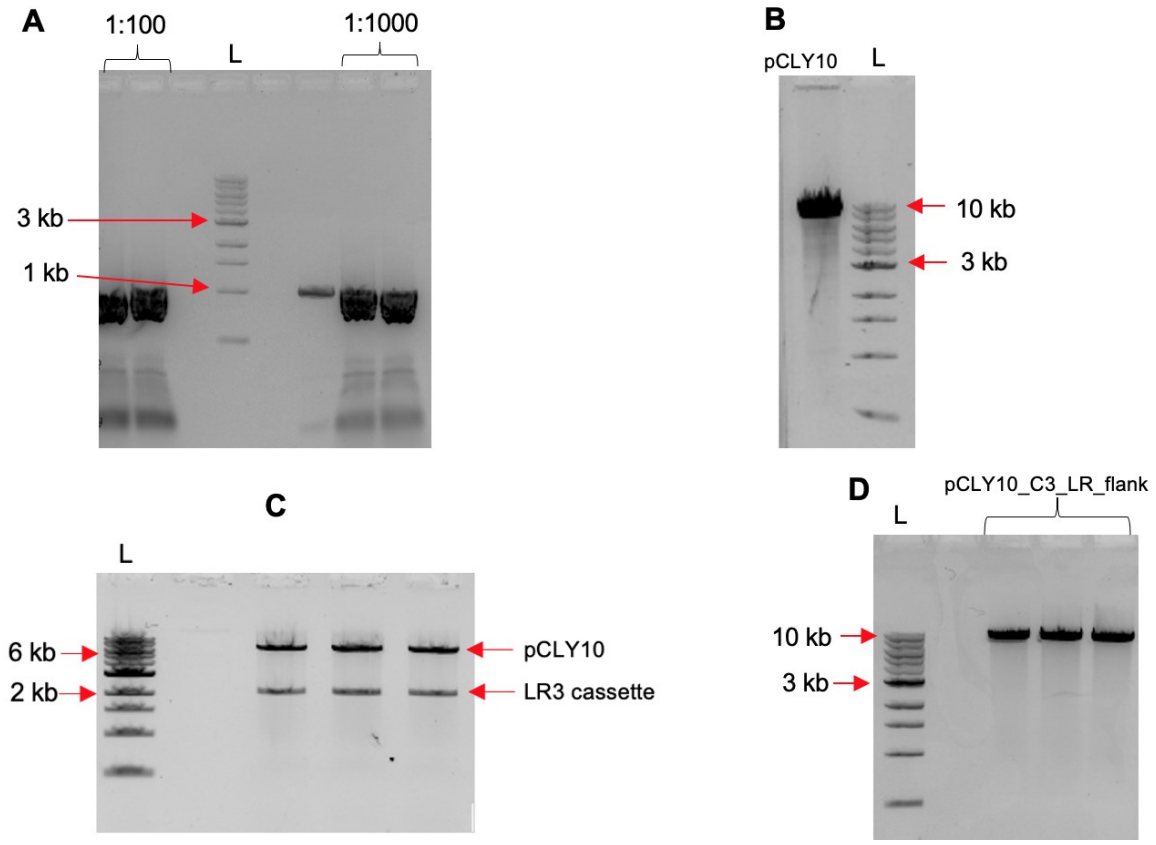


Figure 26: (A) Amplified C3_LR_flank via PCR with primerset SS_LR3_fwd/rev and Q5 High-Fidelity 2X Master Mix. Two dilutions of the inserts were prepared: 1:100 and 1:1000. (B) Restriction of pCLY10 with *Hind*III and *Not*I. (C) Analytical restriction of pCLY10_C3_LR_flank with X*ho*I and *Pme*I to verify the success of the ligation. The band at the top with a size of 6894 bp is representing pCLY10, while the band at the bottom with a size of 2096 bp represents the C3_LR3 fragment. (D) Gel image of the linearized vector pCLY10_C3_LR_flank with a size of 9002 bp. 1 kb ladder (L) was used as a reference.

Before C3_LR_flank and the shuttle vector pCLY10 were ligated, a digestion restriction with the enzymes *Hind*III and *Not*I was performed. Finally, the modified vector, namely pCLY10_C3_LR_flank was produced. To prove if the ligation was successful, an analytical restriction with *Xho*I and *Pme*I was performed. In Figure 26C two bands appeared on the gel at a size of 2096 bp and 6894 bp, representing the C3_LR3 fragment and the linearized vector pCLY10. The modified vector was next linearized with the endonuclease *Pme*I to enable homologous recombination between the “capture arms” of the modified vector and the ends of the DNA-fragments 11A7_pCC1FOS and 13H6_pCC1FOS. Fig. 26D shows a gel image with the expected band of the linearized vector (9002 bp). The process of the construction of the modified vector is shown as a schematic overview in Fig. 27.

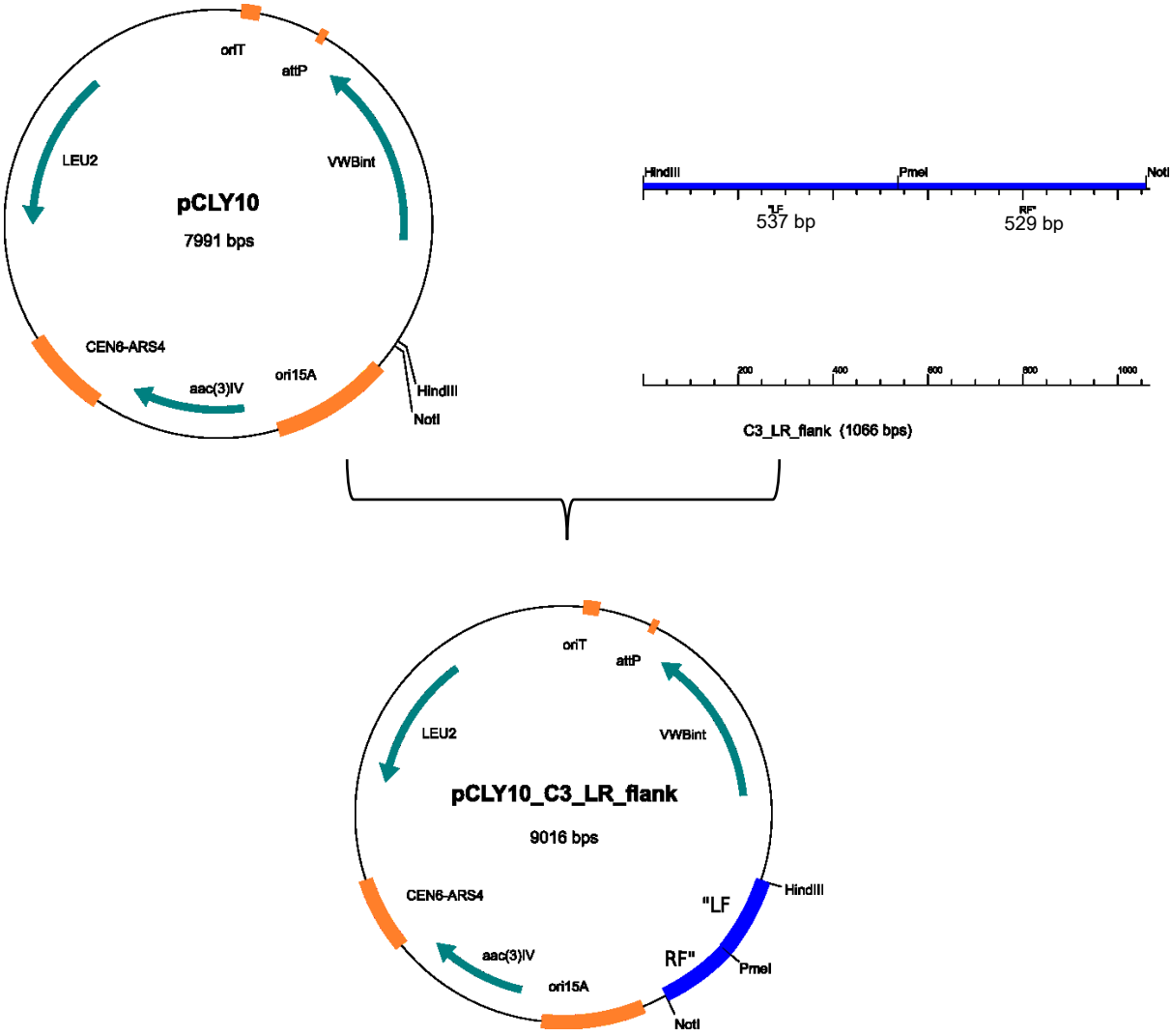


Figure 27: Schematic overview of the construction of pCLY10_C3_LR_flank. The modified vector has a size of 9016 bp.

8.5.2. Modification of pCLY10 for assembly of BGC4

C4_LR_flank which contains flanking regions of BGC4 which shall function as “capture arms” was synthesized in a pUC57 vector by BioCat (Germany). For the production of pCLY10_C4_LR_flank, the shuttle vector pCLY10 and C4_LR_flank were digested with the same enzymes *Apa*I and *Hind*III. After performing a gel electrophoresis as shown in Fig. 28A and 28B, the DNAs of pCLY10 and C4_LR_flank were recovered and subsequently ligated.

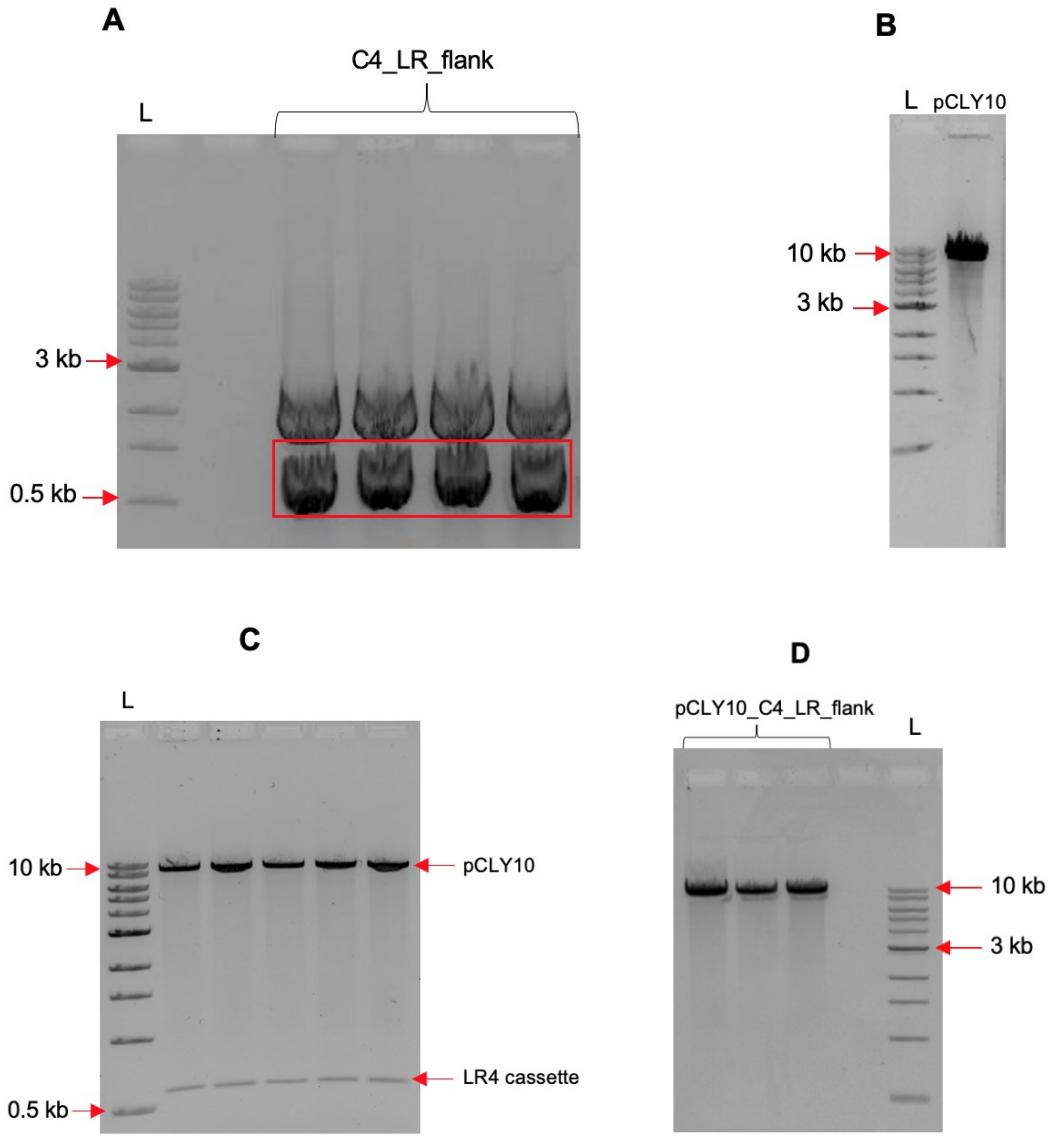


Figure 28: (A) Digestion of C4_LR_flank with *Apa*I and *Hind*III. (B) Digestion of pCLY10 with *Apa*I and *Hind*III. (C) Analytical digestion of pCLY10_C4_LR_flank with *Hind*III and *Pme*I to prove if the cloning was successful. The band at the top with a size of 8483 bp represents the shuttle vector and the band at the bottom with a size of 593 bp represents the LR4 cassette. (D) Gel image of the linearized vector pCLY10_C4_LR_flank with a size of 9095 bp. 1 kb ladder (L) was used as a reference.

For the verification of the successful ligation, an analytical digestion was done with *Hind*III and *Pme*I. Figure 28C shows a gel image where two bands appeared. The band above (8483 bp) belongs to the shuttle vector pCLY10 and the smaller band (593 bp) belongs to the LR4 cassette. Subsequently, a linearization of the modified vector was performed with the enzyme *Pme*I to enable homologous recombination between the “capture arms” of the modified vector and the ends of the DNA-fragments pFOS1_5H3 and pFOS1_1H1. Fig. 28D shows the expected single band of pCLY10_C4_LR_flank of 9095 bp. The process of the construction of the modified vector is shown as schematic overview in Fig. 29.

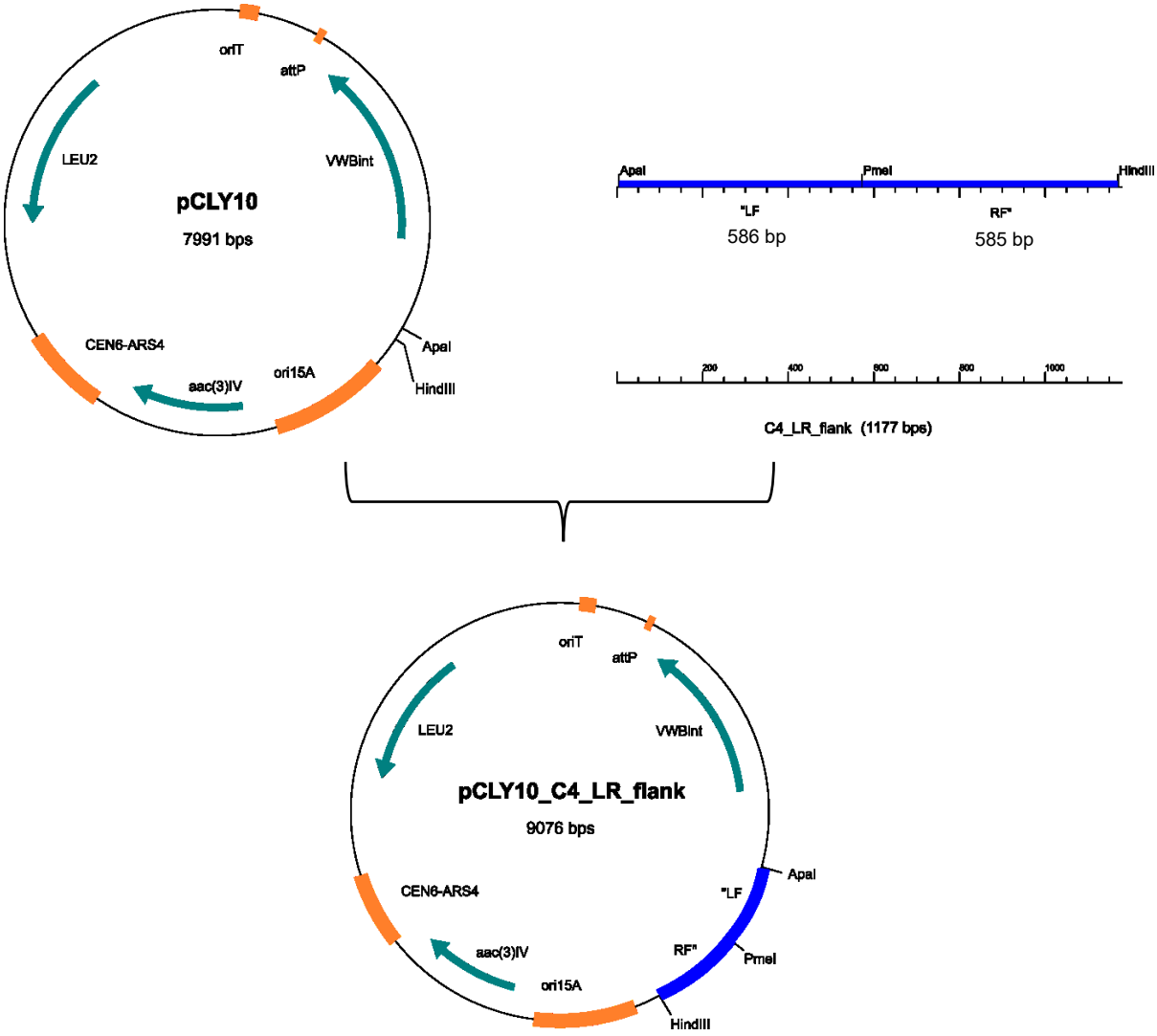


Figure 29: Schematic overview of the construction of pCLY10_C4_LR_flank, which was created by ligating the shuttle vector pCLY10 with C4_LR_flank. The modified vector is 9076 bp in size.

8.6. Yeast Assembly of BGC3 and BGC4

8.6.1. BGC3

For performing yeast assembly, previously prepared fragments were transformed into *Saccharomyces cerevisiae.* In case of BGC3, the linearized vector pCLY10_C3_LR_flank and the DNA-fragments 11A7_pCC1FOS and 13H6_pCC1FOS were used. After the transformation, a yeast colony PCR was performed with primersets a, b and c and 33 out of 40 randomly picked clones showed a positive signal on the gel for all primersets. This was a demonstration of a successful assembly, which means that those 33 colonies contain the vector with the complete BGC3, named pSS_C3 (Fig. 30).

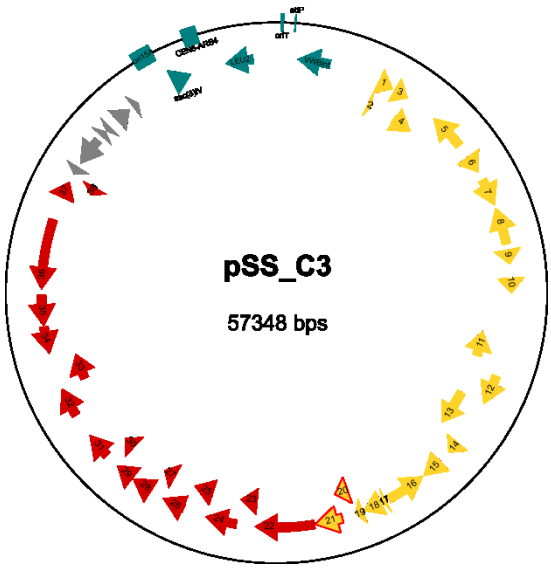


Figure 30: Schematic overview of the vector pSS_C3. The colored arrows are representing the genes of BGC3 and numbered like the gene composition of cluster 3 listed in Table 16. The yellow arrows belong to 11A7_pCC1FOS, while the red ones belong to 13H6_pCC1FOS. Grey arrows are out of the cluster. Genes of pCLY10 are colored in petrol. pSS_C3 has a size of 57348 bp.

After the pDNA was isolated from yeast, a transformation in *E.coli* EPI300 was performed in order to enhance the plasmid copy number. An analytical digestion with *EcoR*I was carried out to test if the right clones were obtained. Fig. 31 shows the gel image of the pSS_C3 digestion after it was transformed into *E.coli* EPI300. As expected, three bands with the sizes of 3661 bp, 14522 bp and 39165 bp were visible on the gel.

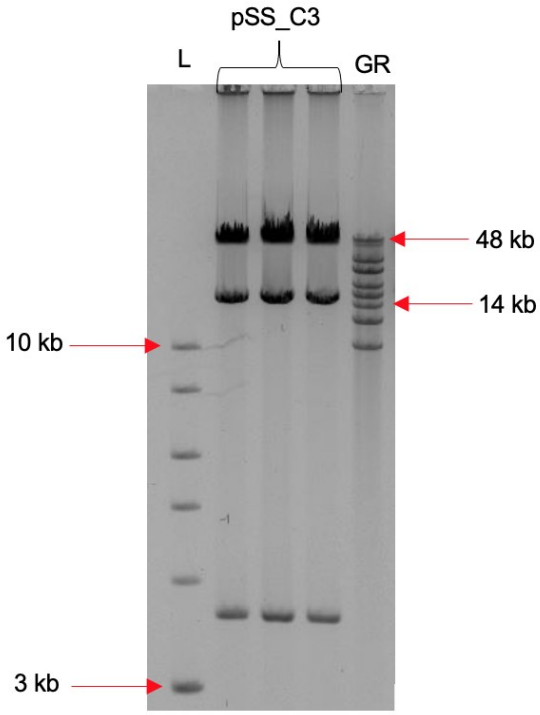


Figure 31: Gel image of the restriction with *EcoR*I after the transformation of pSS_C3 into *E.coli* EPI300. Three bands appeared as expected with a size of 3661 bp, 14522 bp and 39165 bp. GeneRuler High Range DNA Ladder (GR) and 1 kb Ladder (L) were used as marker.

Verified pSS_C3 plasmid was transformed into *E.coli* ET12567 for intergenic conjugation into the *Streptomyces* strains chosen for heterologous expression. During this project the following strains were used: *Streptomyces albus* wt, *Streptomyces albus* B4 and *Streptomyces coelicolor* M1154. The pSS_C3 plasmid was conjugated from *E.coli* ET12567 into the *Streptomyces* strains following the established protocol (Methods 7.17). The obtained transconjugants were fermented and their extracts obtained according to the protocols described in Methods 7.18.1 and 7.18.2.

8.6.2. BGC4

For BGC4, the linearized vector pCLY10_C4_LR_flank and the DNA-fragments pFOS1_5H3 and pFOS1_1H1 were used for yeast assembly. Additionally, the synthesized PCR-fragments, which contain missing genes of BGC4 (5H3_PCR and 1H1_PCR) were introduced into *S.cerevisiae.* Fig. 32 shows physical map of the expected plasmid pSS_C4. After yeast transformants appeared on the plates, colony PCR was performed with primersets d, e and f, to test for correct assembly of pSS_C4. 12 out of 40 randomly picked yeast colonies showed expected PCR products on the gel.

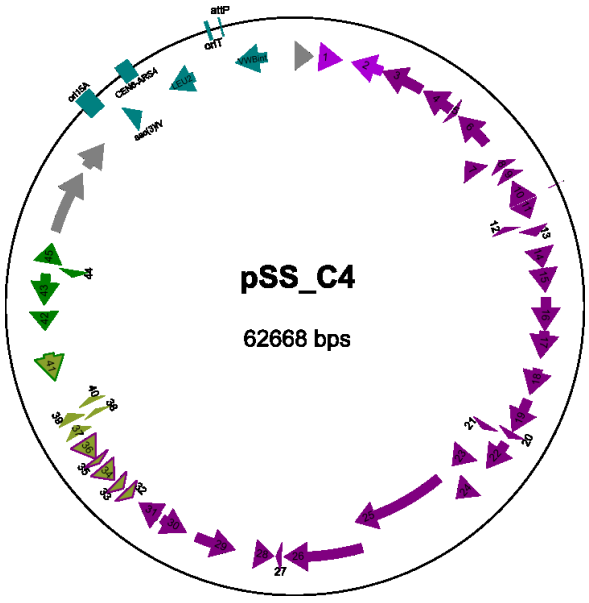


Figure 32: Schematic overview of the plasmid pSS_C4. The colored arrows represent the genes of BGC4, numbered as for BGC4 listed in Table 17. Grey arrows represent genes outside of the cluster. The light purple arrows belong to 5H3_PCR, the majority of genes, colored in dark purple represent pFOS1_5H3. Light green ones belong to 1H1_PCR and dark green arrows belong to pFOS1_1H1. Genes of pCLY10 are colored in petrol. pSS_C4 has a size of 62668 bp.

After the isolation of pDNA from yeast, a transformation in *E.coli* EPI300 was performed in order to amplify the plasmid from single copy. To verify if the plasmid was generated correctly, an analytical restriction with *Hind*III was carried out. Fig. 33 is showing the gel

image of the restriction of pSS_C4 after the transformation in *E.coli* EPI300, where three expected bands with a size of 1890 bp, 12931 bp and 47843 bp were visible.

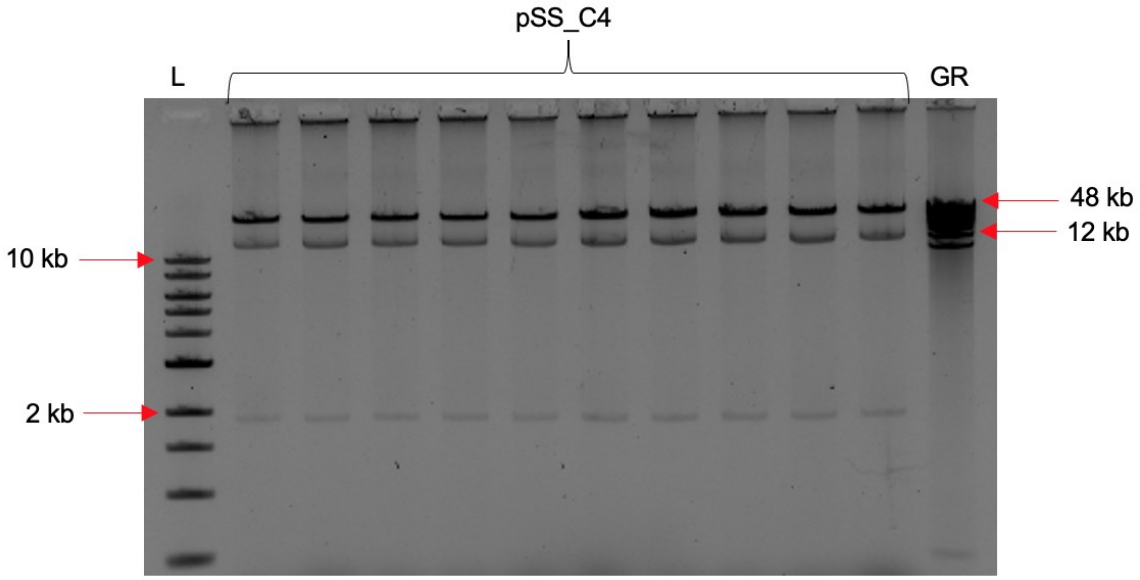


Figure 33: Gel image of the digestion with *Hind*III after transforming pSS_C4 into *E.coli* EPI300. As expected, three bands with a size of 1890 bp, 12931 bp and 47843 bp appeared on the gel. GeneRuler High Range DNA Ladder (GR) and 1 kb Ladder (L) were used as marker.

Subsequently, pSS_C4 carrying BGC4 was transformed into *E.coli* ET12567 and then conjugated into the same *Streptomyces* strains as those use for pSS_C3. The obtained transconjugants were fermented and their extracts obtained according to the protocols described in Methods 7.18.1 and 7.18.2.

8.7. Analysis of the extracts

Assembled BGC3 and BGC4 were thus transformed into different heterologous hosts: *Streptomyces albus* wt, *Streptomyces albus* B4 and *Streptomyces coelicolor* M1154, which were cultivated in different fermentation media. The extracts from the supernatant and the pellet were analyzed via HPLC and tested for bioactivity using disc diffusion assay. As controls, the strains which carried unmodified pCLY10 vector were used. No additional peaks were detected in HPLC chromatograms comparing to the control extracts and no bioactivity could be observed against microorganisms tested in this work. As an example, in Figure 34 shows the chromatograms of two *S.albus* recombinant strains containing pSS_C3 plasmid (blue and red) and one control strain (in green), which were fermented in SM17 medium, the culture was freeze-dried and extracted with 100 % methanol. The absence of new metabolites, which could be the products of heterologously expressed cluster, was confirmed

also via LC-MS, indicating that the expression of BGC3 and BGC4 in recombinant strains failed.

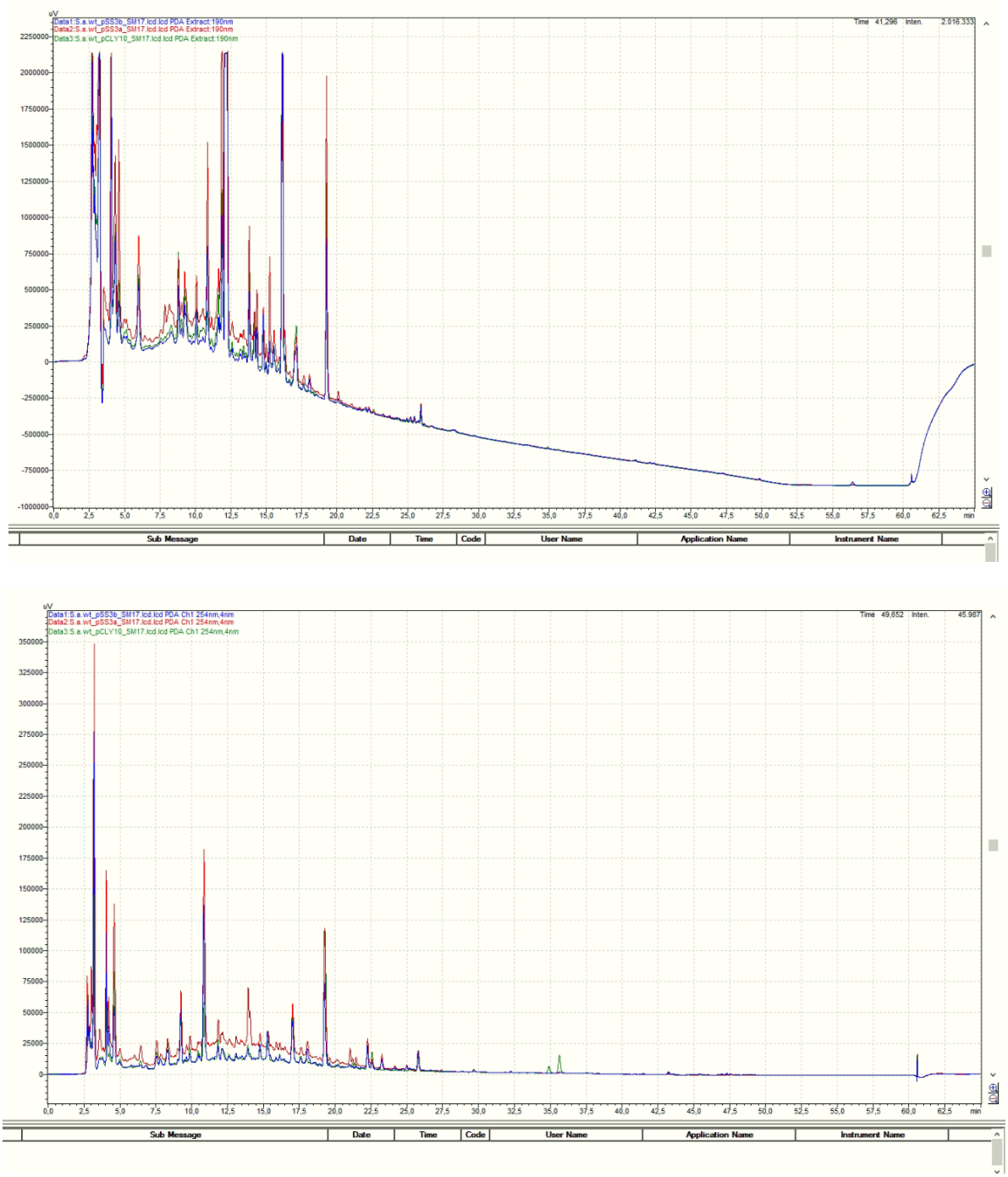


Figure 34: HPLC chromatograms of of the extracts from recombinant *Streptomyces albus* wt and *Streptomyces albus* B4 habrouring BGC3 and BGC4. The blue and red lines represent two *S. albus* recombinant strains, while the green one belongs to the control strain containing the unmodified pCLY10 vector. No additional peaks were detected in HPLC chromatograms comparing to the control extracts.

9. Discussion

Since many years the discovery of NPs originating from animals, plants or microorganisms is very important, for they are extensively used in human and animal medicine, as well as in agriculture. Because of the rise in the infections caused by antibiotic-resistant pathogens, the need for finding new NPs representing novel antibiotics became even more important.

The goal of this project was to discover new NPs by heterologously expressing two unique biosynthetic gene clusters, BGC3 and BGC4 from the actinobacterial strain *Actinoalloteichus fjordicus* DSM 46856. The bioinformatic online tool, antiSMASH, predicted that BGC3 should encode for a ladderane-like compound, which is described to be a lipid membrane structure in bacteria. [53] BGC4 could specify a NRP-PK hybrid compound produced by a biosynthetic machinery of type I PKSs and NRPSs enzyme systems. [10]

Actinoalloteichus fjordicus is a marine and slowly growing rare actinobacterial strain. The genetic modification of this strain would be very difficult, and thus was not considered good strategy.

In this work, a genome library of *Actinoalloteichus fjordicus* was successfully constructed in order to obtain clones of DNA fragments covering the full length of its genome. The actinobacterial strain *Actinoalloteichus fjordicus* has a genome size of 7275385 bp. According to the protocol for the creation of a complete fosmid library, 835 clones were needed, using a formula to determine the number of required fosmid clones. To ensure that the complete gDNA was covered, a total of 1728 *E.coli* clones were used. Even though twice more clones were used, extensive screening for BGC4-containing clones showed that still some fragments were stil missing, and had to be PCR-amplified from the genomic DNA.

It is unclear why some parts of the gDNA of *Actinoalloteichus fjordicus* were not represented in the genomic library, but most probably the secondary structures of GC-rich actinobacterial DNA could be a reason for inefficient cloning of such fragments in *E. coli*. It is possible, that because of the specificity in folding of some parts in DNA, some sequences of gDNA were broken into smaller fragments and were not cut out from the gel within the first steps of library construction and thus could not be packaged into the phages. As result, those fragments could have been lost and not included into the library.

Since the number of clones required for 10x genome coverage was determined on the example of *E. coli* genome-library production [54], it could be shown in this work, that the screening of higher clone numbers is required, at least for DNA originated from *Actinoalloteichus fjordicus* strain.

In this project, yeast assembly was successful for both gene clusters chosen. The pCLY10 vector harboring BGC3 and BGC4 were transformed into three heterologous hosts: *Streptomyces albus* J1074, *Streptomyces albus* B4 and *Streptomyces coelicolor* M1154. The expression of BGC3 and BGC4 was tested in different media with different conditions and a metabolite profile was analyzed by HPLC.

Unfortunately, in HPLC data the production of new compounds could not be shown, and no bioactivity against microorganisms tested in this work could be observed. Host strains for heterologous expression can be a limiting factor for the activation of BGCs in order to produce NPs. Around 90% of the BGCs are inactive or only active under specific fermentation conditions. [55] For achieving BGC expression in different host organisms special genetic parts, tools and techniques are necessary. Maybe different fermentation media or optimization of fermentation conditions could influence the production of SMs, as well as the use of other host strains, such as *Streptomyces lividans* or *Streptomyces venezuelae,* which could be more suitable for heterologous expression of BGC3 and BGC4. In comparison to *Streptomyces coelicolor*, the host organism *Streptomyces lividans* is able to accept foreign DNA more efficient, because it has no restriction on foreign methylated DNA. As an example that a change of host strain can influence the production of SMs, it was shown that a mutant of *S.lividans* could heterologously produce larger number of NPs compared to *S.coelicolor*. [55][56][57]

Due to the fact that many BGCs are silent and not expressed under laboratory conditions, several strategies to activate them and increase the possibility of the production of the target compound have been developed. Since in both clusters the genes encoding SARP-regulator were identified, the overexpression of an activator gene provides the possibility to activate BGCs of interest. For example, the regulator DnrI belonging to the family of the SARP-regulators has proven such potential by upregulating the biosynthesis of the antibiotic and anticancer agent daunorubicin in *Streptomyces peucetius*. [34][31][58]

Another possibility is the inactivation of TetR-like regulators, which were also detected in both BGCs, and known to act as repressors for SMs production in *Streptomyces*. Many reports showed a successful production of heterologously expressed compounds in *Streptomyces* strains after targeting of regulator genes. For example, the deletion of TetR-like repressor encoded genes yielded the production of kanamycin in *Streptomyces ambofaciens* and gaburedins in *Streptomyces venezuelae*. [31]

Taking the above into account, it can be stated that the modification of cluster-specific regulator genes could be a key for the activation of BGC3 and BGC4 in a heterologous host and production of potentially novel NPs.

10. Conclusion and Outlook

Increasing improvements in genome sequencing and analyses, and new approaches such as genome mining, are facilitating the discovery of new NPs. This greatly improves our knowledge on the microorganisms' silent BGCs and allows genetic manipulation aimed at their activation and therefore production of novel NPs.

Within this work, several important goals have been achieved. A gene library for the rare marine actinobacetrium *Actinoalloteichus fjordicus* DSM 46856 has been constructed. To date, no such library has been reported in the literature. The following screening of this library successfully identified clones covering complete BGC3 and most of the BGC4. Using yeast-based assembly and PCR-assisted amplification of the missing parts of BGC4, both clusters were assembled into the shuttle vector. Resulting assembled BGCs were introduced into various *Streptomyces* hosts for heterologous expression.

Although no functional expression of the clusters was demonstrated, this work sets an essential background for follow-up genetic manipulations aimed at their activation.
Expression of the SARP-regulators encoded by the BGC3 and BGC4 in the recombinant *Streptomyces* carrying respective clusters will be the next logical step in attempt to activate them and to achieve production of the NPs they specify.

11. Acknowledgements

First of all, I would like to thank my supervisor Univ.- Prof. Dr. Sergey B. Zotchev for giving me the opportunity to accomplish my diploma thesis and to work in this interesting research field with his friendly working group.
It was a great experience to research in the laboratory for a while and I do not want to miss this time. I am thankful for what I have learned practically and theoretically in this period and also for all the good advices and exchange of experience with members from the whole group.

Special thanks to my co-supervisor Dr. Olha Schneider. She supported me in any way, no matter what and was always there when I needed something. I really appreciate her endurous support and motivation from the beginning of my practical work until the end of my diploma thesis.

Next I want to thank every other member of the working group of the pharmaceutical biotechnology department for the helping hands, the pleasant working atmosphere and the celebrations of different occasions, even though they had to be under special conditions during this extraordinary time.

Moreover, I would like to thank my family for their continuous support during my whole studies and for always being there for me whenever I needed something, my boyfriend for always keeping my back, motivating me in every situation and helping me finish my studies. I also want to thank my best friends from my handball team, who made this time unforgettable, by making parties, vacation trips or simply by just sitting together and talking about everything. Last but not least, a big thank you belongs to “Dönerbanditen”, which made studying easier, because of shared learning sessions, trips and parties.

12. References

[1] L. Kourkouta, A. Tsaloglidou, and K. Koukourikos, “History of Antibiotics,” *Sumerianz Journal od Medical and Healthcare*, vol. 1, no. 2. pp. 51-54, 2018

[2] L. Katz and R. H. Baltz, “Natural product discovery: past, present, and future,” *Journal of Industrial Microbiology and Biotechnology*, vol. 43, no. 2–3, pp. 155–176, doi: 10.1007/s10295-015-1723-5, 2016

[3] S. Mushtaq, B. H. Abbasi, B. Uzair, and R. Abbasi, “Natural products as reservoirs of novel therapeutic agents,” *EXCLI J*, vol. 17, pp. 420–451, doi: 10.17179/excli2018-1174, 2018

[4] S. Bernardini, A. Tiezzi, V. Laghezza Masci, and E. Ovidi, “Natural products for human health: an historical overview of the drug discovery approaches,” *Nat Prod Res*, vol. 32, no. 16, pp. 1926–1950, doi: 10.1080/14786419.2017.1356838, 2018

[5] O. N. Sekurova, O. Schneider, and S. B. Zotchev, “Novel bioactive natural products from bacteria via bioprospecting, genome mining and metabolic engineering,” *Microb. Biotechnol.*, vol. 12, no. 5, pp. 828–844, 2019, doi: 10.1111/1751-7915.13398, 2019

[6] M. H. Medema and M. A. Fischbach, “Computational approaches to natural product discovery,” *Nat Chem Biol*, vol. 11, no. 9, pp. 639–648, doi: 10.1038/nchembio.1884, 2015

[7] S. B. Zotchev, “Antibiotics: Biosynthesis,” in *Wiley Encyclopedia of Chemical Biology*, American Cancer Society, pp. 1–9. doi: 10.1002/9780470048672.wecb014, 2008

[8] C. Olano, F. Lombó, C. Méndez, and J. A. Salas, “Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering,” *Metabolic Engineering*, vol. 10, no. 5, pp. 281–292, doi: 10.1016/j.ymben.2008.07.001, 2008

[9] M. J. Calcott and D. F. Ackerley, “Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products,” *Biotechnol Lett*, vol. 36, no. 12, pp. 2407–2416, doi: 10.1007/s10529-014-1642-y, 2014

[10] J. Soltani, “Secondary Metabolite Diversity of the Genus Aspergillus: Recent Advances,” in *New and Future Developments in Microbial Biotechnology and Bioengineering*, Elsevier, pp. 275–292. doi: 10.1016/B978-0-444-63505-1.00035-X, 2016

[11] K. L. Tedesco and M. J. Rybak, “Daptomycin,” *Pharmacotherapy*, vol. 24, no. 1, pp. 41–57, doi: 10.1592/phco.24.1.41.34802, 2004

[12] Y.-S. Wu, S.-C. Ngai, B.-H. Goh, K.-G. Chan, L.-H. Lee, and L.-H. Chuah, “Anticancer Activities of Surfactin and Potential Application of Nanotechnology Assisted Surfactin Delivery,” *Front. Pharmacol.*, vol. 8, p. 761, doi: 10.3389/fphar.2017.00761, 2017

[13] Q. Guan, S. Huang, Y. Jin, R. Campagne, V. Alezra, and Y. Wan, “Recent Advances

in the Exploration of Therapeutic Analogues of Gramicidin S, an Old but Still Potent Antimicrobial Peptide," *J. Med. Chem.*, vol. 62, no. 17, pp. 7603–7617, doi: 10.1021/acs.jmedchem.9b00156, 2019

[14] X. Yang and A. E. Yousef, "Antimicrobial peptides produced by *Brevibacillus* spp.: structure, classification and bioactivity: a mini review," *World J Microbiol Biotechnol*, vol. 34, no. 4, p. 57, doi: 10.1007/s11274-018-2437-4, 2018

[15] K. J. Weissman, "The structural biology of biosynthetic megaenzymes," *Nat Chem Biol*, vol. 11, no. 9, pp. 660–670, doi: 10.1038/nchembio.1883, 2015

[16] B. Shen, "Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms," *Current Opinion in Chemical Biology*, vol. 7, no. 2, pp. 285–295, doi: 10.1016/S1367-5931(03)00020-6, 2003

[17] A. Nivina, K. P. Yuet, J. Hsu, and C. Khosla, "Evolution and Diversity of Assembly-Line Polyketide Synthases: Focus Review," *Chem. Rev.*, vol. 119, no. 24, pp. 12524–12547, doi: 10.1021/acs.chemrev.9b00525, 2019

[18] H. Jenke-Kodama, A. Sandmann, R. Müller, and E. Dittmann, "Evolutionary Implications of Bacterial Polyketide Synthases," *Molecular Biology and Evolution*, vol. 22, no. 10, pp. 2027–2039, doi: 10.1093/molbev/msi193, 2005

[19] S. Caulier, C. Nannan, A. Gillis, F. Licciardi, C. Bragard, and J. Mahillon, "Overview of the Antimicrobial Compounds Produced by Members of the *Bacillus subtilis* Group," *Front. Microbiol.*, vol. 10, p. 302, doi: 10.3389/fmicb.2019.00302, 2019

[20] Y. A. Chan, A. M. Podevels, B. M. Kevany, and M. G. Thomas, "Biosynthesis of polyketide synthase extender units," *Nat. Prod. Rep.*, vol. 26, no. 1, pp. 90–114, doi: 10.1039/B801658P, 2009

[21] J. Staunton and B. Wilkinson, "Biosynthesis of Erythromycin and Rapamycin," *Chem. Rev.*, vol. 97, no. 7, pp. 2611–2630, doi: 10.1021/cr9600316, 1997

[22] M. Jiang, H. Zhang, S.-H. Park, Y. Li, and B. A. Pfeifer, "Deoxysugar pathway interchange for erythromycin analogues heterologously produced through *Escherichia coli*," *Metabolic Engineering*, vol. 20, pp. 92–100, doi: 10.1016/j.ymben.2013.09.005, 2013

[23] B. S. Moore and J. N. Hopke, "Discovery of a New Bacterial Polyketide Biosynthetic Pathway," p. 4, 2001

[24] H. Zhu, S. K. Sandiford, and G. P. van Wezel, "Triggers and cues that activate antibiotic production by actinomycetes," *Journal of Industrial Microbiology and Biotechnology*, vol. 41, no. 2, pp. 371–386, doi: 10.1007/s10295-013-1309-z, 2014

[25] L. Huo, J. J. Hug, C. Fu, X. Bian, Y. Zhang, and R. Müller, "Heterologous expression of bacterial natural product biosynthetic pathways," *Nat. Prod. Rep.*, vol. 36, no. 10, pp. 1412–1436, doi: 10.1039/C8NP00091C, 2019

[26] J. P. Gomez-Escribano and M. J. Bibb, “Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: from genome mining to manipulation of biosynthetic pathways,” *J Ind Microbiol Biotechnol*, vol. 41, no. 2, pp. 425–431, doi: 10.1007/s10295-013-1348-5, 2014

[27] M. H. Medema, K. Blin, and P. Comermancic, “antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences,” *Nucleic Acids Research*, vol. 39, no. suppl_2, pp. W339–W346, doi: 10.1093/nar/gkr466, 2011

[28] L. Li, W. Jiang, and Y. Lu, “New strategies and approaches for engineering biosynthetic gene clusters of microbial natural products,” *Biotechnology Advances*, vol. 35, no. 8, pp. 936–949, doi: 10.1016/j.biotechadv.2017.03.007, 2017

[29] A. K. Chavali and S. Y. Rhee, “Bioinformatics tools for the identification of gene clusters that biosynthesize specialized metabolites,” *Brief Bioinform*, vol. 19, no. 5, pp. 1022–1034, doi: 10.1093/bib/bbx020, 2018

[30] R. Pan, X. Bai, J. Chen, H. Zhang, and H. Wang, “Exploring Structural Diversity of Microbe Secondary Metabolites Using OSMAC Strategy: A Literature Review,” *Front Microbiol*, vol. 10, p. 294, doi: 10.3389/fmicb.2019.00294, 2019

[31] P. J. Rutledge and G. L. Challis, “Discovery of microbial natural products by activation of silent biosynthetic gene clusters,” *Nat Rev Microbiol*, vol. 13, no. 8, pp. 509–523, doi: 10.1038/nrmicro3496, 2015

[32] O. N. Sekurova, I. Pérez-Victoria, and J Martín, “New Deferoxamine Glycoconjugates Produced upon Overexpression of Pathway-Specific Regulatory Gene in the Marine Sponge-Derived *Streptomyces albus* PVA94-07,” *Molecules*, vol. 21, no. 9, p. 1131, doi: 10.3390/molecules21091131, 2016

[33] B. Aigle and C. Corre, “Waking up *Streptomyces* Secondary Metabolism by Constitutive Expression of Activators or Genetic Disruption of Repressors,” in *Methods in Enzymology*, vol. 517, Elsevier, pp. 343–366. doi: 10.1016/B978-0-12-404634-4.00017-6, 2012

[34] A. Romero-Rodríguez, I. Robledo-Casados, and S. Sánchez, “An overview on transcriptional regulators in *Streptomyces*,” *Biochim Biophys Acta*, vol. 1849, no. 8, pp. 1017–1039, doi: 10.1016/j.bbagrm.2015.06.007, 2015

[35] L. Cuthbertson and J. R. Nodwell, “The TetR family of regulators,” *Microbiology and molecular biology reviews : MMBR*, vol. 77, no. 3, doi: 10.1128/MMBR.00018-13, 2013

[36] J. L. Ramos, M. Martínez-Bueno, and A. J. Molina-Henares, “The TetR Family of Transcriptional Repressors,” *Microbiol Mol Biol Rev*, vol. 69, no. 2, pp. 326–356, doi: 10.1128/MMBR.69.2.326-356.2005, 2005

[37] M. J. Bibb, G. R. Janssen, and J. M. Ward, “Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*” *Elsevir,* Gene, 38, 215-226, 1985

[38] R. H. Baltz, “Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes,” *J Ind Microbiol Biotechnol*, p. 28., 2015

[39] O. Bilyk, O. N. Sekurova, S. B. Zotchev, and A. Luzhetskyy, “Cloning and Heterologous Expression of the Grecocycline Biosynthetic Gene Cluster,” *PLoS ONE*, vol. 11, no. 7, p. e0158682, doi: 10.1371/journal.pone.0158682, 2016

[40] J. Krause, I. Handayani, K. Blin, A. Kulik, and Y. Mast, “Disclosing the Potential of the SARP-Type Regulator PapR2 for the Activation of Antibiotic Gene Clusters in *Streptomycetes*,” *Front. Microbiol.*, vol. 11, p. 225, doi: 10.3389/fmicb.2020.00225, 2020

[41] S. E. Ongley, X. Bian, B. A. Neilan, and R. Müller, “Recent advances in the heterologous expression of microbial natural product biosynthetic pathways,” *Nat. Prod. Rep.*, vol. 30, no. 8, p. 1121, doi: 10.1039/c3np70034h, 2013

[42] A. Tocchetti, S. Donadio, and M. Sosio, “Large inserts for big data: artificial chromosomes in the genomic era,” *FEMS Microbiology Letters,* 365(9). doi: 10.1093/femsle/fny064, 2018

[43] W. Jiang and T. F. Zhu, “Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments,” *Nat Protoc*, vol. 11, no. 5, pp. 960–975, doi: 10.1038/nprot.2016.055, 2016

[44] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. H. Iii, and H. O. Smith, “Enzymatic assembly of DNA molecules up to several hundred kilobases,”*Nature Methods*, 6(5), pp. 343-345.doi: 10.1038/nmeth.1318, 2009

[45] H.-J. Nah, H.-R. Pyeon, S.-H. Kang, S.-S. Choi, and E.-S. Kim, “Cloning and Heterologous Expression of a Large-sized Natural Product Biosynthetic Gene Cluster in *Streptomyces* Species,” *Front. Microbiol.*, vol. 8, doi: 10.3389/fmicb.2017.00394, 2017

[46] J. H. Kim, Z. Feng, J. D. Bauer, D. Kallifidas, P. Y. Calle, and S. F. Brady, “Cloning large natural product gene clusters from the environment: Piecing environmental DNA gene clusters back together with TAR,” *Biopolymers*, vol. 93, no. 9, pp. 833–844, doi: 10.1002/bip.21450, 2010

[47] D. L. Court, J. A. Sawitzke, and L. C. Thomason, “Genetic Engineering Using Homologous Recombination,” *Annu. Rev. Genet.*, vol. 36, no. 1, pp. 361–388, doi: 10.1146/annurev.genet.36.061102.093104, 2002

[48] S.-J. Lee, Y.-S. Jeong, D.-U. Kim, J.-W. Seo, and B.-K. Hur, “Eicosapentaenoic Acid

(EPA) Biosynthetic Gene Cluster of *Shewanella oneidensis* MR-1: Cloning, Heterologous Expression, and Effects of Temperature and Glucose on the Production of EPA in *Escherichia coli*" *Biotechnology and Bioprocess Enginering*, 11, pp. 510-515, 2006

[49] S.-U. Choi, C.-K. Lee, Y.-I. Hwang, H. Kinoshita, and T. Nihira, "Intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to *Kitasatospora setae* , a bafilomycin B 1 producer," *Archives of Microbiology*, vol. 181, no. 4, pp. 294–298, doi: 10.1007/s00203-004-0654-8, 2004

[50] S. A. Nicolaou, S. M. Gaida, and E. T. Papoutsakis, "Exploring the combinatorial genomic space in *Escherichia coli* for ethanol tolerance," *Biotechnology Journal*, vol. 7, no. 11, pp. 1337–1345, doi: 10.1002/biot.201200227, 2012

[51] M. Myronovskyi, "Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters," *Metabolic Engineering*, 49, pp. 316-324, 2018

[52] C. B. Brachmann, A. Davies, G. J. Cost, E. Caputo, J. Lee, P. Hieter, and J. D. Boeke, "Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications," *Yeast*, vol. 14, pp. 115-132, 1998

[53] E. N. Hancock and M. K. Brown"Ladderane Natural Products: From the Ground Up" *Chem. Eur. J.*, 10.1002/chem.202002866, 2020

[54] Epicentre, "CopyControl™ Fosmid Library Production Kit with pCC1FOS™-Vector", 2012

[55] J. Ke and Y. Yoshikuni, "Multi-chassis engineering for heterologous production of microbial natural products," *Current Opinion in Biotechnology*, vol. 62, pp. 88–97, doi: 10.1016/j.copbio.2019.09.005, 2020

[56] M. D. McMahon, C. Guan, J. Handelsman, and M. G. Thomas, "Metagenomic Analysis of *Streptomyces lividans* Reveals Host-Dependent Functional Expression," *Appl. Environ. Microbiol.*, vol. 78, no. 10, pp. 3622–3629, doi: 10.1128/AEM.00044-12, 2012

[57] J. Kormanec, R. Novakova, D. Csolleiova, L. Feckova, B. Rezuchova, B. Sevcikova, and D. Homerova, "The antitumor antibiotic mithramycin: new advanced approaches in modification and production," *Appl Microbiol Biotechnol*, vol. 104, no. 18, pp. 7701–7721, doi: 10.1007/s00253-020-10782-x, 2020

[58] N. Bate, G. Stratigopoulos, and E. Cundliffe, "Differential roles of two SARP-encoding regulatory genes during tylosin biosynthesis: Regulation of tylosin production," *Molecular Microbiology*, vol. 43, no. 2, pp. 449–458, doi: 10.1046/j.1365-2958.2002.02756.x, 2002