Doctoral Thesis

Evolution, environmental distribution, and engineering of the abyssomicin biosynthetic gene cluster

Alba Iglesias Vilches

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

in the

School of Natural and Environmental Sciences

January 2021

Abstract

Alba IGLESIAS VILCHES

Evolution, environmental distribution, and engineering of the abyssomicin biosynthetic gene cluster.

Microbial secondary metabolites constitute a great source of pharmaceutically interesting biomolecules. In particular, the tetronate family of natural products is a structurally and functionally diverse group of secondary metabolites whose potent bioactivities make them attractive targets for clinical and industrial exploitation. The abyssomicins are an actively growing family of small spirotetronate natural products that has been widely studied due to the unique structural features and bioactivities that some of its members exhibit, including antimicrobial activity against Gram-positive bacteria such as methicillin- and vancomycin-resistant *Staphylococcus aureus* and different *Mycobacteria* strains, HIV inhibitory and reactivator properties and anti-influenza A virus activity.

Abyssomicin C and its atrop- isomer, produced by the slow growing marine *Actinobacteria Micromonospora maris* AB-18-032 ^{T}, are type I polyketide antibiotics that inhibit the formation of *p*-aminobenzoic acid, a constituent of the folate pathway. Abyssomicin biosynthesis is highly amenable to reengineering, as the enzymes involved in the synthesis of the tetronate (AbyA1) and the spiro-tetronate-forming Diels-Alderase (AbyU) are both capable of accepting structurally diverse substrates. The aim of this project was to set up the grounds for the discovery and production of novel abyssomicins with applications in the biopharmaceutical industry. First, in order to understand the environmental distribution and evolution of the abyssomicin biosynthetic gene clusters (BGCs) present in nature, an analysis of publicly available genomic and metagenomic data was carried out. The strategy of selecting a pathway-specific enzyme to direct the mining proved to be an excellent strategy; 74 new Diels–Alderase homologs were identified and a surprising prevalence of the abyssomicin BGC within terrestrial habitats, mainly soil and plant-associated, was unveiled. Five complete and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs were also identified, suggesting that a plethora of abyssomicins remain to be discovered. A preliminary study on the abyssomicin production potential of five of the strains containing potential abyssomicin BGCs was also carried out although no abyssomicins were found.

After that, with the final goal of producing abyssomicins of various lengths and different saturation/oxidation patterns, it was necessary to express the *aby* BGC of *M. maris* AB-18-032 in a well-established heterologous host. This cluster was successfully moved into *E. coli* and various *Streptomyces* species, the abyssomicin production potential of these strains was evaluated in various conditions and some of the hosts were promoter engineered to force the expression of the *aby* BGC. Active gene expression was demonstrated, but despite the efforts, none of the heterologous hosts produced abyssomicins. Later analysis unveiled the presence of several mutations within *abyB1*, the first polyketide synthase gene in the *aby* BGC, suggesting this could be the reason for the lack of production.

Since the approach to heterologously produce abyssomicins was not fruitful, this work then focused on increasing abyssomicin production in *M. maris* AB-18-032 and developing genetic tools for this system. First, through ribosome engineering, a library of *M. maris* drug-resistant mutants capable of producing up to 3.4-fold abyssomicin C in comparison to the wild-type strain was generated. Then, using statistical Design of Experiments (DOE), an efficient electroporation protocol that could accelerate targeted genetic manipulations in *M. maris* was developed. Together, increased abyssomicin production and a quick and easy electroporation protocol for *M. maris*, will facilitate future engineering of the *aby* BGC directly in *M. maris* to produce diverse non-natural abyssomicins.

Acknowledgements

I would like to thank my supervisor Jem Stach for his support and encouragement during these four years. I would also like to acknowledge Joe Gray for his patience and kindness in the LC-MS, Monica Gómez Malmierca and Carlos Olano for teaching me the basics of a natural products's laboratory and Adriel Latorre, Javier Pascual and Manel Porcar for all their support and for helping me develop bioinformatic skills. I wish to thank Newcastle University for funding this research and for providing such a good opportunity for my personal and professional development.

For their friendship and for making my time inside and outside the lab more enjoyable I would like to thank Yanhua, Lina, Jui, Victor, Ewelina, Adnan and Alidi. For sharing their knowledge, life experiences and being always there when I needed to talk, thank you Carmen, Imen and Lorena. I would also like to acknowledge Lucía, Miri, Ferran and my Science-me a Story mates for being there and for helping me disconnect from work when I needed it.

I would not want to miss the opportunity of thanking all the researchers that gifted me their time, knowledge and support before I embarked in this PhD, Cristina and Manel, Małgorzata and Paweł, Tristan, Magali and Andy, you all inspired me to pursue this journey, that would have been much harder if I had not learnt by your side before.

Finally, none of this work would have been possible without the support of my family. I will be eternally grateful to my parents Soledad and Agustín, for their unconditional support, for giving me all the opportunities I asked for since I was a child, for being open minded and for educating me to be a good person. It also makes me extremely happy to have lived this PhD experience in Newcastle together with my sister, thanks for listening to me, for your support and for everything you have taught me, I have enjoyed very much spending time with you, Pep and Ruso. Last, I would like to acknowledge the person that has been by my side each day during these four years (and before). Jonathan, thanks for your support, for listening to me, for celebrating, working long hours and crying with me and specially for making me keep going on all those days in which nothing worked. Thanks for giving me stability, for taking care of me and for making me feel valuable and self-confident.

Contributions to this work

Chapter 2. Alba Iglesias, Adriel Latorre-Pérez, James E. M. Stach, Manuel Porcar, Javier Pascual.

AI, AL-P, JS, MP and JP conceived and designed this study; AI and AL-P performed the analyses; AI, AL-P and JP and analysed the data; AI wrote this chapter; MP, JP and JS reviewed this chapter.

Chapter 3. Alba Iglesias, Monica Gómez Malmierca, Maria del Carmen Montero-Calasanz, Bernhard Kepplinger, Jonathan Tellechea-Luzardo, Carlos Olano, Thomas Howard & James E. M. Stach.

AI, TH and JS conceived and designed this study; AI performed the experiments presented in this chapter; MG-M and BK provided guidance on organic extractions and analytical chemistry; MG-M and CO provided guidance on gene expression analysis; MCM-C provided Nanopore sequencing training; JS provided bioinformatic training; JTL helped carrying out the codon usage analysis; AI wrote this chapter; JS reviewed this chapter.

Chapter 4. Alba Iglesias, Jonathan Tellechea-Luzardo, Yanhua Zhong, Joseph Gray & James E. M. Stach.

AI, JTL, JG and JS conceived and designed this study; AI performed the experiments presented in this chapter except the abyssomicin calibrated standard curve that was carried out by YZ; AI wrote this chapter; JS reviewed this chapter.

Chapter 5. Alba Iglesias, Jonathan Tellechea-Luzardo, Alex Laverick, Thomas Howard & James E. M. Stach.

AI and JS conceived and designed this study; AI performed the experiments presented in this chapter; JTL helped carrying out the electroporations; AL, TH and JS provided guidance for DOE; AI wrote this chapter; JS reviewed this chapter.

Publications from this work

Iglesias, A., Latorre-Pérez, A., Stach, J. E. M., Porcar, M. & Pascual, J. (2020). Out of the abyss: Genome and metagenome mining reveals unexpected environmental distribution of abyssomicins. *Front Microbiol* **11**, 645.

Congress attendance

I am very grateful to Newcastle University, the Spanish Microbiology Society and the Federation of European Microbiological Societies for supporting this work and my professional development as a researcher. Thanks to their economical support I was able to present the work carried out during this thesis at: Designer Biology (Newcastle), the 8th Congress of European Microbiologists - FEMS2019 (Glasgow), the XXVII Spanish Congress of Microbiology, (Málaga), the Chemistry and Biology of Natural Products Symposium XIII (University of Warwick), the John Innes/Rudjer Bošković Summer School in Applied Molecular Microbiology (Dubrovnik), the Chemistry and Biology of Natural Products Symposium XII, (University of Warwick) and the Synthetic & Systems Biology Summer School (University of Cambridge).

Contents

List of Figures

List of Tables

List of Supplementary Material

The file “Supplementary material 2.1.pdf” contains supplementary data for Chapter 2.

The file “Supplementary material 2.2.xlsx” contains supplementary data for Chapter 2.

The file “Supplementary material 3.pdf” contains supplementary data for Chapter 3.

The file “Supplementary material 4.pdf” contains supplementary data for Chapter 4.

The file “Supplementary material 5.pdf” contains supplementary data for Chapter 5.

List of abbreviations

aby BGC	*M. maris* AB-18-032 abyssomicin biosynthetic gene cluster
abm BGC	*S. koyangensis* SCSIO 5802 abyssomicin biosynthetic gene cluster
ACP	Acyl carrier protein domain
ADC	4-amino-4-deoxychorismate
ADCL	4-amino-4-deoxychorismate lyase
ADCS	Aminodeoxychorismate synthase
AMR	Antimicrobial resistance
AT	Acyltransferase domain
BGC	Biosynthetic gene cluster
CDM	Chemically Defined Medium
CRISPR	Clustered regularly interspaced short palindromic repeat
crRNA	CRISPR RNA
DH	Dehydratase domain
DOE	Design of experiments
DSB	Double-strand break
DSDs	Definitive Screening Designs
ER	Enoylreductase domain
GAT	Glutamine amidotransferase
GAT-ADCS	Amidotransferase–aminodeoxychorismate synthase
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
KR	Ketoreductase domain
KS	Ketosynthase domain
LB	Luria Bertani
LC	Liquid chromatography
MH	Mueller Hinton
MM	Minimal medium
MOA	Mechanism of action
MRSA	Methicillin-resistant *S. aureus*
MS	Mass spectrometry
NA	Nutrient agar
NB	Nutrient Broth
NRPS	Non-ribosomal peptide synthetase
PABA	4-aminobenzoic acid
PKS	Polyketide synthase
RHB	Rhizobia Helper Bacteria
rif	Rifampin-resistant
RSM	Response surface methodology
sGFP	Superfolder GFP
str	Streptomycin-resistant
tracrRNA	Trans-activating crRNA
TSA	Tryptic Soy Agar
TSB	Tryptic Soy Broth
VRSA	Vancomycin-resistant *S. aureus*
ZOI	Zone of inhibition

Chapter 1. Introduction to this work.

1.1 This thesis in a nutshell.

The rise in antimicrobial resistance (AMR) is currently one of the biggest threats to global health. During the last decades, the misuse of antibiotics, among other causes, has led to the accelerated development of drug-resistant microorganisms. This, coupled to the slow discovery rate of new antibiotics, is a real concern that leaves society in a situation worryingly similar to that of the pre-antibiotic age. Fortunately, today, researchers have more information than ever at their disposal and benefit from the quick developments in technology and infrastructure to generate knowledge at a level quantitatively and qualitatively without precedents. In this dissertation, I present three different synthetic biology-based approaches for the discovery of new antimicrobial compounds. The main aim of this work is to set up the grounds for the discovery of new abyssomicins with potentially novel bioactivities by (1) exploring the diversity of abyssomicin BGCs present in nature and analysing *in vivo* the abyssomicin production potential of some of the strains containing these BGCs, (2) expressing the *M. maris* abyssomicin biosynthetic gene cluster (*aby* BGC) in a heterologous host and (3) developing molecular microbiology and genome editing tools to engineer abyssomicin biosynthesis in *M. maris*.

This dissertation tries to answer the following research questions: (RQ1) How is the abyssomicin BGC distributed in nature and how does it evolve? (RQ2) Could genomic and metagenomic data be used to identify microbial strains with a natural ability to produce abyssomicins? (RQ3) Would it be possible to express the *aby* BGC from *M. maris* in a heterologous host? (RQ4) Could abyssomicin production be enhanced in *M. maris* so that this strain could be used as a production chassis? (RQ5) Which molecular biology protocols would need to be optimised so that *M. maris* could be engineered to produce new abyssomicins?

To solve these questions, first, I investigated the environmental distribution and evolution of the abyssomicin BGC through the analysis of publicly available genomic and metagenomic data and then, I analysed under laboratory conditions the abyssomicin production potential of five bacterial strains harbouring putative abyssomicin BGCs. Then, I focused on using the *M. maris* abyssomicin BGC for the heterologous production of these compounds. For that, first, I

carried out a comprehensive literature review to identify which chassis are generally used for the expression of bacterial polyketide natural products. Subsequently, the *aby* BGC was transferred into *Escherichia coli* and various *Streptomyces* species that were then evaluated for their abyssomicin production potential under different culturing conditions. Finally, in order to use *M. maris* as an engineerable chassis for the production of abyssomicins, ribosome engineering was applied to increase the abyssomicin production yield in this organism and, later, statistical design of experiments (DOE) was used to develop a quick and easy electroporation method to transform it.

With this work, I hope to have paved the way for future research work looking at the discovery and production of new abyssomicins with potentially novel bioactivities.

1.2 Importance.

The discovery, development and introduction into the clinic of the first antibiotics is still today one of the most relevant landmarks of modern medicine as it is estimated that, for example, just in the USA, the deaths caused by infectious diseases decreased from 797 deaths per 100,000 persons in 1900 to 45.6 deaths per 100,000 persons in 2014 (Armstrong *et al.*, 1999; Hansen *et al.*, 2016). Although more importance is given to this topic nowadays, AMR to clinically used antibiotics has been an issue since these started to be used in human and veterinary medicine. During the golden age of antimicrobial drug discovery, this emerging thread was covered up with the discovery of new clinically relevant antimicrobials (Figure 1.1). However, during the last few decades, the absence of new antibiotics and the spread of AMR has limited the treatment options against life-threatening pathogens (Powers, 2004).

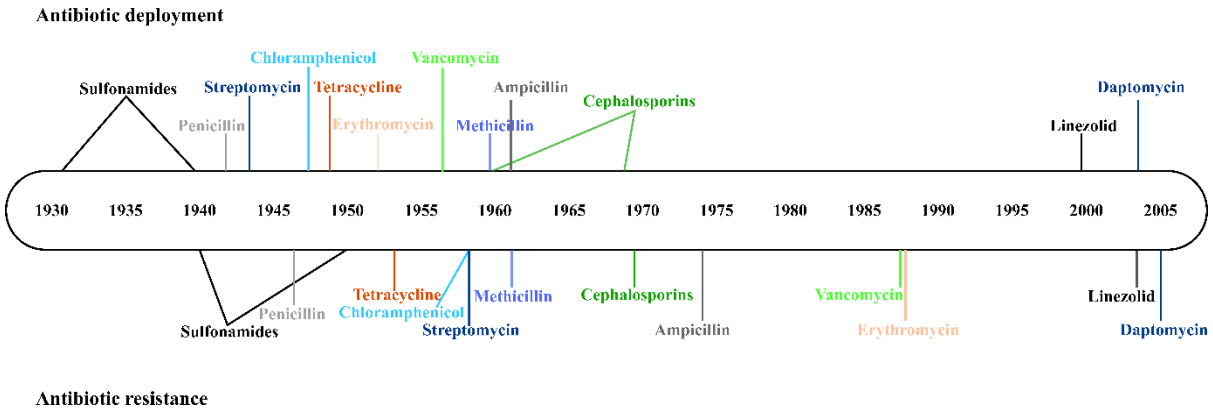

Figure 1.1. Timeline of antibiotic deployment (above the timeline) and the appearance of antibiotic resistance (below the timeline). Adapted from (Clatworthy *et al.*, 2007).

Research in the field of antimicrobial drug discovery faces challenges such as the high rediscovery rate of already known compounds, the 15-20 years process until drug approval and the loss of interest of pharma companies on the exploitation of this kind of drugs. Despite this, it is necessary to keep looking for and developing antibiotics, as for example, the global pandemic of resistant *Staphylococcus aureus* alone already kills more people in the USA each year than HIV/AIDS, Parkinson's disease, emphysema, and homicide combined (Ventola, 2015). Indeed, in the USA alone, antibiotic-resistant pathogen-associated hospital-acquired infections cause 99,000 deaths annually and economic losses of about $35 billion annually were recorded due to antibiotic resistance in health care systems (Aslam *et al.*, 2018).

1.3 Motivation.

Many efforts are currently being done in different research fields to tackle AMR from as many perspectives as possible, from physics researchers generating models on how the complex network of reactions that constitutes bacterial physiology responds to the antibiotic-induced stress (Allen & Waclaw, 2016) to medical researchers developing rapid point-of-care resistance diagnostics tools (McAdams *et al.*, 2019). Synthetic biologists, on their side, have found on the modularity of biosynthetic pathways responsible for the production of natural products the perfect opportunity to generate novel nature-inspired biomolecules with antimicrobial activity (Zakeri & Lu, 2013).

1.3.1 Microbial secondary metabolites vs chemically synthesised molecules.

Plant, fungi and bacteria constitute a great source of secondary metabolites with antimicrobial activity. Over 16,500 antibiotics have been isolated as natural products from these organisms and interestingly over half of them came from environmental bacteria from the phylum *Actinobacteria* (Zakeri & Lu, 2013). During years, researchers assumed that these compounds with antimicrobial activity were involved in inter-microbial competition for nutrients or space, but recent studies on the natural roles of secondary metabolites showed that these have dual roles and display unsuspected activities at low concentrations (Davies, 2013; van Bergeijk *et al.*, 2020). Natural products are synthesised through large secondary metabolic pathways, a subset of which, is characterised by an inherent modularity that makes them attractive platforms for synthetic biology (Khosla *et al.*, 2009). That is the case of polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS) systems. These pathways are organised

in the genomes in the form of biosynthetic gene clusters (BGCs) comprised of large enzymatic complexes that allow for the coordinated action of several enzymes to build complex small molecules from basic building blocks in a similar way to car assembly lines. At a chemistry level, previous research revealed that microbial secondary metabolites contain a much larger fraction of sp^3-hybridised bridgehead atoms and chiral centres than synthetic small molecules (Henkel *et al.*, 1999; Rodrigues *et al.*, 2016). In general, natural products also present a lower nitrogen content, a higher oxygen presence and more aliphatic rings than aromatic ones (Rodrigues *et al.*, 2016). Whilst the traditional antimicrobial screenings of hundreds of isolates are still one of the technically easiest strategies followed in microbiology laboratories for the discovery of new molecules with antimicrobial activity, limitations such as the high rediscovery rate have led to the development of alternative approaches.

The chemical synthesis of urea in 1828 by Friedrich Wöhler is considered the achievement that launched the total synthesis and synthetic organic chemistry fields. It was then when organic chemists started using simple building blocks to build complex molecular architectures, although the purpose of this research field has been continuously evolving to justify the chemical synthesis of natural products (Hong, 2014). At the beginning, the chemical synthesis of a natural product was the last step to confirm the structure of the compound. Then, in the 20^{th} century, X-ray crystallography and spectroscopic methods took over this task, and thus, organic chemistry was relegated to create and discover new reactions and patterns of chemical reactivity. Since then, the natural product synthesis field has focused on synthesising complex molecules at scalable amounts to supply enough material for clinical studies to be carried out. With the rise of AMR, work in this field has also centred on expanding the chemical diversity of already known antibiotics through the chemical synthesis of natural product-inspired compounds (Charest *et al.*, 2005). For that, different approaches have been followed including fragment-based drug discovery, the screening of synthetic chemical libraries for antimicrobial activity, computer-assisted *de novo* design of small molecules that mimic nature's chemistry and combinatorial chemistry (Rodrigues *et al.*, 2016).

These approaches, however, have not been very successful, as synthetic chemical libraries have been unable to mimic the structural complexity of bioactive natural products (Wang *et al.*, 2019). Natural products and marketed drugs are structurally more diverse than combinatorial chemistry products (Figure 1.2). This is mostly a consequence of trying to make combinatorial synthesis more efficient and, in general, the differences with higher impact include the number

of chiral centres, the presence of aromatic rings and other complex ring systems, the degree of the saturation and the presence of different heteroatoms (Feher & Schmidt, 2003).

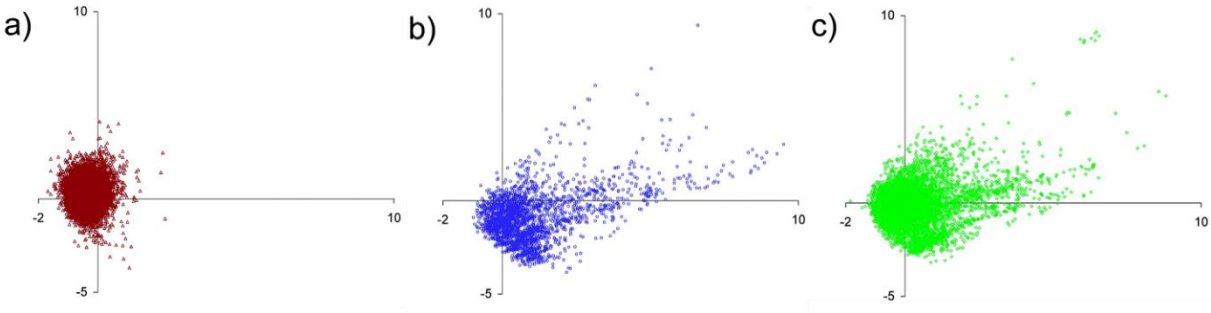

Figure 1.2. Product space (2D) covered by **A)** combinatorial compounds (n=13,506), **B)** natural products (n=3,287) and **C)** marketed drugs (n=10,968). Adapted from (Feher & Schmidt, 2003).

1.3.2 Exploitation of non-streptomycete natural products.

One could think that the chemical diversity of natural products is limited, and that chemical synthesis is necessary for the creation of derivatives with optimised bioactivities. It is well-known that more than 99% of the predicted 10^{11}–10^{12} microbial species on Earth remain undiscovered and only a tiny fraction of these can be cultured in laboratories (Locey & Lennon, 2016; Bodor *et al.*, 2020). The sequencing revolution unveiled the presence of dozens of BGCs in single bacterial genomes and a plethora in metagenomes, that, independently of whether *a priori* these were silent or even vestiges of ancient BGCs, could be studied, activated, restored and combined with parts of each other to create an infinite number of molecules.

It is evidently easier to start studying those BGCs that are active under laboratory conditions and whose natural products already display interesting bioactivities. *Streptomyces* is the best studied genus among *Actinobacteria*, with approximately 900 species described and over 1,620 genomes deposited in the NCBI, although only 189 of these and 35 assemblies would be complete (Lee *et al.*, 2020). Each *Streptomyces* genome possesses between 25–70 BGCs (Belknap *et al.*, 2020) and are natural producers of drugs such as the anticancers daunorubicin (Otten *et al.*, 1995), dactinomycin (Waksman & Gregory, 1954) and mitomycin C (Yingqing *et al.*, 1999), the antimicrobials streptomycin (Waksman *et al.*, 1946), chloramphenicol (Akagawa *et al.*, 1975) and tetracycline (McCormick *et al.*, 1958), the antifungals nystatin (Sekurova *et al.*, 1999), amphotericin B (Linke *et al.*, 1974) and natamycin (El-Enshasy *et al.*,

2000), the immunosuppressants sirolimus (Sehgal, 2003), ascomycin (Arai *et al*., 1962) and tacrolimus (Barreiro *et al*., 2012) and the antiparasitic ivermectin (Campbell *et al*., 1983).

Apart from the genus *Streptomyces*, other genera have also proved to be very valuable sources of natural products with clinical applications. That is the case of for example the genera *Micromonospora*, *Salinispora* and *Frankia*, that are producers of compounds such as gentamicin (Weinstein *et al*., 1963), diazepinomicin (Charan *et al*., 2004), crisamicin (Nelson *et al*., 1986), thiocoraline (Romero *et al*., 1997), salinisporamycin (Matsuda *et al*., 2009), sporolides (Buchanan *et al*., 2005), arenimycin (Asolkar *et al*., 2010), arenicolides (Williams *et al*., 2007), cyanosporasides (Oh *et al*., 2006) and frankiamicin (Ogasawara *et al*., 2015) among many others. The exploitation of secondary metabolites from non-streptomycete *Actinobacteria* tend to be undervalued due to the inherent difficulties to culture some of the strains in terms of time and nutritional requirements, as well as for the absence of well-developed tools to manipulate them and the lack of closely related heterologous hosts. Lately, efforts have focused on easing these limitations and non-streptomycete chassis such as *Salinispora tropica* CNB-4401 have been developed for the heterologous expression of secondary metabolic pathways (Zhang *et al.*, 2018). Some genome engineering tools have also been recently adapted to suit non-streptomycete *Actinobacteria* such as *Verrucosispora* sp. (Wang *et al.*, 2020), however, there is still a need for standardised reproducible protocols.

1.4 Background.

1.4.1 An overview on the abyssomicin family of natural products.

The abyssomicins is an actively expanding family of spirotetronate polyketide natural products. Abyssomicins B, C and D, were the first compounds discovered from this family in 2004 during the screening of a library of hundreds of *Actinobacteria* looking for antimicrobial compounds with anti-folate activity (Riedlinger *et al.*, 2004). From these, only abyssomicin C displayed antibiotic activity, however, with the elucidation of these first abyssomicin structures, a hypothesis describing the potential structure-activity relationship of the molecules arose (Bister *et al.*, 2004). Several different chemical approaches were followed to achieve, first, the partial synthesis of the core structure of abyssomicin C (Rath *et al.*, 2005; Snider & Zou, 2005), and finally, the total synthesis of abyssomicin C and structural analogues including

atrop-abyssomicin C (Zapf *et al.*, 2005; Couladouros *et al.*, 2006; Peters & Fischer, 2006; Nicolaou & Harrison, 2006, 2007).

While synthetic chemistry studies were producing loads of valuable data and new bioactive and non-bioactive structures, abyssomicin E was isolated from the fermentation broth of *Streptomyces* sp. HKI0381 (Niu *et al.*, 2007) and abyssomicins G, H and atrop-abyssomicin C were discovered from the first described abyssomicin producer, *Verrucosispora maris* AB-18-032^T (Keller *et al.*, 2007a), later reclassified as *Micromonospora maris* AB-18-032^T (Nouioui *et al.*, 2018). The continuous feed of different abyssomicin structures together with the study of their bioactivities allowed to completely elucidate the molecular mechanism by which abyssomicin C and its atrop- isomer interfere in the biosynthesis of 4-aminobenzoic acid (PABA) (Keller *et al.*, 2007b). Indeed, it was found not only that abyssomicin C and atrop-abyssomicin C were potent antibiotics against Gram positive bacteria, including methicillin (MRSA) and vancomycin-resistant *S. aureus* (VRSA), but also strong *in vitro* inhibitors of *Mycobacterium tuberculosis* (Freundlich *et al.*, 2010).

Abyssomicin I from *Streptomyces* sp. CHI39, recently classified as *Streptomyces abyssomicinicus* CHI39^T (Komaki *et al.*, 2019), and ent-Homoabyssomicins A and B, from *Streptomyces* sp. Ank 210 were the next abyssomicins to be found in nature (Igarashi *et al.*, 2010; Abdalla *et al.*, 2011). Meanwhile, *M. maris* AB-18-032 genome was sequenced and abyssomicin C biosynthesis and atrop-abyssomicin C chemical synthesis totally elucidated (Gottardi *et al.*, 2011; Roh *et al.*, 2011; Bihelovic & Saicic, 2012; Goodfellow *et al.*, 2012). A bit later, a large scale culture of the new isolate *Verrucosispora* sp. MS100128, yielded three new abyssomicins, abyssomicins J, K, and L, and the four already known abyssomicins B, C, D, and H (Wang *et al.*, 2013). Abyssomicin B was later found to be produced by *Verrucosispora* sp. MS100047 too (Wang *et al.*, 2013; Huang *et al.*, 2016). It was the discovery of three new analogues, abyssomicins 3, 4 and 5, produced by the isolate *Streptomyces* sp. RLUS1487, and the characterisation of abyssomicin 2 as reactivator of latent HIV virus what highlighted the relevance of the chemical diversity specifically within this family of secondary metabolites (León *et al.*, 2015) (Table 1.1).

After that, a study published in 2016, demonstrated for the first time the biocatalytic mechanism of the natural Diels-Alderase AbyU present in the *aby* BGC, closing the gap around the intramolecular [4+2] cycloaddition reaction necessary to build the polycyclic abyssomicin

skeleton (Byrne *et al.*, 2016). Some years later the same team reported the thermodynamic and kinetic stability of this enzyme, that is capable of catalysing the [4+2] cycloaddition at elevated temperatures and in the presence of organic solvents and the chemical denaturant guanidinium hydrochloride (Marsh *et al*., 2019). Meanwhile, the unusual architecture of atrop-abyssomicin C kept attracting the attention of synthetic chemists ten years after its discovery. This time, instead of using a Diels-Alder reaction for the formation of the polycyclic framework, which was the approach followed before for the total synthesis of these compounds, researchers followed a different strategy and developed a synthesis system that made available not only the targeted compound but also structural analogues complementary to those obtainable by the Diels-Alder approach (Bihelovic *et al*., 2017).

During 2017, 20 new abyssomicins were discovered. Highlighting the first reported example of a bacterial strain capable of producing a broad array of abyssomicins, the coal mine fire isolate *Streptomyces* sp. LC-6-2 was found to produce 12 new bi- and tetracyclic abyssomicins (M-X) (Wang *et al.*, 2017). This expanded structural diversity of the abyssomicins culminated with the discovery of neoabyssomicins A-C from the deep-sea derived *Streptomyces koyangensis* SCSIO 5802, capable of producing abyssomicins 2 and 4 too (Song *et al.*, 2017; Tu *et al.*, 2018). Further work with that same strain led to the discovery of three new abyssomicin monomers (neoabyssomicins D, E and A2) and the two dimeric neoabyssomicins F and G (Table 1.1) (Huang *et al.*, 2018). At this time, various papers were published on the biochemistry of some of the enzymes involved in abyssomicin biosynthesis. First, researchers demonstrated *in vivo* that AbmV, a cytochrome P450 enzyme, is responsible for the domino reactions involved in the bridged ether formation and the hydroxylation of the C_{11} during neoabyssomicin/abyssomicin biosynthesis in *S. koyangensis* SCSIO 5802 (Li *et al*., 2018). Next, it was demonstrated that AbyA5 catalyses the stereospecific acetate elimination that gives the (R)-tetronate acetate as biosynthetic intermediate (Lees *et al*., 2019) and a study on the structural truncation of abyssomicin C showed that the bicyclic motif is essential for activity against MRSA (Monjas *et al*., 2019). More recently, abyssomicin Y was discovered in fermentation extracts of the marine *Verrucosispora* sp. MS100137 (Zhang *et al*., 2020) and lastly, the pair of enzymes formed by the luciferase-like monooxygenase AbmE2 and its flavin reductase AbmZ, was characterised as a type II Baeyer–Villiger monooxygenase that catalyses a Baeyer–Villiger oxidation during neoabyssomicin biosynthesis in *S. koyangensis* SCSIO 5802 (Ji *et al*., 2020).

Table 1.1. Abyssomicin producer strains known to date and compounds produced by each of them.

Microorganism	Abyssomicin structure	Reference
Micromonospora maris AB-18-032	B, C, atrop-C, D, G, H	(Riedlinger *et al.*, 2004)
Streptomyces sp. HKI0381	E	(Niu *et al.*, 2007)
Streptomyces abyssomicinicus CHI39	I	(Igarashi *et al.*, 2010) (Komaki *et al.*, 2019)
Streptomyces sp. Ank 210	ent-homoA, ent-homoB	(Abdalla *et al.*, 2011)
Verrucosispora sp. MS100128	B, C, atrop-C, D, H, J, K, L	(Wang *et al.*, 2013)
Streptomyces sp. RLUS1487	2, 3, 4, 5	(León *et al.*, 2015)
Verrucosispora sp. MS100047	B,C	(Huang *et al.*, 2016)
Streptomyces sp. LC-6-2	M, N, O, P, Q, R, S, T, U, V, W, X	(Wang *et al.*, 2017)
Streptomyces koyangensis SCSIO 5802	2, 4, neo-A, neo-B, neo-C, neo-D, neo-E, neo-F, neo-G, neo-A2	(Song *et al.*, 2017)
Verrucosispora sp. MS100137	Y, C, D, L	(Zhang *et al.*, 2020)

Although only ten strains so far have been reported to produce abyssomicins, the 38 known abyssomicins for which structures have been elucidated prove that this family of natural products presents a wide structural diversity (Figure 1.3). Indeed, the abyssomicins are classified as type I or type II abyssomicins depending on their structure. The type I family includes abyssomicins B–E, G, H, J–L and atrop-abyssomicin C, and the enantiomeric counterparts of the type I compounds are considered type II abyssomicins (Sadaka *et al.*, 2018). Furthermore, type II abyssomicins can be categorised by the degree of methylation and the presence of an oxygen atom within the polyketide backbone so that type IIA abyssomicins have methyl substitutions at C_4 and C_{12}, type IIB have one methyl substitution at C_{12}, and type IIC have one methyl substitution at C_{12} and an oxygen atom in the macrocycle (Sadaka *et al.*, 2018). This structural diversity is responsible for different clinically relevant biological activities. Atrop-abyssomicin C and abyssomicins C, 2 and J display antibiotic activity against Gram positive bacteria, including MRSA, VRSA and different *Mycobacteria* strains (Sadaka *et al.*, 2018). Abyssomicin 2 also possesses HIV inhibitory and reactivator properties and neoabyssomicins A and C promote HIV-1 replication in a human lymphocyte model (León *et al.*, 2015; Song *et al.*, 2017). More recently, the inhibitory effects against influenza A virus of abyssomicins Y, D, L and H was reported (Zhang *et al.*, 2020).

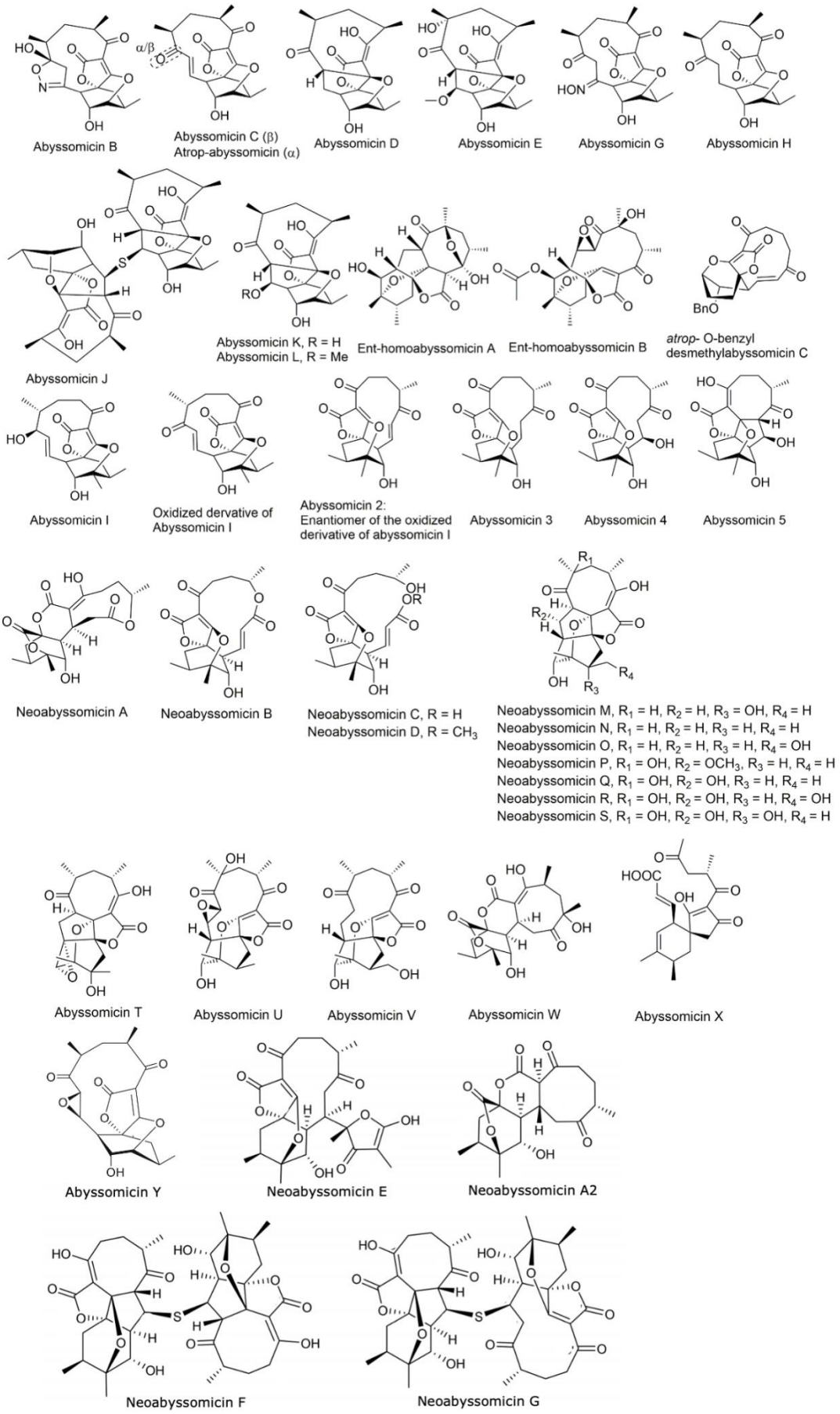

Figure 1.3. Abyssomicin structures discovered to date. Adapted from (Sadaka *et al*., 2018).

1.4.2 Abyssomicin biosynthesis in M. maris.

As described before, *M. maris* AB-18-032 was the first microbial strain reported to produce abyssomicins. Previous work unveiled the biosynthesis of this small but structurally complex spirotetronate natural product, that is as follows. The genome of *M. maris* AB-18-032 harbours a complete abyssomicin BGC that comprises 25 genes, distributed along 56 kb (Figure 1.4A). It is a PKS cluster with three central genes (*abyB1, abyB2* and *abyB3*) encoding for a seven-module type I PKS. The first of these genes, *abyB1*, is made up of four modules each with a ketosynthase (KS), acyltransferase (AT), acyl carrier protein (ACP), dehydratase (DH) and ketoreductase (KR) domain except for the first module that lacks the DH and KR domains. The second PKS gene, *abyB2*, consists of two modules, the first one has active KS, AT and ACP domains, but also displays inactive DH and KR domains. The active site of the DH domain displays the residues Leu–Thr instead of the His–Asp catalytic dyad (Akey *et al*., 2010) while the active site triad Lys–Ser–Tyr required for KR activity is instead Ile–Ser–Val (Haydock *et al*., 1995). The second module of *abyB2* displays KS, AT, ACP, DH and KR domains plus an enoylreductase (ER) domain between the DH and KR domains. The last PKS gene, *abyB3*, makes up for only one module with a KS, AT and ACP domain (Figure 1.4B) (Gottardi *et al*., 2011).

In a bit more detail, the synthesis of the abyssomicin scaffold starts with the incorporation of an acetate molecule by the AT domain of module 1 into the ACP domain of module 1. The KS domain in module 1 acts, in this case, as a loading module for the formation of the acetate starter unit by catalysing the decarboxylation of an ACP-tethered malonate. The ACP of module 1 translocates the precursor chain into the KS domain of module 2. The AT domain selects another acetate precursor as extender unit and incorporates it into the ACP domain, where an elongation reaction or C-C bond is formed between the chain from the downstream module and the precursor. Then various chain modification events can occur. In this case, the KR reduces the oxygen from the acetate molecule incorporated in module 1 to a hydroxyl group that is finally lost as a water molecule by the dehydratase activity of the DH domain that results in the formation of a C-C double bond. After that, the ACP from module 2 translocates the chain generated in module 2 into the KS domain of module 3 through a thioester linkage and the AT from module 3 selects and incorporates into the ACP domain from module 3 another acetate molecule. Modules 3 and 4 are functionally identical to module 2 and elongate the abyssomicin backbone through the incorporation of two more acetates. After that, the ACP

domain of module 4 translocates that C_8 precursor into the KS domain of module 5. The AT domain in module 5 selects this time a propionate unit that is incorporated into the ACP domain of module 5 and then the chain is translocated by the ACP of module 5 into the KS domain of module 6. The AT domain in module 6 selects for a second propionate molecule that is incorporated into the ACP domain of module 6 and the KS domain catalises the decarboxylation reaction between the propionate in the ACP domain and the chain synthesised by the preceding modules. Then, the KR domain reduces the oxygen from the propionate molecule incorporated in module 5 to a hydroxyl group that is finally lost as a water molecule by the dehydratase activity of the DH domain and that results in the formation of a C-C double bond. This double bond is finally reduced by the ER domain in module 6 and the chain is translocated by the ACP domain of module 6 into the KS domain of the terminator module 7. At this point the AT domain of module 7 ends the elongation of the chain introducing a last acetate molecule into the ACP domain of module 7. Finally, the KS domain in module 7 catalises the union of the C_{14} chain and the acetate in the ACP domain (Figure 1.4B).

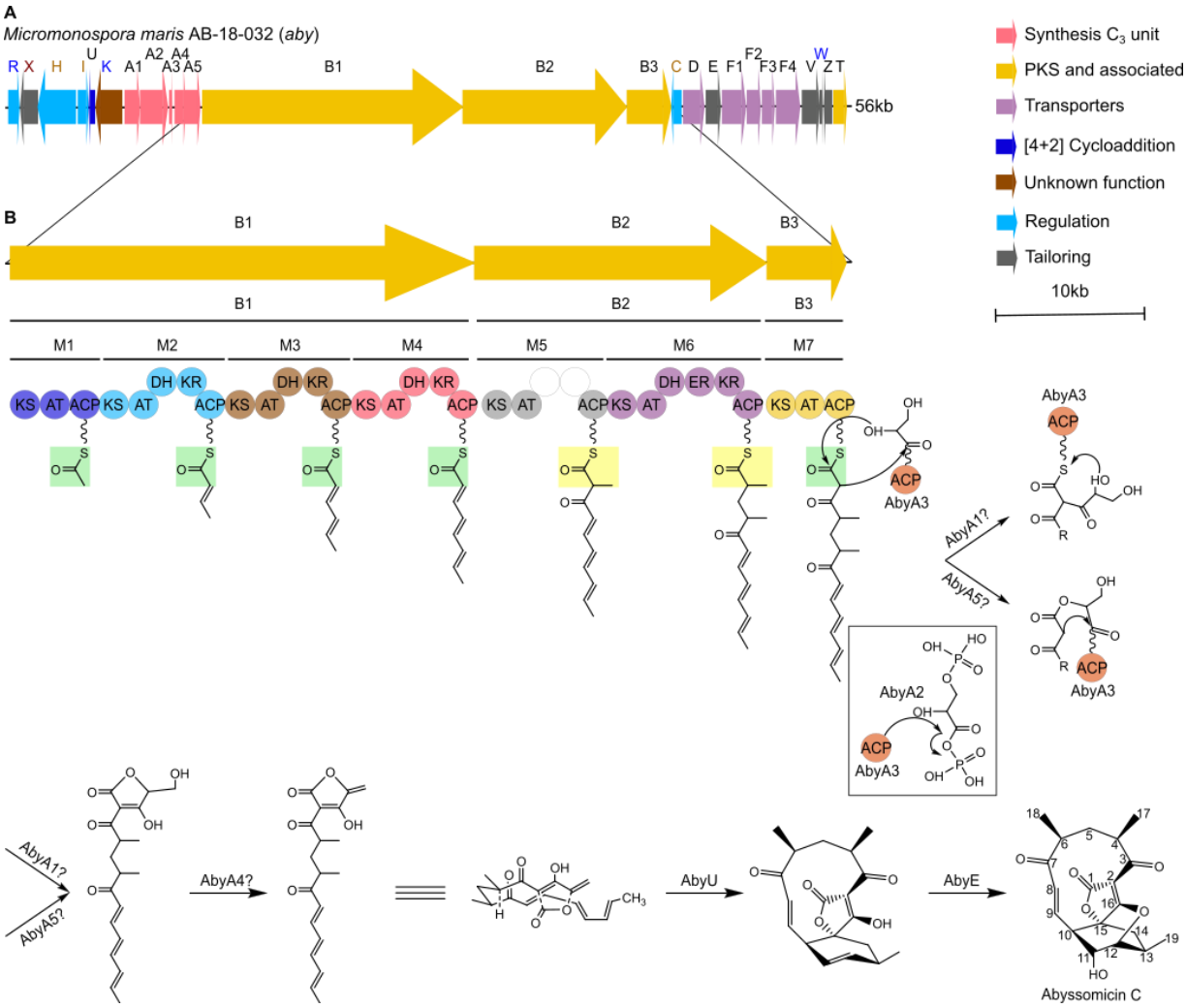

Figure 1.4. A) *M. maris* AB-18-032 abyssomicin BGC. **B)** Abyssomicin biosynthesis in *M. maris*. Adapted from (Gottardi *et al.*, 2011). The incorporation of the five acetates into the linear abyssomicin precursor is highlighted in green and the incorporation of the two propionates is highlighted in yellow.

Upstream the PKS genes that create the abyssomicin linear precursor from two propionates and five acetates, there are five genes involved in the assembly of the tetronic acid moiety (*abyA1-A5*). AbyA2 is an acyl transferase that catalyses the attachment of a bisphosphoglycerate from the glycolytic pathway to the ACP of AbyA3. The linear polyketide chain is then transferred on AbyA3 carrying the glyceraldehyde by the β-ketoacyl-ACP-synthase III AbyA1 and the α/β-hydrolase AbyA5. AbyA4, a dehydrogenase catalytic domain-containing protein would then catalyse a dehydration step. After the formation of the tetronate moiety, the Diels-Alderase AbyU would catalyse the [4+2] cycloaddition reaction and the monooxygenase AbyE would be responsible for the formation of the epoxide cycle (Figure 1.4B). The cluster also contains four transcriptional regulators consisting of a LuxR transcriptional regulator (*abyH*), two putative pathway activators from the SARP family (*abyR*

and *abyI*) and a TetR-like regulatory protein (*abyC*), an ABC exporter system (*abyF1-F4*), a drug resistance transporter (*abyD*), a cytochrome P450 system that would be responsible for the bridged ether formation and the hydroxylation of the C_{11} (*abyV, abyW* and *abyZ*), a cytochrome P450 gene (*abyX*) and a type II thioesterase (*abyT*) (Figure 1.4A) (Gottardi *et al.*, 2011; Byrne *et al.*, 2016; Li *et al.*, 2018; Lees *et al.*, 2019).

1.5 Aims of this thesis.

In this thesis, I present three different approaches for the discovery of new antimicrobial secondary metabolites from the abyssomicin family of natural products. In a bit more detail, the aims of this dissertation are as follows.

1.5.1 Study of the diversity of abyssomicin BGCs present in nature, evolutionary analysis and identification of potential abyssomicin producers.

The first aim of this thesis is to provide with an insight of the diversity of abyssomicin BGCs present in nature. Since genomes started to be sequenced, thousands of microbial genomes and metagenomes have been sequenced and data made available. Through the exploration of this reservoir of raw microbial data, it would be possible to study how abyssomicin producers are distributed across a wide range of natural and artificial environments on Earth and how abyssomicin BGCs evolve. Moreover, this approach allows the identification of potential abyssomicin producing strains that can be studied under laboratory conditions for their ability to produce compounds from this family. In order to carry out this work, I choose to mine publicly-available genomes and metagenomes using the three abyssomicin Diels-Alderases (AbyU/AbsU/AbmU) known to date that catalyse the intramolecular [4+2] cycloaddition reaction of the linear abyssomicin polyketide precursor. Then, the abyssomicin production potential of five bacterial strains harbouring potential abyssomicin BGCs was analysed under laboratory conditions. Finally, the evolutionary relationships between the newly discovered potential abyssomicin BGCs was studied.

1.5.2 Heterologous expression of the **M. maris** *abyssomicin biosynthetic gene cluster.*

While searching for new abyssomicin producers in nature, and with the ultimate goal of producing novel abyssomicins with bioactivities of potential medical interest, a second

research line focused on the exploitation of the *M. maris* abyssomicin BGC started. For that, it was found necessary to express the *M. maris aby* BGC in a well-established heterologous host. This approach provides several advantages, as these chassis tend to be well-characterised and there are effective genome editing tools to engineer the host and achieve production yields that allow the compounds to be further exploited. In this thesis, I present my attempt to express the *M. maris* AB-18-032 *aby* BGC in various heterologous hosts. That includes a systematic analysis of the hosts that have been used to express polyketide synthase (PKS) clusters from bacteria between 2006-2016, the transformation of *E. coli* BAP1 and various *Streptomyces* hosts, the application of two different approaches, ribosome engineering and clustered regularly interspaced short palindromic repeat (CRISPR-Cas9) mediated promoter knock-in, to activate abyssomicin production and whole genome sequencing and a gene expression analysis.

1.5.3 Development of molecular microbiology tools to engineer abyssomicin biosynthesis in M. maris.

Finally, a third approach to exploit abyssomicin biosynthesis would be through the genetic manipulation of the natural producer, *M. maris* AB-18-032. The antimicrobial atrop-abyssomicin C is the main fermentation product of this strain, however, its production yield allows for improvement and various basic molecular tools would need to be developed for it to be engineerable. In this thesis, ribosome engineering was used to generate *M. maris* drug-resistant mutants with enhanced abyssomicin C production and statistical DOE was used to create an electroporation method that allows quick and easy introduction of foreign DNA into this strain. Moreover, in order to facilitate abyssomicin C detection via antimicrobial plate diffusion assay, the influence that media composition and extracellular pH exert on the sensitivity of test strains to antimicrobial compounds was analysed. Preliminary work on the application of CRISPR-Cas9 systems for genome editing *M. maris* was also carried out.

1.6 Research questions and methodology.

This section describes briefly the methods used when looking for an answer to each of the research questions of this thesis.

1.6.1 (RQ1) How is the abyssomicin BGC distributed in nature and how does it evolve?

In order to study how microorganisms harbouring abyssomicin BGCs are distributed in nature and how these BGC have evolved, I used the sequences of the natural Diels-Alderases present in the three abyssomicin BGCs described to date to mine the genomes present in NCBI database plus over 3,000 randomly-selected metagenomes available in the JGI metagenomes database. For the Diels-Alderase containing genomes, putative abyssomicin BGCs were identified using antiSMASH, and in those cases in which the Diels-Alderase homolog was not part of an abyssomicin BGC, the surrounding genomic region was reconstructed manually. For the evolutionary analysis of the isolates harbouring a Diels-Alderase homolog, phylogenetic and phylogenomic trees were created using MAFFT and UBCG respectively and a synteny analysis was carried out manually.

1.6.2 (RQ2) Could genomic and metagenomic data be used to identify microbial strains with a natural ability to produce abyssomicins?

Following the identification of 40 new BGCs potentially involved in the biosynthesis of abyssomicins from genomic data, I randomly selected and analysed the abyssomicin production potential of five bacterial strains harbouring potential abyssomicin BGCs. These strains were cultured in various solid and liquid media and the antimicrobial activity of the cultures was assessed via antimicrobial plate assay from plugs (solid media) and extracts (liquid media). The extracts from liquid media were also analysed by LC-MS.

1.6.3 (RQ3) Would it be possible to express the **aby** *BGC in a heterologous host?*

In order to express the *M. maris aby* BGC in a heterologous host, I first carried out a review of the literature published during the ten years prior to the start of this work focusing on finding which microbial strains had been used more successfully for the production of polyketide natural products from bacteria. Then, via bi- and tri-parental mating *E. coli* BAP1 and various *Streptomyces* strains were transformed with a phage artificial chromosome harbouring the *aby* BGC. Fermentations in liquid and solid media, antimicrobial plate assays and organic extractions and analytical chemistry analysis were carried out to determine abyssomicin production. Ribosome engineering and CRISPR-Cas9 mediated promoter knock-in were

applied to activate abyssomicin production in the heterologous hosts and gene expression and genome sequencing analysis were carried out using standard molecular laboratory practices.

1.6.4 (RQ4) Could abyssomicin production be enhanced in **M. maris** *so that this strain could be used as a production chassis?*

To improve abyssomicin C production in *M. maris* for future scaling-up projects or simply to facilitate its detection in routinely production assays the ribosome engineering methodology was followed. It makes use of basic microbiology techniques to isolate and screen drug-resistant mutants obtained from culturing a strain under sub-lethal concentrations of ribosomal targeting antibiotics.

1.6.5 (RQ5) Which molecular biology protocols would need to be optimised so that **M. maris** *could be engineered to produce new abyssomicins?*

Being able to produce novel abyssomicins in *M. maris* would require mainly two basic molecular microbiology methods to be optimised for this strain. First, an efficient transformation method and second a quick and effective genome engineering tool. In this dissertation, I present the use of statistical DOE for the development of an electroporation protocol tailored for *M. maris* and some preliminary work to evaluate the use of CRISPR-Cas9 technology for genome editing *M. maris.*

1.7 Main findings of this study.

During the course of this thesis, the environmental distribution and evolution of the abyssomicin BGC was investigated through the analysis of publicly available genomic and metagenomic data. The strategy of selecting a pathway-specific enzyme to direct the mining proved to be an excellent strategy, as 74 new Diels–Alderase homologs were identified and a surprising prevalence of the abyssomicin BGC within soil and plant-associated environments was unveiled. Five complete and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs were also identified, suggesting that a plethora of abyssomicins are still to be discovered.

In parallel, the expression of the *aby* BGC in various heterologous hosts was pursued. I was able to move the *aby* BGC into *E. coli* BAP1 and various *Streptomyces* species. I evaluated the abyssomicin production potential of these strains under different conditions and some of the hosts were promoter engineered to force the expression of the *aby* BGC. I demonstrated active gene expression, however, none of the heterologous hosts produced abyssomicins. Re-sequencing the genome of one of the heterologous hosts revealed the presence of several mutations at four different points within *abyB1*, the first polyketide synthase gene in the *aby* BGC, suggesting this could be the reason for the lack of production.

Finally, I focused on increasing abyssomicin production in *M. maris* AB-18-032 and developing genetic tools for this system. First, through ribosome engineering, a library of *M. maris* drug-resistant mutants capable of producing up to 3.4-fold abyssomicin C in comparison to the wild-type strain was generated. Using statistical DOE, an efficient electroporation protocol that could accelerate targeted genetic manipulations in *M. maris* was developed. Together, increased abyssomicin production and a quick and easy electroporation protocol for *M. maris*, will facilitate future engineering of the *aby* BGC directly in *M. maris* to produce diverse non-natural abyssomicins.

1.8 Thesis outline.

1.8.1 Chapter 1.

In this chapter (current), I introduce the work presented in this dissertation. It briefly describes the importance, motivation and aims of the thesis as well as the background, research questions and methodology used to answer them. It also provides an overview of the main findings.

1.8.2 Chapter 2.

Chapter 2 presents an approach to study the environmental distribution and evolution of the abyssomicin BGC based on the use of a pathway-specific enzyme to mine publicly available genomic and metagenomic data.

1.8.3 Chapter 3.

Chapter 3 presents our efforts to express the *M. maris* AB-18-032 *aby* BGC in various heterologous hosts. It also includes the use of ribosome engineering and CRISPR-Cas9 mediated promoter knock-in to drive abyssomicin production.

1.8.4 Chapter 4.

In chapter 4 the application of ribosome engineering to generate *M. maris* mutants with enhanced abyssomicin C production is presented. This chapter also includes a study on how extracellular pH influences antimicrobial plate diffusion assays.

1.8.5 Chapter 5.

In chapter 5, the reader will find how a statistical DOE approach was followed to set up an electroporation method to introduce foreign DNA into *M. maris*. This chapter also includes some preliminary work on the application of CRISPR-Cas9 technology to engineer *M. maris*.

1.8.6 Chapter 6.

This last chapter summarises the conclusions derived from the main findings of this thesis.

1.9 References.

Abdalla, M. A., Yadav, P. P., Dittrich, B., Schüffler, A. & Laatsch, H. (2011). Ent-Homoabyssomicins A and B, two new spirotetronate metabolites from *Streptomyces* sp. Ank 210. *Org Lett* **13**, 2156–2159.

Akagawa, H., Okanishi, M. & Umezawa, H. (1975). A plasmid involved in chloramphenicol production in *Streptomyces venezuelae*: evidence from genetic mapping. *J Gen Microbiol* **90**, 336–346.

Akey, D. L., Razelun, J. R., Tehranisa, J., Sherman, D. H., Gerwick, W. H. & Smith, J. L. (2010). Crystal structures of dehydratase domains from the curacin polyketide biosynthetic pathway. *Structure* **18**, 94–105.

Allen, R. & Waclaw, B. (2016). Antibiotic resistance: A physicist's view. *Phys Biol* **13**.

Arai, T., Kouama, Y., Suenaga, T. & Honda, H. (1962). Ascomycin, an antifungal antibiotic. *J Antibiot* **15**, 231–232.

Armstrong, G. L., Conn, L. A. & Pinner, R. W. (1999). Trends in infectious disease mortality in the United States during the 20th century. *J Am Med Assoc* **281**, 61–66.

Aslam, B., Wang, W., Arshad, M. I., Khurshid, M., Muzammil, S., Rasool, M. H., Nisar, M. A., Alvi, R. F., Aslam, M. A. & other authors. (2018). Antibiotic resistance: a rundown of a global crisis. *Infect Drug Resist* **11**, 1645–1658.

Asolkar, R. N., Kirkland, T. N., Jensen, P. R. & Fenical, W. (2010). Arenimycin, an antibiotic effective against rifampin- and methicillin-resistant *Staphylococcus aureus* from the marine actinomycete *Salinispora arenicola*. *J Antibiot* **63**, 37–39.

Barreiro, C., Prieto, C., Sola-Landa, A., Solera, E., Martínez-Castro, M., Pérez-Redondo, R., García-Estrada, C., Aparicio, J. F., Fernández-Martínez, L. T. & other authors. (2012). Draft genome of *Streptomyces tsukubaensis* NRRL 18488, the producer of the clinically important immunosuppressant tacrolimus (FK506). *J Bacteriol (194-)* **14**, 3756-3757.

Belknap, K. C., Park, C. J., Barth, B. M. & Andam, C. P. (2020). Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep* **10**.

van Bergeijk, D. A., Terlouw, B. R., Medema, M. H. & van Wezel, G. P. (2020). Ecology and genomics of *Actinobacteria*: new concepts for natural product discovery. *Nat Rev Microbiol* 1–13.

Bihelovic, F. & Saicic, R. N. (2012). Total synthesis of (-)-atrop-abyssomicin C. *Angew Chemie - Int Ed* **51**, 5687–5691.

Bihelovic, F., Vulovic, B. & Saicic, R. N. (2017). Synthesis of natural products and the development of synthetic methodology: The case study of (-)-atrop-abyssomicin C. *Nat Prod Commun*.

Bister, B., Bischoff, D., Ströbele, M., Riedlinger, J., Reicke, A., Wolter, F., Bull, A. T., Zähner, H., Fiedler, H. P. & Süssmuth, R. D. (2004). Abyssomicin C - A polycyclic antibiotic from a marine *Verrucosispora* strain as an inhibitor of the p-aminobenzoic acid/tetrahydrofolate biosynthesis pathway. *Angew Chemie - Int Ed* **43**, 2574–2576.

Bodor, A., Bounedjoum, N., Vincze, G. E., Erdeiné Kis, Á., Laczi, K., Bende, G., Szilágyi, Á., Kovács, T., Perei, K. & Rákhely, G. (2020). Challenges of unculturable bacteria: environmental perspectives. *Rev Environ Sci Biotechnol* **19**, 1–22.

Buchanan, G. O., Williams, P. G., Feling, R. H., Kauffman, C. A., Jensen, P. R. & Fenical, W. (2005). Sporolides A and B: Structurally unprecedented halogenated macrolides from the marine actinomycete *Salinispora tropica*. *Org Lett* **7**, 2731–2734.

Byrne, M. J., Lees, N. R., Han, L.-C., van der Kamp, M. W., Mulholland, A. J., Stach, J. E. M., Willis, C. L. & Race, P. R. (2016). The catalytic mechanism of a natural Diels–Alderase revealed in molecular detail. *J Am Chem Soc* **138**, 6095–6098.

Campbell, W. C., Fisher, M. H., Stapley, E. O., Albers-Schönberg, G. & Jacob, T. A. (1983). Ivermectin: A potent new antiparasitic agent. *Science (80-)* **221**, 823–828.

Charest, M. G., Lerner, C. D., Brubaker, J. D., Siegel, D. R. & Myers, A. G. (2005). A convergent enantioselective route to structurally diverse 6-deoxytetracycline antibiotics. *Science* **308**, 395–398.

Charan, R. D., Schlingmann, G., Janso, J., Bernan, V., Feng, X. & Carter, G. T. (2004). Diazepinomicin, a new antimicrobial alkaloid from a marine *Micromonospora* sp. *J Nat Prod* **67**, 1431–1433.

Clatworthy, A. E., Pierson, E. & Hung, D. T. (2007). Targeting virulence: A new paradigm for antimicrobial therapy. *Nat Chem Biol* **3**, 541–548.

Couladouros, E. A., Bouzas, E. A. & Magos, A. D. (2006). Formal synthesis of abyssomicin C. *Tetrahedron* **62**, 5272–5279.

Davies, J. (2013). Specialized microbial metabolites: Functions and origins. *J Antibiot (Tokyo)* **66**, 361–364.

El-Enshasy, H. A., Farid, M. A. & El-Sayed, E. S. A. (2000). Influence of inoculum type and cultivation conditions on natamycin production by *Streptomyces natalensis*. *J Basic Microbiol* **40**, 333–342.

Feher, M. & Schmidt, J. M. (2003). Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* **43**, 218–227.

Freundlich, J. S., Lalgondar, M., Wei, J. R., Swanson, S., Sorensen, E. J., Rubin, E. J. & Sacchettini, J. C. (2010). The abyssomicin C family as *in vitro* inhibitors of *Mycobacterium tuberculosis*. *Tuberculosis* **90**, 298–300.

Goodfellow, M., Stach, J. E. M., Brown, R., Bonda, A. N. V., Jones, A. L., Mexson, J., Fiedler, H.-P., Zucchi, T. D. & Bull, A. T. (2012). *Verrucosispora maris* sp. nov., a novel deep-sea actinomycete isolated from a marine sediment which produces abyssomicins. *Antonie Van Leeuwenhoek* **101**, 185–193.

Gottardi, E. M., Krawczyk, J. M., Von Suchodoletz, H., Schadt, S., Mühlenweg, A., Uguru, G. C., Pelzer, S., Fiedler, H. P., Bibb, M. J. & other authors. **(2011).** Abyssomicin biosynthesis: Formation of an unusual polyketide, antibiotic-feeding studies and genetic analysis. *ChemBioChem* **12**, 1401–1410.

Hansen, V., Oren, E., Dennis, L. K. & Brown, H. E. (2016). Infectious disease mortality trends in the United States, 1980-2014. *J Am Med Assoc* **20**, 2149-2151.

Haydock, S. F., Aparicio, J. F., Molnár, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F. A., Galloway, I. S., Staunton, J. & Leadlay, P. F. (1995). Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett* **374**, 246–248.

Henkel, T., Brunne, R. M., Müller, H. & Reichel, F. (1999). Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chemie Int Ed* **38**, 643–647.

Hong, J. (2014). Natural product synthesis at the interface of chemistry and biology. *Chem - A Eur J* **20**, 10204–10212.

Huang, H., Song, Y., Li, X., Wang, X., Ling, C., Qin, X., Zhou, Z., Li, Q., Wei, X. & Ju, J. (2018). Abyssomicin monomers and dimers from the marine-derived *Streptomyces koyangensis* SCSIO 5802. *J Nat Prod* **81**, 1892–1898.

Huang, P., Xie, F., Ren, B., Wang, Q., Wang, J., Wang, Q., Abdel-Mageed, W. M., Liu, M., Han, J. & other authors. **(2016).** Anti-MRSA and anti-TB metabolites from marine-derived *Verrucosispora* sp. MS100047. *Appl Microbiol Biotechnol* **100**, 7437–7447.

Igarashi, Y., Yu, L., Miyanaga, S., Fukuda, T., Saitoh, N., Sakurai, H., Saiki, I., Alonso-Vega, P. & Trujillo, M. E. (2010). Abyssomicin I, a modified polycyclic polyketide from *Streptomyces* sp. CHI39. *J Nat Prod* **73**, 1943–1946.

Ji, X., Tu, J., Song, Y., Zhang, C., Wang, L., Li, Q. & Ju, J. (2020). A luciferase-like monooxygenase and flavin reductase pair AbmE2/AbmZ catalyzes Baeyer-Villiger oxidation in neoabyssomicin biosynthesis. *ACS Catal* **10**, 2591–2595.

Keller, S., Nicholson, G., Drahl, C., Sorensen, E., Fiedler, H.-P. & Süssmuth, R. D. (2007a). Abyssomicins G and H and atrop-abyssomicin C from the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **60**, 391–394.

Keller, S., Schadt, H. S., Ortel, I. & Süssmuth, R. D. (2007b). Action of atrop-abyssomicin C as an inhibitor of 4-amino-4-deoxychorismate synthase PabB. *Angew Chemie - Int Ed* **46**, 8284–8286.

Khosla, C., Kapur, S. & Cane, D. E. (2009). Revisiting the modularity of modular polyketide synthases. *Curr Opin Chem Biol* **2**, 135–143.

Komaki, H., Sakurai, K., Hosoyama, A., Kimura, A., Trujilo, M. E., Igarashi, Y. & Tamura, T. (2019). Diversity of PKS and NRPS gene clusters between *Streptomyces abyssomicinicus* sp. nov. and its taxonomic neighbor. *J Antibiot (Tokyo)*.

Lee, N., Kim, W., Hwang, S., Lee, Y., Cho, S., Palsson, B. & Cho, B. K. (2020). Thirty complete *Streptomyces* genome sequences for mining novel secondary metabolite biosynthetic gene clusters. *Sci Data* **7**, 1–9.

Lees, N. R., Han, L.-C., Byrne, M. J., Davies, J. A., Parnell, A. E., Moreland, P. E. J., Stach, J. E. M., van der Kamp, M. W., Willis, C. L. & Race, P. R. (2019). An esterase-like lyase catalyzes acetate elimination in spirotetronate/spirotetramate biosynthesis. *Angew Chemie* **131**, 2327–2331.

León, B., Navarro, G., Dickey, B. J., Stepan, G., Tsai, A., Jones, G. S., Morales, M. E., Barnes, T., Ahmadyar, S. & other authors. **(2015).** Abyssomicin 2 reactivates latent HIV-1 by a PKC- and HDAC-independent mechanism. *Org Lett* **17**, 262–265.

Li, Q., Ding, W., Yao, Z., Tu, J., Wang, L., Huang, H., Li, S. & Ju, J. (2018). AbmV catalyzes tandem ether installation and hydroxylation during neoabyssomicin/abyssomicin biosynthesis. *Org Lett* **20**, 4854–4857.

Linke, H. A. B., Mechlinski, W. & Schaffner, C. P. (1974). Production of amphotericin B-14C by *Streptomyces nodosus* fermentation, and preparation of the amphotericin B-14C-methyl-ester. *J Antibiot* **27**, 155–160.

Locey, K. J. & Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc Natl Acad Sci U S A* **113**, 5970–5975.

Matsuda, S., Adachi, K., Matsuo, Y., Nukina, M. & Shizuri, Y. (2009). Salinisporamycin, a novel metabolite from *Salinispora arenicora*. *J Antibiot* **62**, 519–526.

Marsh, C. O., Lees, N. R., Han, L., Byrne, M. J., Mbatha, S. Z., Maschio, L., Pagden-Ratcliffe, S., Duke, P. W., Stach, J. E. M. & other authors. **(2019).** A natural Diels-Alder biocatalyst enables efficient [4+2] cycloaddition under harsh reaction conditions. *ChemCatChem* **11**, 5027–5031.

McAdams, D., Waldetoft, K. W., Tedijanto, C., Lipsitch, M. & Brown, S. P. (2019). Resistance diagnostics as a public health tool to combat antibiotic resistance: a model-based evaluation. *PLoS Biol* **17**.

McCormick, J. R. D., Sjolander, N. O., Miller, P. A., Hirsch, U., Arnold, N. H. & Doerschuk, A. P. (1958). The biological reduction of 7-chloro-5a(11a)-dehydrotetracycline to 7-chloro-tetracycline by *Streptomyces aureofaciens*. *J Am Chem Soc* **80**, 6460–6461.

Monjas, L., Fodran, P., Kollback, J., Cassani, C., Olsson, T., Genheden, M., Larsson, D. G. J. & Wallentin, C.-J. (2019). Synthesis and biological evaluation of truncated derivatives of abyssomicin C as antibacterial agents. *Beilstein J Org Chem* **15**, 1468–1474.

Nelson, R. A., Pope Jr., J. A., Luedemann, G. M., MCDaniel, L. E. & Schaffner, C. P. (1986). Crisamicin A, a new antibiotic from *Micromonospora*. I. Taxonomy of the producing strain, fermentation, isolation, physico-chemical characterization and antimicrobial properties. *J Antibiot* **39**, 335–344.

Nicolaou, K. C. & Harrison, S. T. (2007). Total synthesis of abyssomicin C, atrop-abyssomicin C, and abyssomicin D: Implications for natural origins of atrop-abyssomicin C. *J Am Chem Soc* **129**, 429–440.

Nicolaou, K. C. & Harrison, S. T. (2006). Total synthesis of abyssomicin C and atrop-abyssomicin C**. *Synthesis (Stuttg)* 3256–3260.

Niu, X. M., Li, S. H., Görls, H., Schollmeyer, D., Hilliger, M., Grabley, S. & Sattler, I. (2007). Abyssomicin E, a highly functionalized polycyclic metabolite from *Streptomyces* species. *Org Lett* **9**, 2437–2440.

Nouioui, I., Carro, L., García-López, M., Meier-Kolthoff, J. P., Woyke, T., Kyrpides, N. C., Pukall, R., Klenk, H.-P., Goodfellow, M. & Göker, M. (2018). Genome-based taxonomic classification of the phylum *Actinobacteria*. *Front Microbiol* **9**.

Ogasawara, Y., Yackley, B. J., Greenberg, J. A., Rogelj, S. & Melançon, C. E. (2015). Expanding our understanding of sequence-function relationships of type II polyketide biosynthetic gene clusters: Bioinformatics-guided identification of Frankiamicin A from *Frankia* sp. EAN1pec. *PLoS One* **10**, e0121505.

Oh, D. C., Williams, P. G., Kauffman, C. A., Jensen, P. R. & Fenical, W. (2006). Cyanosporasides A and B, chloro- and cyano-cyclopenta[a]indene glycosides from the marine actinomycete '*Salinispora pacifica*'. *Org Lett* **8**, 1021–1024.

Otten, S. L., Ferguson, J. & Hutchinson, C. R. (1995). Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR2* locus. *J Bacteriol* **177**, 1216–1224.

Peters, R. & Fischer, D. F. (2006). Total syntheses of the antibacterial natural product abyssomicin C. *Angew Chemie - Int Ed* **45**, 5736–5739.

Powers, J. H. (2004). Antimicrobial drug development – the past, the present, and the future. *Clin Microbiol Infect* **10**, 23–31.

Rath, J. P., Kinast, S. & Maier, M. E. (2005). Synthesis of the fully functionalized core structure of the antibiotic abyssomicin C. *Org Lett* **7**, 3089–3092.

Riedlinger, J., Reicke, A., Zähner, H., Krismer, B., Bull, A. T., Maldonado, L. A., Ward, A. C., Goodfellow, M., Bister, B. & other authors. **(2004).** Abyssomicins, inhibitors of the para-aminobenzoic acid pathway produced by the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **57**, 271–279.

Rodrigues, T., Reker, D., Schneider, P. & Schneider, G. (2016). Counting on natural products for drug design. *Nat Chem* **8**, 531–541.

Romero, F., Espliego, F., Baz, J. P., Quesada, T. G. De, Grávalos, D., Calle, F. D. La & Fernández-Puentes, J. L. (1997). Thiocoraline, a new depsipeptide with antitumor activity produced by a marine *Micromonospora*. I. Taxonomy, fermentation, isolation, and biological activities. *J Antibiot* **50**, 734–737.

Roh, H., Uguru, G. C., Ko, H. J., Kim, S., Kim, B. Y., Goodfellow, M., Bull, A. T., Kim, K. H., Bibb, M. J. & other authors. **(2011).** Genome sequence of the abyssomicin- and proximicin-producing marine actinomycete *Verrucosispora maris* AB-18-032. *J Bacteriol* **193**, 3391–3392.

Sadaka, C., Ellsworth, E., Robert Hansen, P., Ewin, R., Damborg, P. & Watts, J. L. (2018). Review on abyssomicins: Inhibitors of the chorismate pathway and folate biosynthesis. *Molecules* **23**.

Sehgal, S. N. (2003). Sirolimus: Its discovery, biological properties, and mechanism of action. *In Transplant Proc* **35**, S7–S14.

Sekurova, O., Sletta, H., Ellingsen, T. E., Valla, S. & Zotchev, S. (1999). Molecular cloning and analysis of a pleiotropic regulatory gene locus from the nystatin producer *Streptomyces noursei* ATCC11455. *FEMS Microbiol Lett* **177**, 297–304.

Snider, B. B. & Zou, Y. (2005). Synthesis of the carbocyclic skeleton of abyssomicins C and D. *Org Lett* **7**, 4939–4941.

Song, Y., Li, Q., Qin, F., Sun, C., Liang, H., Wei, X., Wong, N. K., Ye, L., Zhang, Y. & other authors. **(2017).** Neoabyssomicins A–C, polycyclic macrolactones from the deep-sea derived *Streptomyces koyangensis* SCSIO 5802. *Tetrahedron* **73**, 5366–5372.

Tu, J., Li, S., Chen, J., Song, Y., Fu, S., Ju, J. & Li, Q. (2018). Characterization and heterologous expression of the neoabyssomicin/abyssomicin biosynthetic gene cluster from *Streptomyces koyangensis* SCSIO 5802. *Microb Cell Fact* **17**, 1–14.

Ventola, C. L. (2015). The antibiotic resistance crisis: causes and threats. *P T J* **40**, 277–283.

Waksman, S. A. & Gregory, F. J. (1954). Actinomycin. II. Classification of organism producing different forms of actinomycin. *Antibiot Chemother* **4**, 1050–1056.

Waksman, S. A., Reilly, H. C. & Johnstone, D. B. (1946). Isolation of streptomycin-producing strains of *Streptomyces griseus*. *J Bacteriol* **52**, 393–7.

Wang, Q., Song, F., Xiao, X., Huang, P., Li, L., Monte, A., Abdel-Mageed, W. M., Wang, J., Guo, H. & other authors. **(2013).** Abyssomicins from the South China Sea deep-sea sediment *Verrucosispora* sp.: Natural thioether Michael addition adducts as antitubercular prodrugs. *Angew Chemie - Int Ed* **52**, 1231–1234.

Wang, Q., Xie, F., Tong, Y., Habisch, R., Yang, B., Zhang, L., Müller, R. & Fu, C. (2020). Dual-function chromogenic screening-based CRISPR/Cas9 genome editing system for actinomycetes. *Appl Microbiol Biotechnol* **104**, 225–239.

Wang, S., Dong, G. & Sheng, C. (2019). Structural simplification of natural products. *Chem Rev* **6**, 4180–4220.

Wang, X., Elshahawi, S. I., Cai, W., Zhang, Y., Ponomareva, L. V., Chen, X., Copley, G. C., Hower, J. C., Zhan, C.-G. & other authors. **(2017).** Bi- and tetracyclic spirotetronates from the Coal Mine Fire isolate *Streptomyces* sp. LC-6-2. *J Nat Prod* **2**.

Weinstein, M. J., Luedemann, G. M., Oden, E. M., Wagman, G. H., Rosselet, J. P., Marquez, J. A., Coniglio, C. T., Charney, W., Herzog, H. L. & Black, J. (1963). Gentamicin, 1 a new antibiotic complex from *Micromonospora*. *J Med Chem* **6**, 463–464.

Williams, P. G., Miller, E. D., Asolkar, R. N., Jensen, P. R. & Fenical, W. (2007). Arenicolides A-C, 26-membered ring macrolides from the marine actinomycete *Salinispora arenicola*. *J Org Chem* **72**, 5025–5034.

Yingqing, M., Varoglu, M. & Sherman, D. H. (1999). Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564. *Chem Biol* **6**, 251–263.

Zakeri, B. & Lu, T. K. (2013). Synthetic biology of antimicrobial discovery. *ACS Synth Biol* **2**, 358–372.

Zapf, C. W., Harrison, B. A., Drahl, C. & Sorensen, E. J. (2005). A Diels-Alder macrocyclization enables an efficient asymmetric synthesis of the antibacterial natural product abyssomicin C. *Angew Chemie - Int Ed* **44**, 6533–6537.

Zhang, J. J., Moore, B. S. & Tang, X. (2018). Engineering *Salinispora tropica* for heterologous expression of natural product biosynthetic gene clusters. *Appl Microbiol Biotechnol* **102**, 8437–8446.

Zhang, J., Li, B., Qin, Y., Karthik, L., Zhu, G., Hou, C., Jiang, L., Liu, M., Ye, X. & other authors. **(2020).** A new abyssomicin polyketide with anti-influenza A virus activity from a marine-derived *Verrucosispora* sp. MS100137. *Appl Microbiol Biotechnol* **104**, 1533–1543.

Chapter 2. Out of the abyss: Genome and metagenome mining reveals unexpected environmental distribution of the abyssomicins.

Abstract

Natural products have traditionally been discovered through the screening of culturable microbial isolates from diverse environments. The sequencing revolution allowed the identification of dozens of biosynthetic gene clusters within single bacterial genomes, either from cultured or uncultured strains. However, researchers are still far from fully exploiting the microbial reservoir, as most of the species are non-model organisms with complex regulatory systems that can be recalcitrant to engineering approaches. Genomic and metagenomic data produced by laboratories worldwide covering the range of natural and artificial environments on Earth, are an invaluable source of raw information from which natural product biosynthesis can be accessed. The present work describes the environmental distribution and evolution of the abyssomicin BGC through the analysis of publicly available genomic and metagenomic data. The results obtained demonstrate that the selection of a pathway-specific enzyme to direct genome mining is an excellent strategy; 74 new Diels-Alderase homologs were identified and unveiled a surprising prevalence of abyssomicin BGCs within terrestrial habitats, mainly soil and plant-associated. Five complete and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs were also identified. These results strongly support the potential of genome and metagenome mining as a key preliminary tool to inform bioprospecting strategies aimed at the identification of new bioactive compounds such as – but not restricted to – abyssomicins.

2.1 Introduction.

2.1.1 Tetronate natural products.

Natural products are the main source of pharmaceutically interesting biomolecules. In particular, the search for microbial secondary metabolites has yielded a broad range of chemical structures with bioactivities, from antibiotics or antimycotics to immunosuppressants and anticancer compounds (Lacoske & Theodorakis, 2015). Among those, compounds featuring tetronate moieties are attractive due to their versatile biological activities. Most of these compounds are produced by bacteria from the phylum *Actinobacteria* and are built of a

linear fatty acid or polyketide chain with a characteristic tetronic acid 4-hydroxy-2(5H)-furanone ring system.

Within the growing family of tetronates, compounds are classified taking into account the linearity or macrocyclization of the carbon backbone and the size of the central ring system (Vieweg *et al*., 2014). Spirotetronates are tetronate compounds in which two rings are linked to each other by a spiroatom, and include, amongst many others, the abyssomicins, chlorothricins, tetrocarcins, lobophorines and quartromicins. This class of tetronates shares important biosynthetic and structural features: a conjugated pair of carbon–carbon double bonds at the end of a linear polyketide chain, a characteristic exocyclic double bond on the tetronate ring system and a Diels–Alder reaction that forms the cyclohexene moiety and an additional macrocycle (Weixin *et al*., 2013; Vieweg *et al*., 2014).

Many spirotetronates, including abyssomicin C and its atrop- isomer, kijanimicin, chlorothricin, lobophorins E and F and decatromicin B, exhibit antimicrobial properties, but in addition to that, most of these compounds show other unique bioactivities. For example, abyssomicin 2 reactivates latent HIV-1 (León *et al*., 2015), lobophorins A and B have anti-inflammatory activity (Jiang *et al*., 1999), lobophorins C and D display cytotoxic activity against human liver and breast cancer cell lines respectively (Wei *et al*., 2011), maklamicin displays moderate cytotoxicity against human cervical cancer and breast cancer cell lines (Igarashi *et al*., 2011), nomimicin has antifungal activity (Igarashi *et al*., 2012), quartromicins A1–A3 show strong activity against herpes simplex virus and quartromicins D1–D3 are active against the influenza virus (Tsunakawa *et al*., 1992).

2.1.2 The abyssomicins: biosynthesis and habitat distribution.

The abyssomicins are an actively growing family of small spirotetronate natural products with a polyketide backbone and a C_{11} central ring system that has been widely studied due to the unique structural features and bioactivities that some of its members exhibit. Abyssomicin biosynthesis occurs in a variety of hosts isolated from different ecosystems. The first abyssomicins (B-D) were discovered in 2004 during the screening of 930 actinomycetes extracts in a successful attempt to find antibacterial compounds targeting folate biosynthesis. Those abyssomicins were fermentation products of the marine actinomycete *Verrucosispora maris* AB-18-032^T, later reclassified as *Micromonospora maris* AB-18-032^T (Nouioui *et al*.,

2018), isolated from sediments of the Sea of Japan (Riedlinger *et al*., 2004). Years later, other research groups found new abyssomicins produced by soil isolates of *Streptomyces* sp. HKI0381, *Streptomyces* sp. CHI39, recently classified as *Streptomyces abyssomicinicus* CHI39^{T} (Komaki *et al*., 2019), and *Streptomyces* sp. Ank 210, in Senegal, Mexico and Germany, respectively (Niu *et al*., 2007; Igarashi *et al*., 2010; Abdalla *et al*., 2011). After that, the production of abyssomicins was again reported in marine isolates: *Verrucosispora* sp. MS100128 (Wang *et al*., 2013), *Streptomyces* sp. RLUS1487 (León *et al*., 2015) and *Verrucosispora* sp. MS100047 (Huang *et al*., 2016). Finally, the last abyssomicins found were synthesised by the soil *Streptomyces* sp. LC-6-2 (Wang *et al*., 2017) and the marine *S. koyangensis* SCSIO 5802 (Song *et al*., 2017; Huang *et al*., 2018). During the review process of this paper, abyssomicin Y was discovered in fermentation extracts of the marine *Verrucosispora* sp. MS100137 (Zhang *et al*., 2020) (Table S2.1).

Despite the limited number of bacterial strains identified so far as abyssomycin producers, this family of natural products presents a wide structural diversity. In fact, there are as many as 38 members classified as type I or type II abyssomicins, where the type I family includes abyssomicins B–E, G, H, J–L and atrop-abyssomicin C, and type II abyssomicins are the enantiomeric counterparts of the type I compounds (Sadaka *et al*., 2018). The type II abyssomicins are further categorised by the degree of methylation and the presence of an inserted oxygen atom within the polyketide backbone. Type IIA abyssomicins have methyl substitutions at C_4 and C_{12}, type IIB have one methyl substitution at C_{12}, and type IIC have one methyl substitution at C_{12} and an inserted oxygen atom in the macrocycle (Sadaka *et al*., 2018). This structural diversity has gifted this family of natural products with different clinically relevant biological activities. Atrop-abyssomicin C and abyssomicins C, 2 and J exhibit antimicrobial activity against Gram positive bacteria, including MRSA, VRSA and different *Mycobacteria* strains (Sadaka *et al*., 2018). Abyssomicin 2 also possesses HIV inhibitory and reactivator properties and neoabyssomicins A and C promote HIV-1 replication in a human lymphocyte model (León *et al*. 2015; Song *et al*., 2017). More recently, abyssomicins Y, D, L and H were described as the first abyssomicins with significant inhibitory effects against influenza A virus (Zhang *et al*., 2020) (Figure S2.1).

Previous works also elucidated the complete abyssomicin BGC present in the genome of *M. maris* AB-18-032 and proposed a model for the biosynthesis of atrop-abyssomicin C, the atrop-isomer of abyssomicin C and main product synthesised by *M. maris* AB-18-032 (Nicolaou &

Harrison, 2006, 2007; Keller *et al*., 2007; Gottardi *et al*., 2011). This abyssomicin biosynthetic gene cluster (*aby*) is made up of 25 genes, distributed along 56 kb in *M. maris* AB-18-032 genome. The cluster consists of several transcriptional regulators (*abyR, abyH, abyI* and *abyC*), an ABC exporter system (*abyF1-F4*), a drug resistance transporter (*abyD*), a cytochrome P450 system (*abyV, abyW* and *abyZ*), a cytochrome P450 gene (*abyX*), a monooxygenase (*abyE*), a type II thioesterase (*abyT*), a Diels-Alderase (*abyU*), the PKS I genes (*abyB1, abyB2* and *abyB3*) and five genes (*abyA1-A5*) involved in the assembly of the tetronic acid moiety (Figure 2.1A and Table S2.10) (Gottardi *et al*., 2011; Byrne *et al*., 2016). The partially sequenced cluster of the isolate *Verrucosispora* sp. MS100047 is 99% similar to the *aby* BGC (Figure 2.1B and Table S2.13).

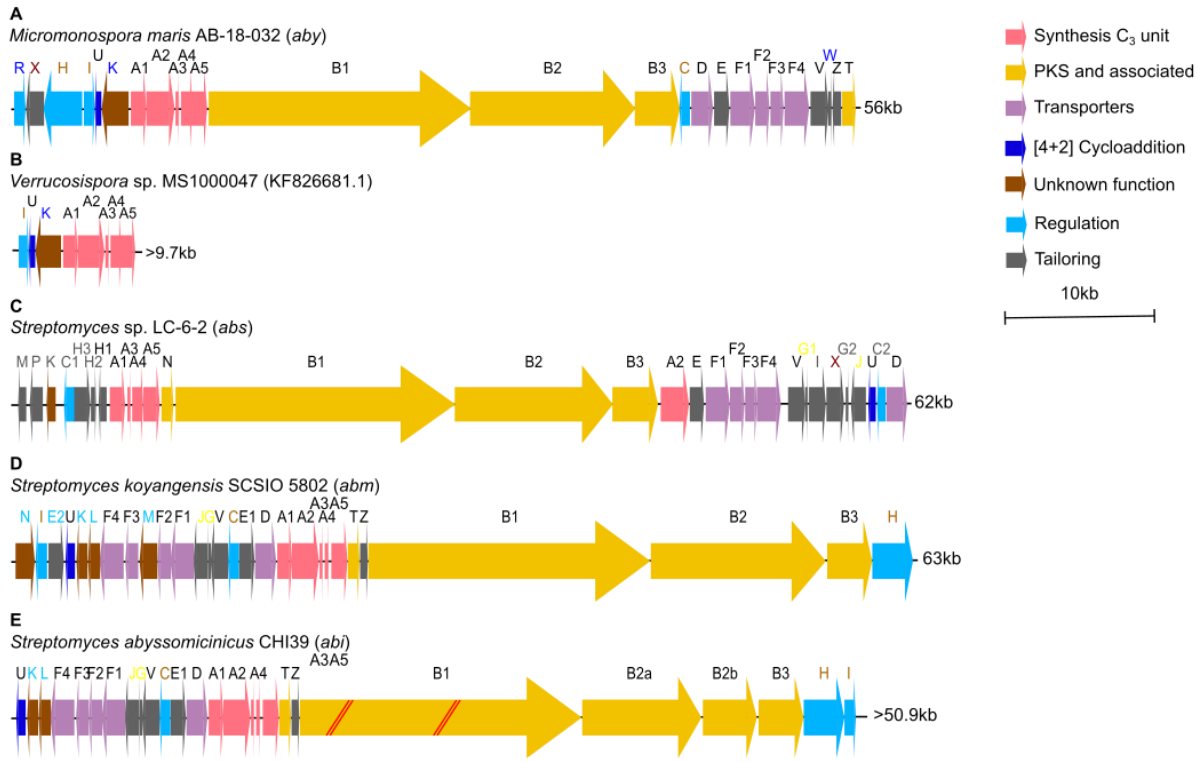

Figure 2.1. A) Abyssomicin BGC (*aby*) of *M. maris* AB-18-032. **B)** Partial abyssomicin BGC of *Verrucosispora* sp. MS1000047. **C)** Abyssomicin BGC (*abs*) of *Streptomyces* sp. LC-6-2. **D)** Neoabyssomicin/abyssomicin BGC (*abm*) of *S. koyangensis* SCSIO 5802. **E)** Abyssomicin BGC (*abi*) of *S. abyssomicinicus*. Gene names in black are common to *aby*, *abs* and *abm* BGCs. Blue font represents genes present only in *M. maris* AB-18-032, grey font represents genes present only in *Streptomyces* sp. LC-6-2 and light blue font represent genes unique to *S. koyangensis* SCSIO 5802. In maroon font appear those genes that appear both in the *aby* and *abs* BGCs, in light brown those genes that appear both in the *aby* and *abm* BGCs and in yellow those genes that appear both in the *abs* and *abm* BGCs.

After that, the discovery of abyssomicins M-X as fermentation products of *Streptomyces* sp. LC-6-2 led to the description of a new abyssomicin BGC (*abs*). This cluster consists of 30 genes disposed along 62 kb and presents homologs to most of the genes within the *aby* BGC (Table S2.12) but displays also two unique regulators (*absC1* and *absC2*) and a set of four new tailoring genes (*absG1, absG2, absI* and *absJ*) (Wang *et al*., 2017) (Figure 2.1C). A third cluster responsible for neoabyssomicin/abyssomicin biosynthesis (*abm*) was identified in *S. koyangensis* SCSIO 5802. Composed of 28 genes distributed along 63 kb, it presents five genes (*abmK, abmL, abmM, abmN* and *abmE2*) with no apparent homologous counterparts in the *aby* cluster and two more (*abmJ* and *abmG*) that appear to be in the *abs* BGC but not in the *aby* BGC (Figure 2.1D and Table S2.11) (Tu *et al*., 2018). The latest abyssomicin BGC (*abi*) was found in *S. abyssomicinicus* CHI39 and is almost identical to the *abm* BGC (Figure 2.1E and Table S2.14) (Komaki *et al*., 2019).

The environmental diversity of the abyssomicin-producing isolates suggests that abyssomicin biosynthesis could be ubiquitously distributed in nature, and bioprospecting could focus on those environments heavily colonised by *Actinobacteria* of the genus *Micromonospora* and *Streptomyces*. There are few studies concerning the driving forces behind the transmissibility and evolution of BGCs (Chevrette *et al*., 2020). In the abyssomicin family, the chemical diversity found is likely to have arisen after horizontal transfer of the abyssomicin gene cluster into new hosts with subsequent domain swapping and point mutations (Ridley *et al*., 2008). As domain swapping is predicted to occur both within and between BGCs (Jenke-Kodama *et al*., 2006), the host background (genomic context) will influence structural diversification. Moreover, the enzymes involved in the synthesis of the tetronate (AbyA1-A5) and the spiro-tetronate-forming Diels-Alderase (AbyU) are capable of accepting structurally diverse substrates (Ye *et al*., 2014; Grabarczyk *et al*., 2015; Abugrain *et al*., 2017). Thus, identification of abyssomicin BGCs in different genomic contexts is a reasonable strategy to identify structurally novel abyssomicins.

In the present work, in order to investigate the environmental colonisation of abyssomicin-producing bacteria as well as the structural diversity of abyssomicin BGCs, the distribution of the abyssomicin BGC was systematically explored, as well as its evolution, through the analysis of publicly available genomic and metagenomic data, targeting the Diels-Alderase

(AbyU/AbsU/AbmU) that catalyses the intramolecular [4 + 2] cycloaddition reaction of the linear abyssomicin polyketide precursor.

2.2 Materials and methods.

2.2.1 Preliminary metagenome mining.

Two solar panel metagenomes (available at https://www.mg-rast.org/, accession numbers: 4629146.3 and 4629147.3; accessed January/February 2019; (Meyer *et al*., 2008)) were analysed using Genome Workbench (https://www.ncbi.nlm.nih.gov/tools/gbench/) for the presence of the *M. maris* AB-18-032 abyssomicin BGC (Figure S2.2). In a first attempt, the metagenomes were set as database and a BLASTx of the *aby* BGC against the database was carried out. In a second attempt, only the coding sequences of the *aby* BGC were used in a BLASTp against the metagenomes database. In a third strategy, a BLASTp of each single protein within the *aby* BGC against the metagenomes database was performed. E-value < 10^{-6} was set as threshold to accept/reject the alignments. Unless the contrary is stated, default parameters were applied to the analyses.

2.2.2 Diels-Alderase directed metagenome mining.

A total of 3,027 metagenomes available in the JGI metagenomes database (IMG; https://img.jgi.doe.gov/; accessed February-April 2019) were analysed for AbyU/AbsU/AbmU homologs presence using the site option to carry out BLASTp (default parameters). E-value < 10^{-6} was set as threshold to accept/reject the alignments and thus identify putative Diels-Alderases. The sequences of AbyU (*Micromonospora maris* AB-18-032), AbsU (*Streptomyces* sp. LC-6-2) and AbmU (*Streptomyces koyangensis* SCSIO 5802) used as query can be found in Table S2.3. Habitats frequently populated by *Micromonospora* and *Streptomyces* species were selected, primarily from soil and aquatic environments but also from other less common environments, including fresh-water, artificial and host-associated environments (Figure S2.3). The detailed classification of the metagenomic samples from aquatic, terrestrial, engineered and host-associated environments mined for AbyU, AbsU and AbmU can be found in Tables S2.5-S2.8. For complete details on the metagenomes analysed and the Diels-Alderase positive metagenomes please refer to Supplementary material 2.2.

In order to investigate possible taxonomic biases between the Diels-Alderase positive and negative metagenomes, the relative abundance of the domain *Bacteria* and the phylum *Actinobacteria* in 50 Diels-Alderase positive metagenomes were compared against that in 50 aquatic and 50 terrestrial Diels-Alderase negative metagenomes, randomly selected from the

3,027 pool. The taxonomic assignments of both the assembled and unassembled metagenomes' reads were carried out using the IMG/JGI site option "Phylogenetic Distribution of Genes – Distribution by BLAST percent identities" and are presented here in form of relative abundance. The Mann-Whitney U test was used to investigate whether the relative abundance of *Bacteria* and *Actinobacteria* was significatively different between the aquatic, terrestrial and Diels-Alderase positive metagenomes (Nachar, 2008). In order to investigate bias in the sequencing depth, the metagenome size (bp) of those same 50 Diels-Alderase positive metagenomes was compared against the 50 aquatic and 50 terrestrial Diels-Alderase negative metagenomes. The Mann-Whitney U test was equally applied to identify significative differences in the sequencing depth.

2.2.3 Diels-Alderase directed genome mining and identification of putative abyssomicin BGCs.

BLASTp of AbyU, AbsU and AbmU were carried out against the non-redundant protein sequences database (NCBI; accessed April 2019). E-value < 10^{-6} was set as threshold to accept/reject the alignments and thus identify putative Diels-Alderases. The Diels-Alderase containing genomes were then submitted to antiSMASH (Blin *et al*., 2019) (accessed April 2019; default parameters used) for BGC mining. The location of the Diels-Alderase homolog within the genome was used to verify BGC presence in antiSMASH. When a BGC was found by antiSMASH in the desired location, ORF, protein size and proposed annotation were collated and BLASTp of every protein was carried out against the non-redundant protein sequences database (NCBI) to obtain the closest homolog (Tables S2.15-S2.85). BLASTp was used to verify/redefine the BGCs limits established by antiSMASH. In cases where antiSMASH did not identify any BGC, reconstruction of the Diels-Alderase homolog nearby genomic region was done manually from the corresponding genome in NCBI. All the recovered BGCs were classified based on their completeness and novelty (Table 2.1).

Table 2.1. Classification of the recovered BGCs found through Diels-Alderase directed genome mining.

BGC	The Diels-Alderase homolog is
Abyssomicin, total	Part of an abyssomicin BGC and it is possible to recover the sequence and structure of the entire BGC.
Abyssomicin, partial	Part of an abyssomicin BGC that is likely to be complete but due to the sequencing technology used there are some incomplete genes, frame shifts, gaps or the cluster is on a contig edge.
Potential abyssomicin, total	Part of a BGC whose product may potentially be an abyssomicin according to antiSMASH and it is possible to recover the sequence and structure of the entire BGC.
Potential abyssomicin, partial	Part of a BGC whose product may potentially be an abyssomicin according to antiSMASH but there are some genes missing or incomplete, frame shifts, gaps or the cluster is on a contig edge.
Potential BGC, total	Surrounded by genes that could form a BGC altogether, but it is unclear which could be its product.
Potential BGC, partial	Surrounded by genes that could form a BGC altogether, but it is unclear which could be its product and there were some incomplete genes, frame shifts, gaps or the cluster was on a contig edge.
Not a BGC	Not likely to be part of any BGC.
Not enough data	In a contig whose length makes it not possible to gain any knowledge.
Quartromycin, total	In a quartromycin BGC and the sequence of the cluster is complete.
Quartromycin, partial	In a quartromycin BGC but the sequence of the corresponding PKS is incomplete.
Potential tetronomycin, total	Part of a potential tetronomycin BGC.
Potential chlorothricin, partial	In a chlorothricin BGC but the sequence of the corresponding PKS is incomplete.

2.2.4 Evolutionary analysis.

All the proteins identified through genome mining that produced significative alignments (E-value < 10^{-6}) against AbyU, AbsU and AbmU were aligned and the Neighbor-Joining algorithm was used to create a phylogenetic tree using MAFFT (https://mafft.cbrc.jp/alignment/server/phylogeny.html; accessed May 2019) (Katoh *et al*., 2002). The RefSeq annotated genomes of the microorganisms harbouring those proteins were used to create a phylogenomic tree using UBCG (default parameters) (Na *et al*., 2018). The phylogenetic and phylogenomic trees were visualised and annotated with iTOL (Letunic & Bork, 2019).

A manual synteny analysis was carried out for all the newly discovered abyssomicin and potential abyssomicin BGCs (both total and partial), which were classified accordingly as described below (Table 2.2). The presence of mobile elements within the Diels-Alderase positive mined genomes was studied using Island Viewer 4 (Bertelli *et al*., 2017).

Table 2.2. Features shared by the potential abyssomicin BGCs described in this study.

Cluster type	Shared motifs
1a	Upstream the PKS genes: Entire *abyA1-A5* operon next to/nearby *abyU* which in most cases is located next to *abyK*, *abyH* and *aby*I. Downstream the PKS genes: Synteny is maintained from *abyC* to *abyW* with punctual rearrangements.
1b	Same conserved blocks as type 1a clusters but all genes are upstream the PKS genes. Extra copy of *abyV* downstream the PKS.
2a	Upstream the PKS genes: Presence of *abyA1-A5* operon except *abyA2* followed by *abyN*. Downstream the PKS genes: Synteny is maintained from *abyA2* to *abyI* with punctual rearrangements.
2b	Same conserved blocks as type 2a clusters but all genes are upstream the PKS genes except *abyA2*.
3	Entire *abyA1-A5* operon followed by *abyT*. ABC transport system divided by the presence of *abyM*.
4	*abyZ*, *abyA3* and *abyA2* are located upstream the first PKS gene. The PKS operon harbours between *abyB1* and *abyB2* a set of genes that includes at least *abyA1* and *abyU*.
5	No synteny between themselves nor with other cluster types.

2.2.5 Preliminary fermentations.

Streptomyces regensis RRL X-5263 (DSM 40551), *Streptomyces niveus* C7655 (DSM 41786), *Streptomyces incarnatus* NRRL 8089 (Prof. Takashi Tamura), *Actinokineospora auranticolor* YU 961-1 (DSM 44650) and *Micromonospora wenchangensis* CCTCC AA 2012002 (Dr. Lorena Carro) were grown in 50 ml of A1 media and 50 ml of SGG media in 250 ml baffled Erlenmeyer flasks at 30 °C and 200 rpm. Samples consisting of 5 ml of culture were taken after 24 h, 48 h, 72 h, 96 h and 120 h from inoculation and frozen for further work.

2.2.6 Organic extractions.

Samples from the preliminary fermentations were adjusted to pH 4 using HCl, mixed 1:1 with ethyl acetate and incubated at room temperature and 250 rpm for 2 hours. Tubes were centrifuged at 4000 rpm for 15 minutes and the upper organic phase was transferred into test tubes for evaporation at 22 °C and 10 psi (TurboVap® LV, Biotage). Pellets were then resuspended in 200 µl of methanol and stored at -20 °C.

2.2.7 Extract agar plate diffusion assay.

CDM agar plates with and without 5 mM PABA were inoculated with *Bacillus subtilis* DSM 10, *Micrococcus luteus* ATCC 831 and *Escherichia coli* DSM 1103 cultures at OD_{600}=1.00 following a 1:100 ratio. 0.6 mm (diameter) filter paper disks were cut using a paper drill and put on top of the agar plates. Frozen extracts were diluted 1:9 in ddH_2O and 10 µl were put onto the filter paper disks. Extracts of *M. maris* AB-18-032 were used as positive control, antimicrobial susceptibility disks of nalidixic acid (Oxoid™, ThermoFisher) were used as negative control to test the effect of PABA and filter disks with 10 µl of 10% methanol were used as negative control. After overnight incubation at 37 ºC, the zone of inhibition (ZOI) generated by each extract was measured.

2.2.8 Analytical chemistry.

Frozen extracts were diluted 1:9 in ddH_2O and centrifuged twice at 13200 rpm for 15 minutes. Supernatants were then analysed in a mass spectrometer ESI microTOF-Q II (Bruker Daltonics). Separation of the compounds was performed by a Raptor™ ARC-18 (2.7 µ, 150 x 2.1 mm, Restex) chromatographic column protected by a guard column. The mobile phase was ultrapure water (solvent A) and acetonitrile (solvent B), both with 0.1% formic acid (v/v). The following was the elution gradient: 0–55 min 20-100% B; 55-59 min 100% B; 59-60 min 100-20% B, 60-75 min 20% B followed by column washing and reconditioning (1 minute). The flow rate was 0.2 mL/min. The column oven was at 35 °C and the injection volume was 3 µL. The UV analyses were performed in the wavelength range of 190–600 nm, with the mass spectrometer operating in positive mode (m/z 50–2800). The identification of the compounds was based on mass spectrometry and the antiSMASH report generated from each genome.

2.2.9 Plug agar plate diffusion assay.

The antimicrobial production potential of *S. regensis*, *S. niveus*, *S. incarnatus*, *A. auranticolor* and *M. wenchangensis* was analysed by plug agar plate diffusion assay. First, the strains were grown in SGG agar, A1 agar, A1Sal* agar (Zhang *et al*., 2018), R5A agar (Kieser *et al*., 2000), SM10 agar (Malmierca *et al*., 2018) and SM17 agar (Malmierca *et al*., 2018) at 30 ºC for 15 days. CDM agar plates with and without 5 mM PABA were prepared and inoculated with *B. subtilis*, *E. coli* and *M. luteus* as described before. 0.6 cm (diameter) agar plugs from each strain grown in each production media were placed into the inoculated CDM agar plates. Plugs of *M. maris* AB-18-032 were used as positive control and antimicrobial susceptibility disks of nalidixic acid (Oxoid™, ThermoFisher) were used as negative control to test the effect of PABA. After overnight incubation at 37 ºC, the ZOI generated around each plug was measured.

2.3 Results.

2.3.1 Preliminary metagenome mining.

Recent studies unveiled the microbial diversity present on one of the most widespread man-made extreme structure: solar panels (Dorado-Morales *et al.*, 2016; Porcar *et al.*, 2018; Tanner *et al.*, 2018). In order to evaluate whether it would be possible to isolate any new abyssomicin producer from Valencia's solar panels, two solar panel microbiomes were mined for abyssomicin biosynthetic genes.

First, the metagenomes were set as database and a BLASTx of the *M. maris* AB-18-032 abyssomicin BGC against the database was carried out. That yielded 1,151 short sequences aligning at different points along the BGC, covering mostly those genes commonly present in the majority of BGCs (LuxR, SARP and TetR family transcriptional regulators, drug resistance exporters and ABC transporters). These results only confirmed the presence of microorganisms with secondary metabolite BGCs, which was expected since both microbiomes contained sequences belonging to the *Actinomycetales* order (Figure S2.2) and it is well known that actinobacterial genomes often contain dozens of BGCs (Bentley *et al.*, 2002; Ikeda *et al.*, 2003).

After that, with the aim of making the search more specific, only the coding sequences of the abyssomicin BGC were used for the BLASTp. This resulted in 10,146 alignments following the same patterns as in the first attempt. Nearly all the alignments appeared to belong to conserved domains of big protein families whose presence in BGCs is very common.

Finally, the approach that produced the most useful information consisted in a one-by-one BLASTp of each protein within the cluster against the metagenomes. All proteins except AbyU, AbyA3, AbyA5 and AbyW had metagenomic sequences significatively (E-value < 10^{-6}) aligned and thus would be present in the microbiome (Figure 2.2). The absence of AbyU, AbyA3, AbyA5 and AbyW was sufficient to determine the impossibility of finding abyssomicin producers in the pool, since AbyU is the natural Diels-Alderase that catalyses the [4 + 2] cycloaddition that results in the polycyclic abyssomicin formation and AbyA3 and AbyA5 are responsible for the tetronic acid moiety formation. The role of AbyW in abyssomicin biosynthesis is still unknown.

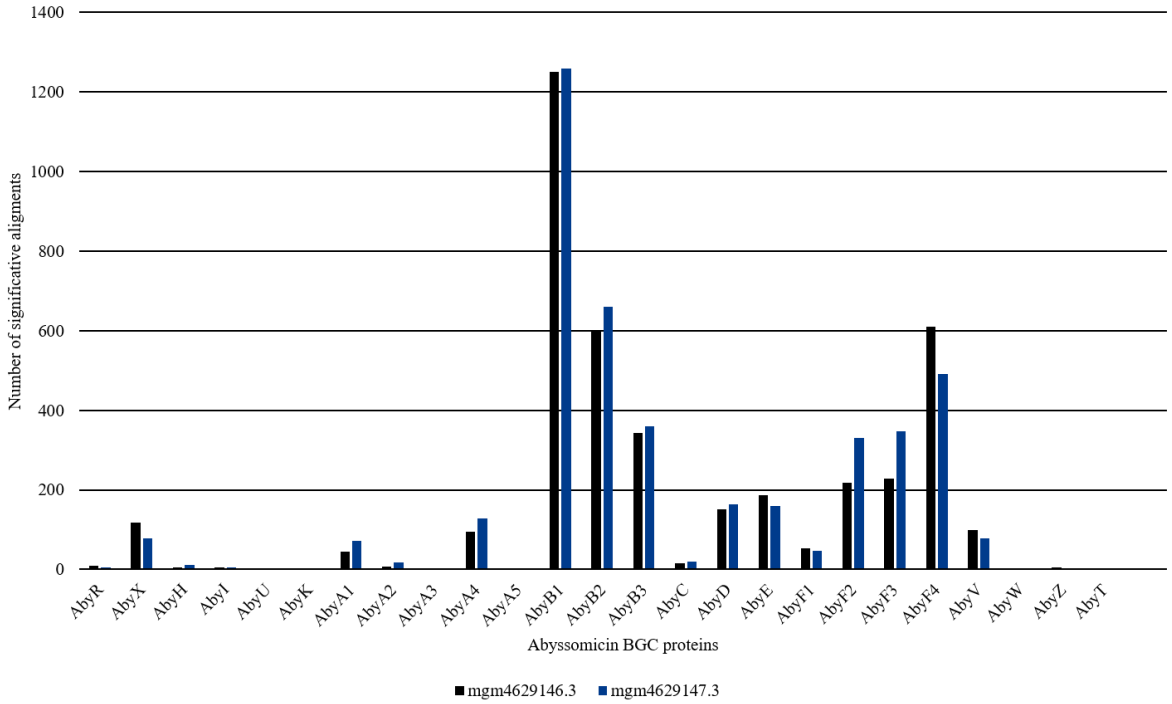

Figure 2.2. Number of significative alignments between each solar panel metagenome and each abyssomicin BGC gene.

Albeit it was not possible to find any potential abyssomicin producer in these samples, this analysis gave us an important insight into which abyssomicin-biosynthesis related proteins are less common in other BGCs and thus could be abyssomicin specific.

2.3.2 Habitat distribution of the Diels-Alderase positive metagenomes.

In order to study the habitat distribution of the bacteria harbouring an abyssomicin BGC, and considering that the Diels-Alderase AbyU could be used as an abyssomicin-biosynthesis specific marker, 3027 publicly available metagenomes were mined for the presence of AbyU and its already known homologs AbsU and AbmU (Table S2.3-S2.4). E-value < 10^{-6} was set as threshold to accept/reject the alignments and thus identify putative Diels-Alderases. 27% of the analysed metagenomes had aquatic origin, 31% belonged to soil samples, 22% were plant-associated and the remaining 20% covered human-built environments and different host-associated microbiomes (Figure S2.2). The results obtained showed that the three Diels-Alderase homologs share a similar habitat distribution, 31% of the AbyU positive metagenomes were from soil, 68% were plant-associated and 1% Arthropoda-associated (Figure 2.2B); 55% of the AbsU-positive had soil origin and 45% were plant-associated (Figure 2.2C) and AbmU displayed a similar distribution to AbyU with the only difference being its

additional presence in an artificial bioreactor environment (Figure 2.2D). Surprisingly, however, none of the AbyU, AbsU or AbmU positive metagenomes had aquatic origin.

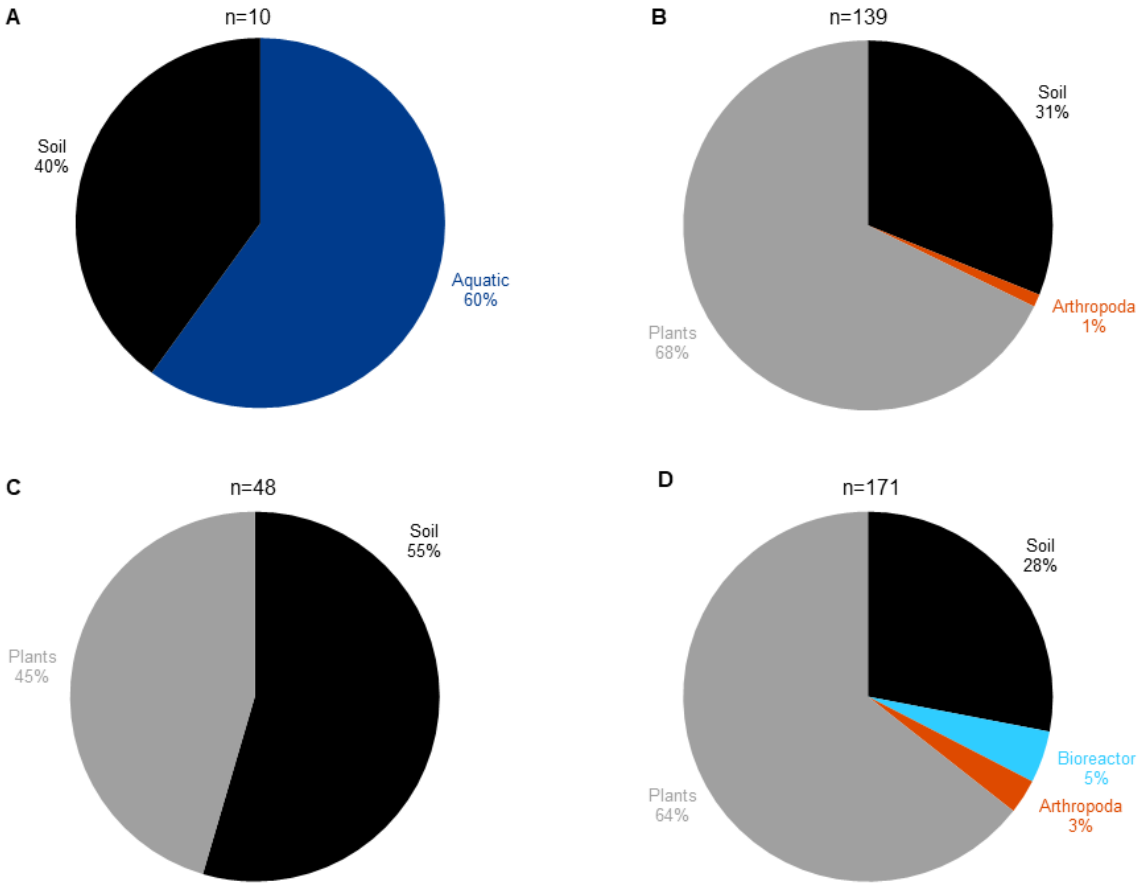

Figure 2.3. Habitat distribution of **A)** The abyssomicin producing bacteria isolated and reported in the literature until this manuscript was written. **B)** Metagenomes containing AbyU homologs. **C)** Metagenomes containing AbsU homologs. **D)** Metagenomes containing AbmU homologs.

In order to find an explanation to the absence of Diels-Alderase positive metagenomes in aquatic environments, possible taxonomic and sequencing depth biases between Diels-Alderase positive and negative metagenomes from aquatic origin were investigated. Specifically, the relative abundance of assembled and unassembled reads belonging to the domain *Bacteria* and the phylum *Actinobacteria* was compared in 50 randomly selected Diels-Alderase positive metagenomes from different environments against 50 aquatic Diels-Alderase negative metagenomes. The Mann-Whitney U test showed that the relative abundance of reads of the domain *Bacteria* and the phylum *Actinobacteria* was higher in Diels-Alderase positive metagenomes than in aquatic Diels-Alderase negative metagenomes (Figure S2.4). Similarly, the relative abundance of *Bacteria* and *Actinobacteria* was lower in terrestrial Diels-Alderase

negative metagenomes than in Diels-Alderase positive metagenomes (Figure S2.4). On the other hand, the sequencing depth of those same 50 randomly selected Diels-Alderase positive metagenomes was compared against the 50 aquatic and 50 terrestrial Diels-Alderase negative metagenomes. The Mann-Whitney U test showed that the sequencing depth of the Diels-Alderase positive metagenomes was significatively higher than the sequencing depth of the aquatic and terrestrial Diels-Alderase negative metagenomes (Figure S2.5).

2.3.3 Diels-Alderase directed genome mining and diversity of abyssomicin BGCs.

In order to gain a better overview over how abyssomicin-producing bacteria are environmentally distributed and the structural diversity of abyssomicin BGCs in nature, both partial and complete genomes available in public databases were mined. In a BLASTp of AbyU, AbsU and AbmU against the RefSeq NR database, 74 Diels-Alderase homologs from 66 different genomes were identified (Table S2.9). E-value < 10^{-6} was set as threshold to accept/reject the alignments and thus identify putative Diels-Alderases.

All the 66 Diels-Alderase positive genomes belonged to culturable bacterial strains. The habitat distribution of these isolates was, overall, similar to that found by metagenome mining. Specifically, about one third of the strains were recovered from soil, one third from plant-associated environments, and the remaining were recovered from mammals, annelids and lichens (Figure S2.6). Unlike the metagenome mining results, some Diels-Alderase positive bacterial species were recovered from marine environments.

The bacterial genomes were analysed in order to locate those Diels-Alderase homologs and study whether they were part of a potential abyssomicin BGC. This way, it was possible to identify and annotate five total and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs never described until now and with similar but not identical architectures to the *aby*, *abs* and *abm* clusters (Figure 2.4). Eleven of the Diels-Alderase homologs could be located in potential BGCs, three more were found in genomic regions apparently unrelated to any BGC and 11 were located in short contigs from which it was impossible to infer any information. Finally, two Diels-Alderase homologs were found in two different quartromicin BGCs and another two in potential tetronomycin and chlorothricin BGCs.

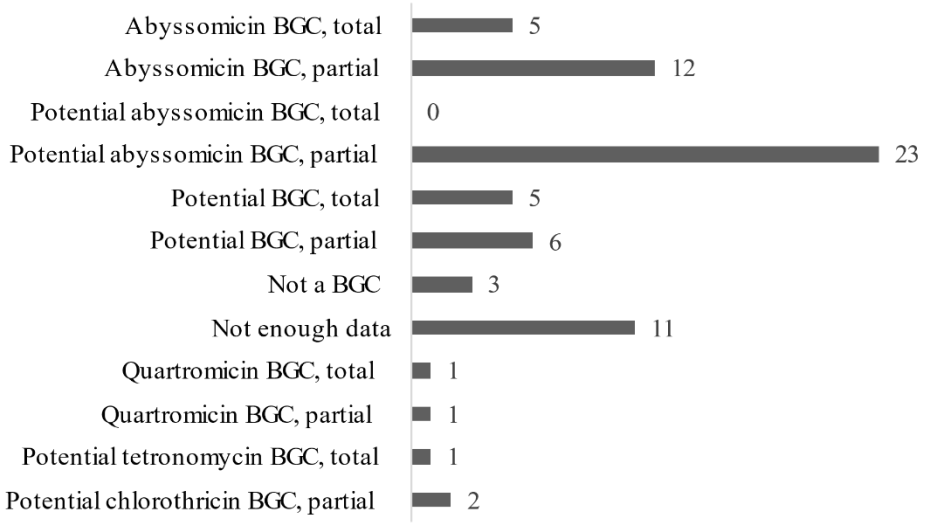

Figure 2.4. Recovered BGCs found through Diels-Alderase directed genome mining.

From the newly identified Diels-Alderase homologs it was possible to recover 40 total or partial new clusters potentially involved in the biosynthesis of abyssomicins (Figure S2.9-S2.12; Table S2.15-S2.85). These clusters were further classified according to their synteny in order to analyse their structural diversity. The analysis was carried out manually, as the modular nature of BGCs made the application of general synteny analysis tools impossible. Considering the diversity of biosynthetic genes and their disposition, abyssomicin and potential abyssomicin BGCs were classified into seven cluster types (Table 2.2). There were four genomes containing type 1a clusters and ten genomes displaying type 1b clusters from the genera *Micromonospora*, *Actinokineospora*, *Frankia*, *Herbidospora* and *Streptomyces* (Figure S2.9). There were seven clusters classified as type 2a and two clusters classified as type 2b. In this case, type 2a clusters were found in *Streptomyces*, *Actinokineospora* and *Micromonospora* and type 2b only in *Frankia* (Figure S2.10). Five clusters were classified as type 3, all belonging to *Streptomyces* and three clusters were type 4 clusters found in *Streptomyces* and *Streptacidiphilus* (Figure S2.11). Finally, there were 13 clusters that did not present enough similarity to any of the cluster types described above. These clusters were found in *Frankia*, *Actinokineospora*, *Lentzea*, *Kutzneria*, *Micromonospora*, *Streptomyces*, *Saccharothrix* and *Actinocrispum* and did not share any outstanding synteny pattern amongst themselves (Figure S2.12) neither with the five potential tetronomycin, chlorothricin, or quartromycin BGCs that were also found from the Diels-Alderase directed genome mining (Figure S2.13). The genomes that harboured a Diels-Alderase that was not part of an abyssomicin or potential abyssomicin BGC were not considered for this classification.

2.3.4 Evolutionary history of the abyssomicin BGCs.

Most of the Diels-Alderase positive bacteria were taxonomically identified as belonging to the phylum *Actinobacteria* and most of them to the genus *Streptomyces* (37 isolates), followed by seven *Frankia*, three *Herbidospora*, three *Actinomadura* and three *Micromonospora* strains (Figure 2.5). As was expected, all the genera formed monophyletic clusters, corroborating their correct taxonomic assignment. The abyssomicin BGCs were only identified in several species of some actinobacterial genera but not in all, suggesting that the abyssomicin BGCs may be acquired through horizontal gene transfer (HGT) events. This hypothesis was reinforced by the fact that the phylogenetic history of the Diels-Alderase (Figure 2.6) does not follow the same evolutionary history as of the species tree (Figure 2.5).

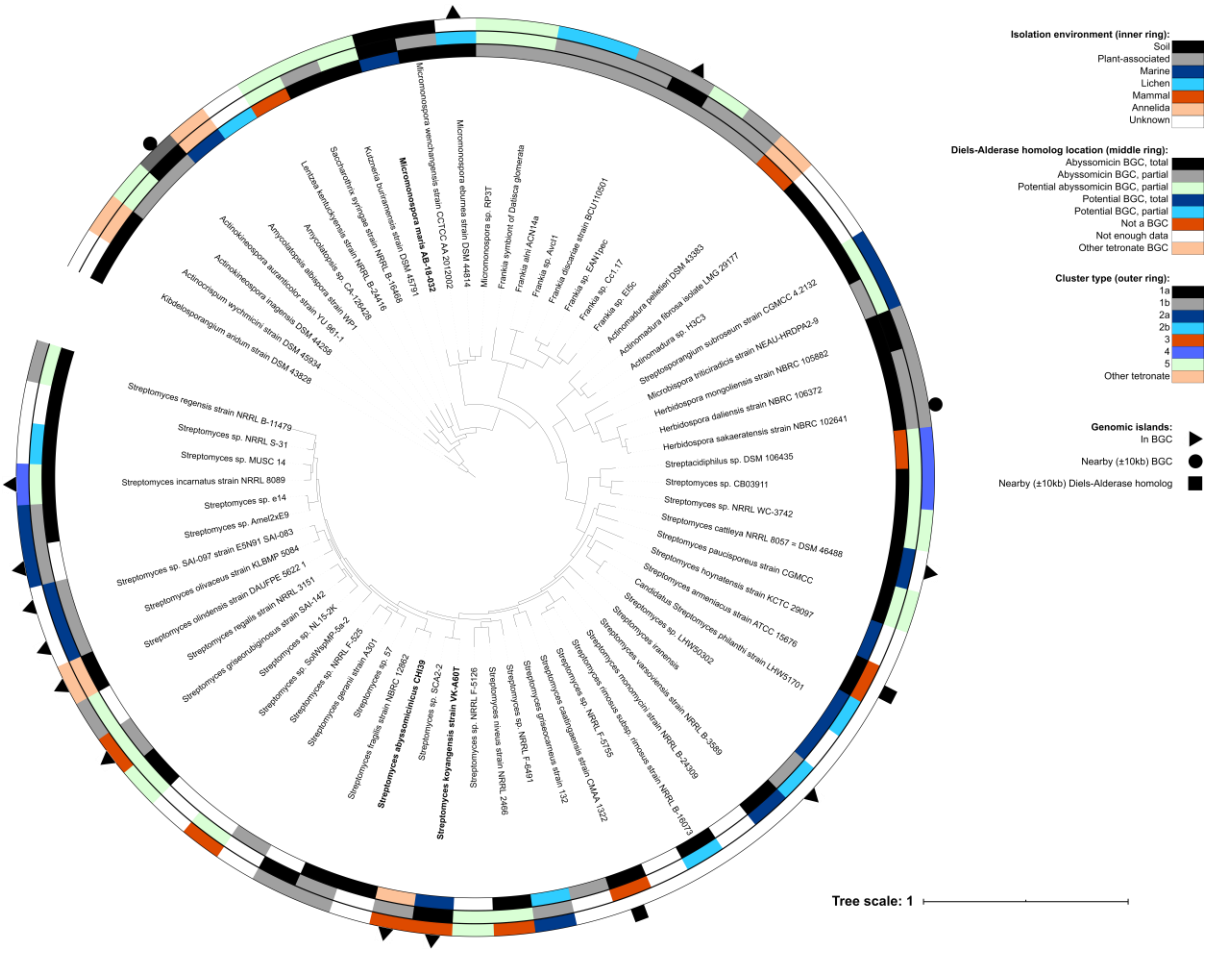

Figure 2.5. Phylogenomic tree of bacterial genomes containing Diels-Alderase homologs. The inner ring represents the environment where each strain was isolated, the middle ring depicts the location of the Diels-Alderase homolog and the outer ring shows the cluster type for those isolates found to have abyssomicin and potential abyssomicin BGCs both total and partial. Outer symbols indicate presence of genomic island inside the abyssomicin or potential abyssomicin BGC, nearby it (±10 kb upstream or downstream BGC) or nearby the Diels-Alderase (±10 kb upstream or downstream) when the isolate did not present an abyssomicin BGC. Bar 1 fixed nucleotide substitutions per site.

Interestingly, abyssomicin BGCs are usually associated with genomic islands (GI; Figure S2.9-S2.12; Table S2.86-S2.101) which may allow its HGT among taxa. Specifically, GIs were identified in the abyssomicin BGC of some *Streptomyces*, *Frankia*, *Herbidospora*, *Micromonospora* and *Actinokineospora*, nearby it (±10 kb upstream or downstream BGC) or nearby the Diels-Alderase (±10 kb upstream or downstream) (Figure 2.5). Albeit the vast majority of these HGT events take place among members of the phylum *Actinobacteria*, two *Proteobacteria*, namely *Pantoea* sp. A4 and *Photobacterium ganghwense* JCM 12487, harboured a Diels-Alderase homolog. The Diels-Alderase genes of both *Proteobacteria* strains were phylogenetically related among them and in turn related to other *Streptomyces* strains

(Figure 2.5). The closest neighbour to both proteobacterial strains was *S. caatingaensis*. Thereby, these intra- and inter- phyla HGT events may be explained by the presence of mobile elements such as transposases and integrases flanking or within the BGCs (Table S2.86-S2.101). Moreover, several Diels-Alderase paralogs were found in the mined genomes of *A. auranticolor* YU 961-1, *Frankia* sp. Cc1.17, *Streptomyces* sp. NL15-2K, *Streptomyces* sp. NRRL F-525, *Streptomyces* sp. NRRL S-31 and *S. syringae* NRRL B-16468 (Figure 2.5).

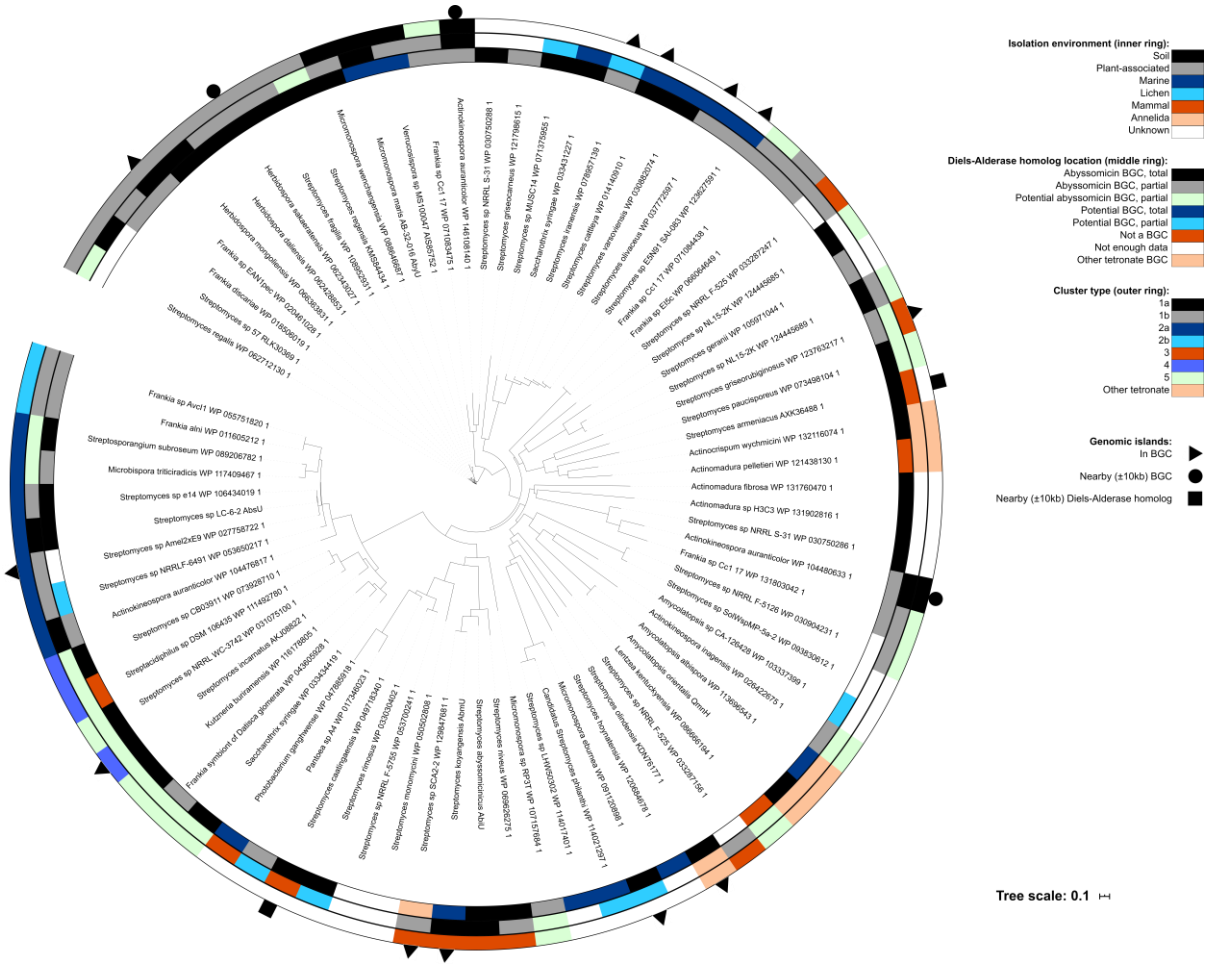

Figure 2.6. Phylogenetic tree of the Diels-Alderase homologs. The inner ring represents the environment where each strain was isolated, the middle ring depicts the location of the Diels-Alderase homolog and the outer ring shows the cluster type for those isolates found to have abyssomicin and potential abyssomicin BGCs both total and partial. Outer symbols indicate presence of genomic islands inside the abyssomicin or potential abyssomicin BGC, nearby it (±10 kb upstream or downstream BGC) or nearby the Diels-Alderase (±10 kb upstream or downstream) when the isolate did not present an abyssomicin BGC. Bar 0.1 fixed nucleotide substitutions per site.

On the other hand, evolutionary pressure has shaped the abyssomicin BGCs, widening the functional and structural diversity of this secondary metabolite. In fact, the presence of tailoring genes is variable among species as well as the Diels-Alderase gene location within the BGCs

(Figure S2.9-S2.12; Table S2.15-S2.85). However, the synteny of abyssomicin BGCs lacks phylogenetic signal and hence the abyssomicin BGC classification proposed in the present study could not be used to trace its evolutionary history.

2.3.5 Preliminary abyssomicin production screening.

In order to close up the story, a preliminary abyssomicin production screening was carried out with five of the strains containing potential abyssomicin BGCs: *Streptomyces regensis* RRL X-5263, *Streptomyces niveus* C7655, *Streptomyces incarnatus* NRRL 8089, *Actinokineospora auranticolor* YU 961-1 and *Micromonospora wenchangensis* CCTCC AA 2012002. Three different approaches were followed.

First, the strains were cultured in liquid SGG and A1 media, the two media in which abyssomicin production has been reported to date (Gottardi *et al*., 2011; Tu *et al*., 2018). Samples of these cultures were taken for further analysis from day 1 to day 5 of fermentation. Organic extractions with ethyl acetate were carried out to purify any possible abyssomicin-like compounds and the final extracts were analysed both through agar plate diffusion assays with and without PABA against *B. subtilis*, *E. coli* and *M. luteus* and through LC-MS. The agar plate diffusion assay of the fermentation extracts revealed that *S. regensis* produced, both in A1 and SGG, one or more compounds with strong inhibitory activity against *B. subtilis* that generated ZOI of 10-14 mm in radius measured from the edge of the disk (Figure 2.7A and 2.7B). This strain, grown in A1, also produced ZOI up to 8 mm against *M. luteus*. *S. niveus* extracts produced ZOI of 8-10 mm against *B. subtilis* when the strain was cultured over 3 days in A1 and ZOI of 12 mm after growing 4 days in SGG, and ZOI of 6-8 mm against *M. luteus* when grown over 4 days (Figure 2.7C and 2.7D). The fermentation extracts of *S. incarnatus* in A1 produced ZOI of up to 8 mm against *M. luteus* and between 2-4 mm against *B. subtilis* (Figure 2.7E). Similar results were observed for *S. incarnatus* grown in SGG (Figure 2.7F). *A. auranticolor* produced, both in A1 and SGG, small molecules that inhibited the growth of *B. subtilis* producing ZOI of 3 mm, however, ZOI were not observed against *M. luteus* (Figure 2.7G and 2.7H). Finally, extracts from *M. wenchangensis* cultures produced minimal ZOI against *B. subtilis* and *M. luteus* (Figure 2.7I and 2.7J). As a positive control, fermentation extracts of *M. maris*, grown in A1 and SGG, containing mainly abyssomicin C (Figure 2.7K and 2.7L), reported ZOI against *B. subtilis* of up to 3 mm and no activity against *M. luteus*. None of the fermentation extracts had any inhibitory effect against *E. coli*. The antagonistic

effect of PABA was only barely observed for *S. niveus* and *M. maris* extracts and no differences in the ZOI generated by nalidixic acid disks were appreciated. The filter disks containing 10% methanol did not generate any ZOI against any of the targeted strains.

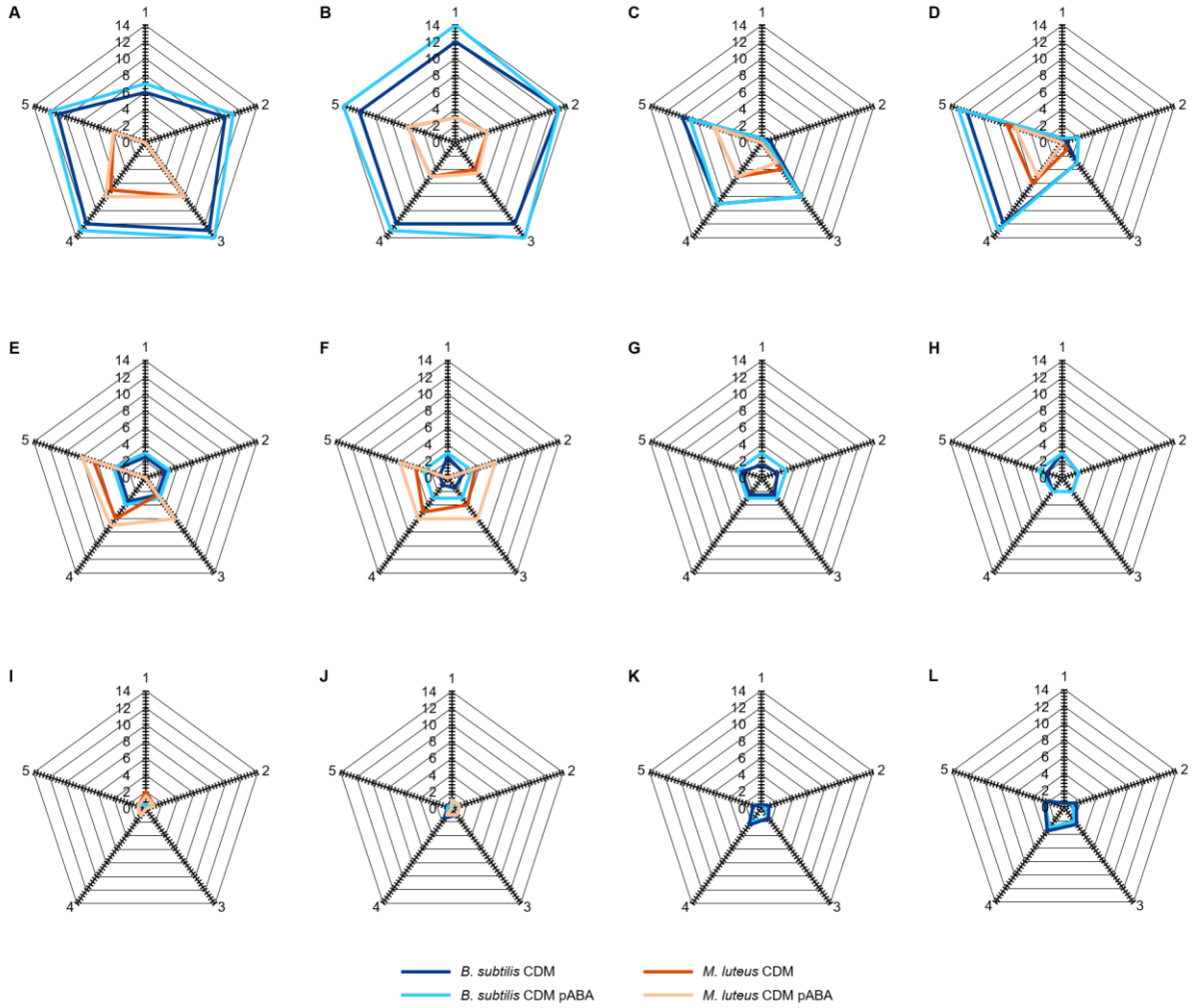

Figure 2.7. Zones of inhibition in mm generated by the fermentation extracts at days 1-5 of **A)** *S. regensis* in A1, **B)** *S. regensis* in SGG, **C)** *S. niveus* in A1, **D)** *S. niveus* in SGG, **E)** *S. incarnatus* in A1, **F)** *S. incarnatus* in SGG, **G)** *A. auranticolor* in A1, **H)** *A. auranticolor* in SGG, **I)** *M. wenchangensis* in A1, **J)** *M. wenchangensis* in SGG, **K)** *M. maris* in A1 and **L)** *M. maris* in SGG against *B. subtilis* and *M. luteus* grown in CDM plates with and without PABA. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

Then, those same fermentation extracts were run through a mass spectrometer ESI microTOF-Q II (Figure S2.14-S2.24). It was possible to identify at least three of the compounds based on their mass and the antiSMASH predictions, however, none of the already known abyssomicins/neoabyssomicins were found. The remarkably potent antimicrobial activity of *S.*

regensis extracts was associated to the production of actinomycin D based on LC-MS data and the molecular weight of the compound (Figure S2.14-S2.15).

Finally, plug agar plate diffusion assays were carried out. Plugs from *S. regensis*, *S. niveus*, *S. incarnatus*, *A. auranticolor* and *M. wenchangensis* grown in six different production media were used to further investigate the production profile of these strains in different solid media. *S. regensis* plugs from SGG, A1, R5A and A1Sal* plates generated ZOI of 10 mm against *B. subtilis* and of 8 and 4 mm when the plugs were from SM10 and SM17 respectively. Plugs from SGG, A1, R5A and A1Sal* plates also generated ZOI between 2 and 4 mm against *M. luteus*. Only plugs of *S. regensis* grown in SM17 produced ZOI against *E. coli* (Figure 2.8A). *S. niveus* plugs generated ZOI that varied between 1-7 mm against *B. subtilis* and 0.5-2 mm against *M. luteus* and ZOI were not observed against *E. coli* (Figure 2.8B). For *S. incarnatus*, only plugs from A1 plates showed inhibitory activity against *B. subtilis*, while also A1Sal* and R5A plugs showed antimicrobial activity against *M. luteus* (Figure 2.8C). *A. auraticolor* plugs generated ZOI between 3-8 mm against *B. subtilis* and between 0.5-2 mm against *M. luteus*. *A. auraticolor* plugs from SGG, A1, R5A and SM10 also produced ZOI of 5-10 mm against *E. coli* (Figure 2.8D). *M. wenchangensis* plugs from the SGG, SM10 and SM17 plates only produced ZOI of 0.5 mm against *B. subtilis* and *M. luteus* (Figure 2.8E). As a reference, *M. maris* plugs generated ZOI of up to 2.5 mm against *B. subtilis* (Figure 2.8 F). No relevant differences were appreciated between the ZOI generated in plates containing PABA and plates without PABA except for the plug of *S. regensis* in SM17, which ZOI was reduced from 7 mm to 4 mm against *E. coli*.

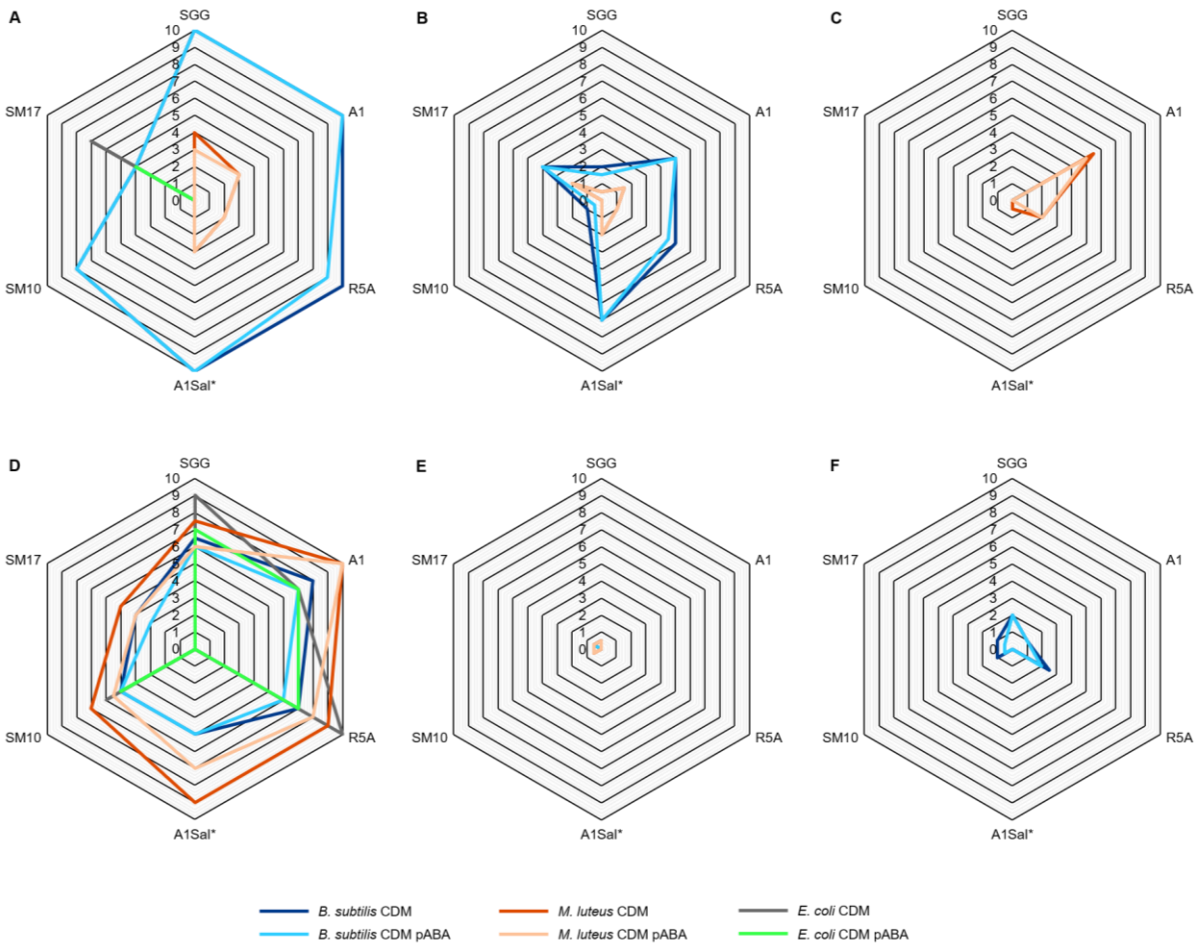

Figure 2.8. Antimicrobial plate diffusion assay results with and without PABA of **A)** *S. regensis*, **B)** *S. niveus*, **C)** *S. incarnatus*, **D)** *A. auranticolor*, **E)** *M. wenchangensis* and **F)** *M. maris* grown in six different production media for 15 days. Antimicrobial activity was measured as the zone of inhibition (mm) against three target strains: *B. subtilis*, *E. coli* and *M. luteus*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

2.4 Discussion.

2.4.1 Preliminary metagenome mining.

To date, only ten bacterial strains have been reported to produce abyssomicins (Table S2.1). From these strains, 38 abyssomicins with differences at structures and bioactivities have been characterised (Sadaka *et al*., 2018). In a preliminary metagenome mining study, two solar panel microbiomes were mined looking for new potential abyssomicin producers. A one-by-one BLASTp of each protein within the *aby* BGC against the metagenomes showed that AbyU, AbyA3, AbyA5 and AbyW would not be present in the microbiomes and thus, there would not be abyssomicin producers in the pools. Despite that, those results provided an important insight into which abyssomicin biosynthetic proteins could be specific to abyssomicin biosynthesis, as AbyU is the natural Diels-Alderase that catalyses the [4 + 2] cycloaddition reaction that results in the polycyclic abyssomicin formation and AbyA3 and AbyA5 are responsible for the tetronic acid moiety formation (Gottardi *et al*., 2011, Byrne *et al*., 2016). These proteins could be used to lead larger mining studies focused on abyssomicins.

2.4.2 Habitat distribution of the Diels-Alderase hosts discovered through metagenome and genome mining.

With the aim of studying the distribution of those microorganisms capable of producing new abyssomicin molecules, an extensive diversity of metagenomes and genomes was analysed *in silico*. AbyU is the natural Diels-Alderase present in the abyssomicin BGC that catalyses the formation of the heterobicyclic ring system that characterises this family of natural products. Very few enzymes in nature can catalyse this reaction and despite being capable of accepting structurally diverse substrates, sequence conservation with the closest known spirotetronate cyclases is minimal (Byrne *et al*., 2016). This enzyme was selected to lead the mining as it is essential in abyssomicin biosynthesis.

Here, 3,027 metagenomes were mined for the presence of AbyU, AbsU and AbmU, and the results showed that Diels-Alderase positive microorganisms have a strikingly diverse environmental distribution, being mainly present in soil and plant-associated microbiomes but totally absent in aquatic habitats (Figure 2.2). Since the few isolates reported in the literature to produce abyssomicins were equally distributed between aquatic and soil environments (Figure 2.2A; Table S2.1), these results were totally unexpected. After examining the taxonomic composition of 50 aquatic Diels-Alderase negative and 50 Diels-Alderase positive metagenomes from different environments, it was possible to conclude that the Diels-Alderase

positive metagenomes have a higher relative abundance of *Bacteria* and *Actinobacteria* than Diels-Alderase negative metagenomes from aquatic environments (Figure S2.4). Furthermore, those Diels-Alderase negative metagenomes from aquatic environments showed, in general, a lower sequencing depth than the Diels-Alderase positive metagenomes (Figure S2.5). Therefore, the fact that metagenomes of aquatic origin have a lower sequencing depth, and that the abundance of *Bacteria* and *Actinobacteria* is lower, could make it less likely to sequence Diels-Alderase homolog genes when shotgun sequencing aquatic metagenomes. On the contrary, by using the appropriate culturing techniques, those low abundant abyssomicin-producing *Actinobacteria* could be enriched from aquatic environments (Table S2.2).

Interestingly, it was observed that all the abyssomicin-producing strains isolated from aquatic environments so far come specifically from marine sediments (Table S2.2). Meaning that the abyssomicins could play a key role in the biology or ecology of bacteria inhabiting benthic regions. Moreover, it is tempting to hypothesise that abyssomicin-producing bacteria may be involved in symbioses with higher organisms, which has been seen before for other different antibiotic-producing strains that play an important role as defensive symbionts both in marine and terrestrial ecosystems (Gunatilaka, 2006; Seipke *et al*., 2012; Adnani *et al*., 2017). The abyssomicins could also act as signal molecule in plant-bacteria communication or as precursors involved in plant growth and development, as reported before in the *Frankia* and *Micromonospora* genera through, for example, the formation of nitrogen fixing actinonodules (Trujillo *et al*., 2010; Sellstedt & Richau, 2013). Further investigations will be needed in order to unravel the biased habitat distribution of Diels-Alderase positive bacteria.

The results of this study also highlight some of the limitations of the bioinformatic approaches used. In first instance, data normalisation prior to including the metagenomes in the analysis should have been carried out. In addition to that, despite the uniqueness of the catalytic activity of the Diels-Alderase, the low similarity percentage between the sequences of the three already known abyssomicin Diels-Alderases could indicate a great sequence diversity within this kind of enzymes, meaning that only by using the sequence of these three proteins, other Diels-Alderases from a different evolutive origin could be being misrepresented. Future work could look into using tools with different algorithms to that of BLASTp to identify more distantly related Diels-Alderases. For example, PSI-BLAST and PHI-BLAST use the query sequence to build a profile of that protein that is used to find closely related homologs. Then, with the sequences of all those homologs another profile is created to search for more distantly related homologs, and this continues as an iterative process. Apart from sequence and motif-based methods, since generally protein structure is more conserved that protein sequence, structure-

based methods represent a strong alternative for the discovery of enzymes with similar functions but less related sequences. Indeed, with the exponential development of machine learning tools and artificial intelligence the paradigm sequence-function of proteins may be more accessible than ever, easing greatly the discovery of proteins with specific functions (Callaway, 2020).

Altogether, 74 Diels-Alderase homologs present in 66 different genomes were identified (Table S2.9) from which it was possible to identify and annotate five total and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs. Indeed, all 40 abyssomicin and potential abyssomicin producers are culturable strains whose habitat distribution follows the same patterns found through metagenome mining as none of them was recovered from aquatic samples (Figure S2.6). In this case, 60.6% of the Diels-Alderase positive genomes displayed an abyssomicin or potential abyssomicin BGC. In the remaining genomes in which the Diels-Alderase was not located in any BGC, it was not possible to predict its metabolic function. Previous studies reported other Diels-Alderases involved in the synthesis of other natural products, with the exception of riboflavin synthases that are involved in primary metabolism (Lichman *et al*., 2019).

Therefore, based on the genome and metagenome mining, it can be concluded that the potential abyssomicin producers have a cosmopolitan distribution albeit their presence in aquatic habitat is limited. This strongly suggests that abyssomicin bioprospecting efforts should not be focused on aquatic environments but rather on soil and plant-associated ones. Also, two Diels-Alderase homologs were found in two different quartromicin BGCs and another two in potential tetronomycin and chlorothricin BGCs. The presence of those four Diels-Alderase homologs within BGCs belonging to other natural products is well justified, as quartromicin, tetronomycin and chlorothricin share the same tetronate cycloaddition as the abyssomicins (Vieweg *et al*., 2014).

Moreover, 11 of the Diels-Alderase homologs detected in the mined genomes were in potential non-abyssomicin BGCs, three more were found in genomic regions *a priori* unrelated to any BGC and 11 appeared in short contigs from which it was impossible to infer any information. In this case, only ten of the 66 genomes analysed were completely sequenced and only seven isolates were sequenced with third generation sequencing technologies (Table S2.9). The identification of the Diels-Alderase homologs location within the genomes and the recovery of potential BGCs was influenced by the quality of the sequencing technology used and the assembly level achieved by each previous individual study. Some factors such as the high G+C

content of actinomycete genomes affect the sequencing reactions and the assembly process (Nakamura *et al*., 2011), however, the biggest challenge appears to be the recovery of the highly-conserved and modular sequences of polyketide synthases characterised by displaying highly similar intragenic and intergenic tandem repeats at nucleotide level, which, in many cases, are longer than the read-length of the sequencing technology used (Gomez-Escribano *et al*., 2016). Moreover, large PKS clusters can often be distributed along several contigs, and it has been demonstrated that sequencing errors can introduce false frameshifts into the large PKS sequences (Blažič *et al*., 2012). Finally, the presence of Diels-Alderase homologs outside abyssomicin BGCs, could be explained by the presence of transposases flanking Diels-Alderase homologs allowing their genetic recombination along the genome (Table S2.86-S2.101). Specifically, the Diels-Alderase homologs of *Streptomyces caatingaensis* CMAA 1322 and *Streptomyces armeniacus* ATCC 15676 were not part of an abyssomicin BGC but showed transposases on both sides (Figure 2.5; Table S2.90; Table S2.92).

2.4.3 Evolutionary history of the abyssomicin BGC.

It is well-known that *Actinobacteria* are characterised by their ability to produce a wide variety of secondary metabolites and, despite the problem of rediscovering already known molecules, bacteria from the phylum *Actinobacteria* are still one of the most prolific sources of chemical diversity (Genilloud, 2017). The presence of abyssomicin BGCs is limited to the phylum *Actinobacteria*, mainly representatives of the genus *Streptomyces* and *Micromonospora*. The constraint of the abyssomicin BGC to some specific strains suggests that speciation was not the primary driver for dissemination of this cluster (Figure 2.4). Instead, HGT may have played an important role in the transmission of abyssomicin BGCs, which may have jumped among taxa through mobile elements (Ziemert *et al*., 2014; Hall *et al*., 2017). Indeed, many integrases and transposases were found surrounding or inside the abyssomicin BGCs (Table S2.86-S2.101).

Many BGCs in *Actinobacteria* evolve through HGT events, but only a few studies have demonstrated it (Choudoir *et al*., 2018). For example, in a genome mining study on 75 *Salinispora* strains, 124 pathways involved in the synthesis of PKS and NRPS natural products were identified and showed that HGT events were responsible for the majority of pathways, which occurred in only one or two strains, as acquired pathways were incorporated into genomic islands (Ziemert *et al*., 2014). In another example, the secondary metabolite clusters on the chromosome of *S. avermitilis* ATCC31267 were found to contain many transposase

genes in the regions near both ends of the clusters, suggesting these transposases might have been involved in the transfer of these clusters (Omura *et al*., 2001). Similarly, it was demonstrated that the rifamycin BGC in *Salinispora arenicola* CNS-205 had been acquired through HGT directly from *Amycolatopsis mediterranei* S699 by genomic island movement (Penn *et al*., 2009).

Although HGT events are more frequent among phylogenetically close taxa, in this case within the phylum *Actinobacteria*, HGT events can take place among different phyla. In the present study, it was possible to identify a potential HGT event of Diels-Alderases from a representative of the genus *Streptomyces* to two strains of the phylum *Proteobacteria*, namely *Pantoea* sp. A4 and *Photobacterium ganghwense* JCM 12487 (Figure 2.5). The transmission of functional BGCs among phyla was also reported by other authors (Zeng *et al*., 2014). Unfortunately, neither transposases nor integrases were identified nearby the Diels-Alderases of *Pantoea* sp. A4 and *Photobacterium ganghwense* JCM 12487, which could have explained the HGT event.

The acquisition of an abyssomicin BGC by a bacterial strain could increase its evolutionary fitness and therefore enhance its competitiveness against other members of the community. In fact, the biological activity of abyssomicins includes antimicrobial activities against Gram-positive bacteria and *Mycobacteria* (Riedlinger *et al*., 2004; Freundlich *et al*., 2010). Other biological activities discovered so far are antitumor properties, latent HIV reactivator, anti-HIV and HIV replication inducer properties (Sadaka *et al*., 2018). The wide diversity of abyssomicin BGCs found through genome mining suggests that a plethora of abyssomicin-like molecules remain undiscovered.

2.4.4 Preliminary abyssomicin production screening.

Finally, the secondary metabolite production profile of five strains with new potential abyssomicin BGCs was analysed. First, the strains were screened for their ability to produce antimicrobial compounds in different production media both in liquid and in plates. *S. regensis*, *S. niveus* and *A. auranticolor* proved to be good producers of antimicrobial compounds, especially against *B. subtilis* and *M. luteus* (Figure 2.7 and 2.8). Then, fermentation extracts from all the strains were also analysed, focusing at finding already known abyssomicins. However, none of the compounds could be identified as an already known abyssomicin.

There are several hypotheses that could explain these results. In first place, although *A. auranticolor* and *M. wenchangensis* display abyssomicin BGCs that are very similar to the abyssomicin BGC of *M. maris*, the genomic context where these clusters are found plays a key role in the resulting natural products produced (Jenke-Kodama *et al*., 2006; Ridley *et al*., 2008). It is likely that these strains could be producing new abyssomicins that I was not able to identify due to the lack of time and expertise to carry out NMR experiments. Another option is the fact that these clusters could be transcriptionally silent under standard laboratory conditions, which has been shown to be very common, as secondary metabolite BGCs are often under strict pleotropic and pathway-specific regulation (Liu *et al*., 2013; Rutledge & Challis, 2015; Ke & Yoshikuni, 2020). These clusters could also be inactive or non-functional at translation level due to mutations that occurred in the course of evolution (Hoogendoorn *et al*., 2018). In this case, characterisation of the enzymatic pathways would be necessary to determine whether the gene clusters are functional or vestiges of evolution (Micallef *et al*., 2015; Amos *et al*., 2017). This was the case, for example, of the 19 cyanobacterial strains that were found to encode a cyanobactin BGC, but it turned out that many of these gene clusters were non-functional due to the lack of some essential genes, the presence of insertions and truncated genes, although active versions had been identified within closely related strains (Leikoski *et al*., 2012; Leikoski *et al*., 2013). Future work could include a deeper analysis of the corresponding abyssomicin BGCs, to estimate the structures of the compounds that could be produced and could look at elucidating the production of novel abyssomicins by these strains.

2.5 Conclusions and future perspectives.

The aim of this study was to shed some light into the structural diversity, habitat distribution and evolutionary history of the abyssomicin BGC. Through metagenome and genome mining, it was discovered that the habitat distribution of microorganisms harbouring putative Diels-Alderases is restricted to that of the phylum *Actinobacteria*, with representatives mainly of the genus *Streptomyces* and *Frankia*, which are primarily present in soil and plant-associated environments. Surprisingly, it was not possible to find any Diels-Alderase positive bacterium in aquatic environments although six out of ten reported abyssomicin producers were isolated from marine sediments. Moreover, the vast structural diversity of abyssomicin BGCs found, could reflect its horizontal evolutionary history, and it is possible to predict that a plethora of abyssomicins remain unknown to date.

Future work could focus on unraveling the ecological role that abyssomicin-like molecules play in the corresponding environments. Apart from that, the identification and characterisation of new Diels-Alderase homologs could hold great potential as part of the synthetic biology toolbox to generate libraries of novel non-natural biomolecules. This could also facilitate the development of environmentally friendly synthetic routes to a wide variety of useful compounds. Taken together, the results of the present work reveal the interest of a new bioprospecting strategy to identify natural products such as abyssomicins out of their currently assumed environmental distribution.

2.6 References.

Abdalla, M. A., Yadav, P. P., Dittrich, B., Schüffler, A. & Laatsch, H. (2011). Ent-Homoabyssomicins A and B, two new spirotetronate metabolites from *Streptomyces* sp. Ank 210. *Org Lett* **13**, 2156–2159.

Abugrain, M. E., Brumsted, C. J., Osborn, A. R., Philmus, B. & Mahmud, T. (2017). A highly promiscuous ß-ketoacyl-ACP synthase (KAS) III-like protein is involved in pactamycin biosynthesis. *ACS Chem Biol* **12**, 362–366.

Adnani, N., Rajski, S. R. & Bugni, T. S. (2017). Symbiosis-inspired approaches to antibiotic discovery. *Nat Prod Rep* **34**, 784–814.

Amos, G. C. A., Awakawa, T., Tuttle, R. N., Letzel, A. C., Kim, M. C., Kudo, Y., Fenical, W., Moore, B. S. & Jensen, P. R. (2017). Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *Proc Natl Acad Sci U S A* **114**, E11121–E11130.

Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoad, G., Winsor, G. L. & Sl Brinkman, F. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* **45**, W30-W35.

Blažič, M., Starcevic, A., Lisfi, M., Baranasic, D., Goranovič, D., Fujs, Š., Kuščer, E., Kosec, G., Petkovic´, H. & other authors. **(2012).** Annotation of the modular polyketide synthase and nonribosomal peptide synthetase gene clusters in the genome of *Streptomyces tsukubaensis* NRRL18488. *Appl Environ Microbiol* **78**, 8183–8190.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H. & Weber, T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 1–7.

Byrne, M. J., Lees, N. R., Han, L.-C., van der Kamp, M. W., Mulholland, A. J., Stach, J. E. M., Willis, C. L. & Race, P. R. (2016). The catalytic mechanism of a natural Diels–Alderase revealed in molecular detail. *J Am Chem Soc* **138**, 6095–6098.

Callaway, E. (2020). 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. *Nature* **588**, 203-204.

Chevrette, M. G., Gutiérrez-García, K., Selem-Mojica, N., Aguilar-Martínez, C., Yañez-Olvera, A., Ramos-Aboites, H. E., Hoskisson, P. A. & Barona-Gómez, F. (2020). Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat Prod Rep* **37**, 566-599.

Choudoir, M. J., Pepe-Ranney, C. & Buckley, D. H. (2018). Diversification of secondary metabolite biosynthetic gene clusters coincides with lineage divergence in *Streptomyces*. *Antibiotics* **7**.

Freundlich, J. S., Lalgondar, M., Wei, J.-R., Swanson, S., Sorensen, E. J., Rubin, E. J. & Sacchettini, J. C. (2010). The abyssomicin C family as *in vitro* inhibitors of *Mycobacterium tuberculosis*. *Tuberculosis* **90**, 298–300.

Genilloud, O. (2017). Actinomycetes: still a source of novel antibiotics. *Nat Prod Rep* **34**, 1203–1232.

Gomez-Escribano, J. P., Alt, S. & Bibb, M. J. (2016). Next generation sequencing of *Actinobacteria* for the discovery of novel natural products. *Mar Drugs* **14**, 6–8.

Gottardi, E. M., Krawczyk, J. M., Von Suchodoletz, H., Schadt, S., Mühlenweg, A., Uguru, G. C., Pelzer, S., Fiedler, H. P., Bibb, M. J. & other authors. **(2011).** Abyssomicin biosynthesis: Formation of an unusual polyketide, antibiotic-feeding studies and genetic analysis. *ChemBioChem* **12**, 1401–1410.

Grabarczyk, D. B., Chappell, P. E., Johnson, S., Stelzl, L. S., Lea, S. M. & Berks, B. C. (2015). Structural basis for specificity and promiscuity in a carrier protein/enzyme system from the sulfur cycle. *Proc Natl Acad Sci U S A* **112**, E7166–E7175.

Gunatilaka, A. A. L. (2006). Natural products from plant-associated microorganisms: Distribution, structural diversity, bioactivity, and implications of their occurrence. *J Nat Prod* 509–526.

Hall, J. P. J., Brockhurst, M. A. & Harrison, E. (2017). Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria. *Philos Trans R Soc Lond B Biol Sci* **372**.

Hoogendoorn, K., Barra, L., Waalwijk, C., Dickschat, J. S., van der Lee, T. A. J. & Medema, M. H. (2018). Evolution and diversity of biosynthetic gene clusters in *Fusarium*. *Front Microbiol* **9**.

Huang, H., Song, Y., Li, X., Wang, X., Ling, C., Qin, X., Zhou, Z., Li, Q., Wei, X. & Ju, J. (2018). Abyssomicin monomers and dimers from the marine-derived *Streptomyces koyangensis* SCSIO 5802. *J Nat Prod* **81**, 1892–1898.

Huang, P., Xie, F., Ren, B., Wang, Q., Wang, J., Wang, Q., Abdel-Mageed, W. M., Liu, M., Han, J. & other authors. **(2016).** Anti-MRSA and anti-TB metabolites from marine-derived *Verrucosispora* sp. MS100047. *Appl Microbiol Biotechnol* **100**, 7437–7447.

Igarashi, Y., Iida, T., Oku, N., Watanabe, H., Furihata, K. & Miyanouchi, K. (2012). Nomimicin, a new spirotetronate-class polyketide from an actinomycete of the genus *Actinomadura*. *J Antibiot (Tokyo)* **65**, 355–359.

Igarashi, Y., Ogura, H., Furihata, K., Oku, N., Indananda, C. & Thamchaipenet, A. (2011). Maklamicin, an antibacterial polyketide from an endophytic *Micromonospora* sp. *J Nat Prod* **74**, 670–674.

Igarashi, Y., Yu, L., Miyanaga, S., Fukuda, T., Saitoh, N., Sakurai, H., Saiki, I., Alonso-

Vega, P. & Trujillo, M. E. (2010). Abyssomicin I, a modified polycyclic polyketide from *Streptomyces* sp. CHI39. *J Nat Prod* **73**, 1943–1946.

Iglesias, A., Latorre-Perez, A., Stach, J., Porcar, M. & Pascual, J. (2019). Out of the abyss: Genome and metagenome mining reveals unexpected environmental distribution of abyssomicins. *bioRxiv* https://doi.org/10.1101/789859.

Jenke-Kodama, H., Börner, T. & Dittmann, E. (2006). Natural biocombinatorics in the polyketide synthase genes of the *Actinobacterium Streptomyces avermitilis*. *PLoS Comput Biol* **2**, e132.

Jiang, Z.-D., Jensen, P. R. & Fenical, W. (1999). Lobophorins A and B, new antiinflammatory macrolides produced by a tropical marine bacterium. *Bioorg Med Chem Lett* **9**, 2003-2006.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–66.

Ke, J. & Yoshikuni, Y. (2020). Multi-chassis engineering for heterologous production of microbial natural products. *Curr Opin Biotechnol* **62**, 88-97.

Keller, S., Schadt, H. S., Ortel, I. & Süssmuth, R. D. (2007). Action of atrop-abyssomicin C as an inhibitor of 4-amino-4-deoxychorismate synthase PabB. *Angew Chemie - Int Ed* **46**, 8284–8286.

Komaki, H., Sakurai, K., Hosoyama, A., Kimura, A., Trujillo, M. E., Igarashi, Y., Tamura, T. (2019). Diversity of PKS and NRPS gene clusters between *Streptomyces abyssomicinicus* sp. nov. and its taxonomic neighbor. *J Antibiot (Tokyo)* **73**, 141–151.

Lacoske, M. H. & Theodorakis, E. A. (2015). Spirotetronate polyketides as leads in drug discovery. *J Nat Prod* **3**, 562–575.

Leikoski, N., Fewer, D. P., Jokela, J., Alakoski, P. & Wahlsten, M. (2012). Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. *PLoS One* **7**, 43002.

Leikoski, N., Liu, L., Jokela, J., Wahlsten, M., Gugger, M., Calteau, A., Permi, P., Kerfeld, C. A., Sivonen, K. & Fewer, D. P. (2013). Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides. *Chem Biol* **20**, 1033–1043.

León, B., Navarro, G., Dickey, B. J., Stepan, G., Tsai, A., Jones, G. S., Morales, M. E., Barnes, T., Ahmadyar, S. & other authors. (2015). Abyssomicin 2 reactivates latent HIV-1 by a PKC- and HDAC independent mechanism. *Org Lett* **17**, 262–265.

Letunic, I. & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256–W259.

Lichman, B. R., O'Connor, S. E. & Kries, H. (2019). Biocatalytic strategies towards [4+2]

cycloadditions. *Chem - A Eur J* **25**, 6864–6877.

Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. (2013). Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol Mol Biol Rev* **77**, 112–143.

Micallef, M. L., D'Agostino, P. M., Sharma, D., Viswanathan, R. & Moffitt, M. C. (2015). Genome mining for natural product biosynthetic gene clusters in the Subsection V *Cyanobacteria*. *BMC Genomics* **16**, 669.

Na, S. I., Kim, Y. O., Yoon, S. H., Ha, S. min, Baek, I. & Chun, J. (2018). UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol* **56**, 281–285.

Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutor Quant Methods Psychol* **4**, 13–20.

Nakamura, K., Oshima, T., Morimoto, T., Ikeda, S., Yoshikawa, H., Shiwa, Y., Ishikawa, S., Linak, M. C., Hirai, A. & other authors. **(2011).** Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res* **39**.

Nicolaou, K. C. & Harrison, S. T. (2007). Total synthesis of abyssomicin C, atrop-abyssomicin C, and abyssomicin D: Implications for natural origins of atrop-abyssomicin C. *J Am Chem Soc* **129**, 429–440.

Nicolaou, K. C. & Harrison, S. T. (2006). Total synthesis of abyssomicin C and atrop-abyssomicin C**. *Synthesis (Stuttg)*, 3256–3260.

Niu, X. M., Li, S. H., Görls, H., Schollmeyer, D., Hilliger, M., Grabley, S. & Sattler, I. (2007). Abyssomicin E, a highly functionalized polycyclic metabolite from *Streptomyces* species. *Org Lett* **9**, 2437–2440.

Nouioui, I., Carro, L., García-López, M., Meier-Kolthoff, J. P., Woyke, T., Kyrpides, N. C., Pukall, R., Klenk, H.-P., Goodfellow, M. & Göker, M. (2018). Genome-based taxonomic classification of the phylum *Actinobacteria*. *Front Microbiol* **9**, 1-119.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H. & other authors. **(2001).** Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc Natl Acad Sci U S A* **98**, 12215–20.

Penn, K., Jenkins, C., Nett, M., Udwary, D. W., Gontang, E. A., McGlinchey, R. P., Foster, B., Lapidus, A., Podell, S. & other authors. **(2009).** Genomic islands link secondary metabolism to functional adaptation in marine *Actinobacteria*. *ISME J* **3**, 1193–203.

Ridley, C. P., Ho, Y. L. & Khosla, C. (2008). Evolution of polyketide synthases in bacteria. *Proc Natl Acad Sci U S A* **105**, 4595–4600.

Riedlinger, J., Reicke, A., Zähner, H., Krismer, B., Bull, A. T., Maldonado, L. A., Ward,

A. C., Goodfellow, M., Bister, B. & other authors. **(2004).** Abyssomicins, inhibitors of the para-aminobenzoic acid pathway produced by the marine *Verrucosispora* strain AB-18-032. *J Antibiot* **57**, 271–279.

Rutledge, P. J. & Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **8**, 509-23.

Sadaka, C., Ellsworth, E., Robert Hansen, P., Ewin, R., Damborg, P. & Watts, J. L. (2018). Review on abyssomicins: Inhibitors of the chorismate pathway and folate biosynthesis. *Molecules* **23**.

Seipke, R. F., Kaltenpoth, M. & Hutchings, M. I. (2012). *Streptomyces* as symbionts: an emerging and widespread theme? *FEMS Microbiol Rev* **36**, 862–876.

Sellstedt, A. & Richau, K. H. (2013). Aspects of nitrogen-fixing *Actinobacteria*, in particular free-living and symbiotic *Frankia*. *FEMS Microbiol Lett* **342**, 179–186.

Song, Y., Li, Q., Qin, F., Sun, C., Liang, H., Wei, X., Wong, N. K., Ye, L., Zhang, Y. & other authors. **(2017).** Neoabyssomicins A–C, polycyclic macrolactones from the deep-sea derived *Streptomyces koyangensis* SCSIO 5802. *Tetrahedron* **73**, 5366–5372.

Trujillo, M. E., Alonso-Vega, P., Rodríguez, R., Carro, L., Cerda, E., Alonso, P. & Martínez-Molina, E. (2010). The genus *Micromonospora* is widespread in legume root nodules: the example of *Lupinus angustifolius*. *ISME J* **4**, 1265–1281.

Tsunakawa, M., Tenmyo, O., Tomita, K., Naruse, N., Kotake, C., Miyaki, T., Konishi, M. & Oki, T. (1992). Quartromicin, a complex of novel antiviral antibiotics. I. Production, isolation, physico-chemical properties and antiviral activity. *J Antibiot (Tokyo)* **45**, 180–8.

Tu, J., Li, S., Chen, J., Song, Y., Fu, S., Ju, J. & Li, Q. (2018). Characterization and heterologous expression of the neoabyssomicin/abyssomicin biosynthetic gene cluster from *Streptomyces koyangensis* SCSIO 5802. *Microb Cell Fact* **17**, 1–14.

Vieweg, L., Reichau, S., Schobert, R., Leadlay, P. F. & Süssmuth, R. D. (2014). Recent advances in the field of bioactive tetronates. *Nat Prod Rep* 1554–1584.

Wang, Q., Song, F., Xiao, X., Huang, P., Li, L., Monte, A., Abdel-Mageed, W. M., Wang, J., Guo, H. & other authors. **(2013).** Abyssomicins from the South China Sea deep-sea sediment *Verrucosispora* sp.: Natural thioether michael addition adducts as antitubercular prodrugs. *Angew Chemie - Int Ed* **52**, 1231–1234.

Wang, X., Elshahawi, S. I., Cai, W., Zhang, Y., Ponomareva, L. V., Chen, X., Copley, G. C., Hower, J. C., Zhan, C.-G. & other authors. **(2017).** Bi- and tetracyclic spirotetronates from the coal mine fire isolate *Streptomyces* sp. LC-6-2. *J Nat Prod* **80**, 1141–1149.

Wei, R.-B., Xi, T., Li, J., Wang, P., Li, F.-C., Lin, Y.-C., Qin, S., Wei, R.-B., Xi, T. & other authors. **(2011).** Lobophorin C and D, new kijanimicin derivatives from a marine sponge-

associated *Actinomycetal* strain AZS17. *Mar Drugs* **9**, 359–368.

Weixin, T., Manghong, Z., Deng, Z. & Yuhui, S. (2013). Biosynthesis of tetronate antibiotics: A growing family of natural products with broad biological activities. *Sci China Chem* **56**, 1364–1371.

Ye, Z., Musiol, E. M., Weber, T. & Williams, G. J. (2014). Reprogramming acyl carrier protein interactions of an acyl-CoA promiscuous trans-acyltransferase. *Chem Biol* **21**, 636–646.

Zhang, J., Li, B., Qin, Y., Karthik, L., Zhu, G., Hou, C., Jiang, L., Liu, M., Ye, X., & other authors (2020). A new abyssomicin polyketide with anti-influenza A virus activity from a marine-derived *Verrucosispora* sp. MS100137. *Appl Microbiol Biotechnol* **4**, 1533-1543.

Zeng, Y., Feng, F., Medováa, H., Dean, J. & Koblízeka, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum *Gemmatimonadetes*. *Proc Natl Acad Sci* **111**, 7795–7800.

Ziemert, N., Lechner, A., Wietz, M., Millañ-Aguiñaga, N., Chavarria, K. L. & Jensen, P. R. (2014). Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci U S A* **111**.

Chapter 3. Heterologous expression and production of abyssomicins.

Abstract

Publically available genomic and metagenomic data are used daily to carry out a variety of analysis such as those presented in Chapter 2 on the evolution and environmental distribution of the abyssomicin BGC. All this raw data allows the identification of strains with a specific biosynthetic potential, however, the probability that these strains produce a target secondary metabolite under standard laboratory conditions is very low. For the last few decades, to exploit the sometimes unaccessible reservoir of BGCs present in bacterial genomes, heterologous hosts that facilitate the activation and production of secondary metabolic pathways have been used. Through heterologous production, it is possible to access the structural and functional diversity of compounds encoded in the genomes of strains that do not produce them under standard laboratory conditions or are recalcitrant to genetic manipulation. Furthermore, it can afford production yields that allow the compounds to be exploited in other fields such as medicine or agriculture. In the present work, various attempts to express the abyssomicin BGC from *M. maris* AB-18-032 in several heterologous hosts are presented. First, a systematic analysis of heterologous hosts typically used for the production of natural products from *Actinobacteria* was carried out. Then, *E. coli* BAP1 and various *Streptomyces* species were transformed with an *aby* BGC-containing phage artificial chromosome and two different approaches, ribosome engineering and CRISPR-Cas9 mediated promoter knock-in, were followed to activate abyssomicin production. Despite these efforts, it was not possible to identify any abyssomicin from the fermentations of the heterologous hosts. Active transcription of the genes was confirmed. Genome sequencing of one heterologous host revealed four points along *abyB1* where mutations and deletions may have happened. This highlights the importance of strain re-sequencing after engineering and opens the possibility of repairing those mutations to restore abyssomicin production in the heterologous hosts.

3.1 Introduction.

3.1.1 Heterologous production of natural products.

Bacteria, and more specifically, bacteria from the phylum *Actinobacteria* constitute a rich source of natural products. Unlike synthetic chemical compounds, biologically synthesised natural products exhibit a wider and sometimes more complex structural diversity that results in their specific and often inimitable biological activities (Davison & Brimble, 2019). Indeed,

most BGCs responsible for the production of these natural products remain uncharacterised and the corresponding products untested for potential bioactivities, not only for the development of new drugs but also for potential agricultural products (Myronovskyi & Luzhetskyy, 2019). Since conventional drug discovery approaches result in a high rediscovery rate of known compounds, current efforts rely on the application of bioinformatic and genetic tools to exploit the BGCs present in *Actinobacteria* genomes. In order to overcome the inherent limitations of working with microorganisms recalcitrant to standard laboratory techniques, it is necessary to use other more amenable strains that allow for genetic manipulation to activate and enhance the production of the compound of interest as well as non-natural derivatives of this (Cook & Pfleger, 2019).

Using heterologous hosts for the discovery and production of novel natural products depends on the ability to capture and express putative BGCs (Nah *et al.*, 2017). In order to maximise the chances of success, it is important to take into account genomic features such as the GC content, the codon usage and the ability of the heterologous host to produce all the necessary metabolite precursors. To facilitate these needs, various research projects during the last two decades have focused on developing heterologous expression workhorses, from *E. coli* (Pfeifer *et al.*, 2001) to *Actinobacteria* such as *Streptomyces coelicolor* (Gomez-Escribano & Bibb, 2011), *Streptomyces lividans* (Ziermann & Betlach, 1999; Ahmed *et al.*, 2020) and *Streptomyces albus* (Myronovskyi *et al.*, 2018; Fazal *et al.*, 2019). These strains, together with *Pseudomonas putida* (Choi *et al.*, 2018), *Myxococcus xanthus* (Yan *et al.*, 2018) and *Synechococcus elongatus* (Roulet *et al.*, 2018) have successfully been used for the production of a wide variety of bacterial natural products, whilst *Saccharomyces cerevisiae* (Billingsley *et al.*, 2016) and species of the genera *Aspergillus* and *Penicillium* (Alberti *et al.*, 2017) have mainly been used for the production of plant and fungal natural products. Among many others, some important achievements in this field include the production of erythromycin in *E. coli* (Pfeifer *et al.*, 2001), artemisinic acid in *S. cerevisiae* (Ro *et al.*, 2006), epothilones in *M. xantus* (Julien & Shah, 2002) and *E. coli* (Mutka *et al.*, 2006), daptomycin in *S. lividans* (Penn *et al.*, 2006) and tetracycline in *S. lividans* (Wang *et al.*, 2012).

Despite notable successes in the field and the new knowledge generated, the heterologous production of secondary metabolites still faces many challenges. The phylogenetic distance between the natural producer and the heterologous host is a key factor to consider and from it depends that the heterologous host recognises the regulatory elements associated to the BGC of interest and is capable of expressing it. Then, refactoring of promoters and regulatory

elements may be necessary, and production can still be affected by the number of copies and the size of the BGC, the BGC position in the genome, the number of BGCs already in the heterologous host and the availability of cellular resources (Huo *et al.*, 2019). Despite taking all those factors into account, there are no guarantees of success and, due to "negative results" going unpublished, it is not possible to calculate the real success percentage of heterologous production projects. Only a few papers have reported that heterologous production was not achieved, for example, epothilone biosynthesis in *E. coli* 10 β (Mutka *et al.*, 2006), *M. xanthus* DK1622 and *P. putida* KT2440 (Fu *et al.*, 2008). Similarly, the expression of the triketide lactone BCG from *Mycobacterium ulcerans* was not achieved in the faster-growing *Mycobacterium marinum* (Porter *et al.*, 2013), although in all these cases, the communication of unsuccessful production was accompanied by the successful production achieved in other hosts. In other cases, "failed" heterologous production was reported prior to refactoring the BGC, but again these published results were always presented together with the successful production results after refactoring (Aso *et al.*, 2004; Alberti *et al.*, 2018; Kallscheuer *et al.*, 2019).

3.1.2 Heterologous expression of **Micromonospora** *biosynthetic gene clusters.*

Micromonospora species account for the second most prolific source of natural products with biomedical applications after the genus *Streptomyces* (Barka *et al.*, 2016; Carro *et al.*, 2019). However, contrary to *Streptomyces* species, it has not been until recently that some molecular tools specific for *Micromonospora* have been developed (Wang *et al.*, 2020), and most of these, still remain species specific. To date, five research projects have reported to heterologously express *Micromonospora* BGCs. First, *S. albus* J1074 and *S. lividans* TK21 were used to produce the antitumor thiocoraline from *Micromonospora* sp. ML1 (Lombó *et al.*, 2006). Researchers attempted to express the thiocoraline BGC with and without the regulator gene *tioA* located under the control of the *ermE* promoter, and only under the presence of the *ermE* promoter, production was achieved in both hosts. After that, another study reported the production of high levels of mycarosyl-erythronolide B after expressing the L-mycarose BGC from *Micromonospora megalomicea* in *E. coli* BL21 (DE3) (Peirú *et al.*, 2007). In this case, expression of the proteins was carried out in standard *E. coli* expression vectors and controlling the expression with inducible promoters. A third paper, reported how different sets of genes from *Micromonospora echinospora* necessary for the biosynthesis of gentamicin and its intermediates were heterologously expressed in *Streptomyces venezuelae* YJ003 (Je *et al.*, 2008). In this case the genes were maintained in *S. venezuelae* via replicative plasmids and any

kind of optimisation such as codon optimisation or promoter engineering was not needed to achieve production. More recently, the fluostatin BGC from *Micromonospora rosaria* SCSIO N160 was expressed in *S. coelicolor* YF11 and led to the discovery of new fluostatin analogues (Yang *et al.*, 2015). The fluostatin BGC was integrated into *S. coelicolor* YF11 chromosome via phiC31 site-specific recombination and no refactoring was needed for the production of these compounds. Finally, the selection of a putative ClpC-containing BGC from *Micromonospora* sp. strain B006, its cloning, promoter exchange, and heterologous expression in *S. coelicolor* M1152, led to the production of a compound yet to be identified (Braesel & Eustáquio, 2019). The BGC was directly captured from genomic DNA via TAR cloning and the *ermE*p* promoter was used to drive transcription.

This study presents an attempt to express the *M. maris* AB-18-032 abyssomicin BGC in various heterologous hosts. First, in order to carefully select the most adequate expression chassis, a systematic analysis of the research previously developed in the field was carried out. From there, the strains that had been used more frequently for the expression of bacterial polyketide natural products were selected and their codon usage was compared to that in the *aby* BGC. Then, *E. coli* BAP1 and various *Streptomyces* hosts were transformed with an *aby* BGC-containing vector and the production profile of the strains when grown in solid and liquid media was analysed via antimicrobial plate diffusion assays and LC-MS. Two different approaches, ribosome engineering and CRISPR-Cas9 mediated promoter knock-in, were used to activate abyssomicin production and the engineered strains were rescreened for abyssomicin production. Finally, a gene expression analysis was carried out and one of the engineered heterologous host was re-sequenced.

3.2 Materials and methods.

3.2.1 Chassis selection.

In order to find the heterologous hosts that would present *a priori* more chances of successful production of abyssomicins, a consistent review of the literature published between 2006-2016 was carried out to determine which chassis have been used for the heterologous expression of bacterial polyketide natural products more fruitfully.

Moreover, in order to predict whether a chassis would be suitable for the expression of the *M. maris* AB-18-032 abyssomicin BGC, the codon usage of the *aby* BGC was compared against the whole genome codon usage of *S. coelicolor*, *S. albus* and *S. lividans*. *M. maris* genome was also included in the analysis to determine if the codon usage of the *aby* BGC followed that of the rest of the genome. Finally, the codon usage in the *aby* BGC was compared against the codon usage in the abyssomicin BGC present in *Streptomyces koyangensis* (*abm*). Since the objective of including the abyssomicin BGC present in *Streptomyces koyangensis* (*abm*) was just to check and compare its codon usage against that of the heterologous host used for its heterologous expression and the codon usage in the *aby* BGC (Li *et al*., 2018), the whole genome codon usage of *S. koyangensis* was not included in this analysis. None of the online tools for codon usage calculation had in their database the genomes of *M. maris*, *S. coelicolor*, *S. albus* and *S. lividans*, thus, codon usage calculations were carried out as follows. First, one python script was created to modify the structure in which the genes are presented when a file containing the genomic DNA sequence of an organism is downloaded from NCBI database. This script puts all the genes within the genome in a single line (Figure S3.1). Then, a second python script divides each gene into its codons, creating a list with all the codons, that then, is used to count how many times each codon appears and to calculate the corresponding abundance in per mil (Figure S3.2).

3.2.2 Transformation of the abyssomicin BGC in E. coli *BAP1 and* Streptomyces *species.*

The mobilisation plasmid pTAMob (Strand *et al.*, 2014) was purified form *E. coli* 10β using the GenElute™ HP Plasmid Maxiprep Kit (Sigma). This vector (52.7 kb) was transferred by electroporation into *E. coli* Top10 strain carrying the phage artificial chromosome pESAC13_*aby* containing the entire abyssomicin BGC as described before (Tu *et al.*, 2016). Correct electrotransformants were checked by colony PCR using primers that amplify the gentamycin resistance gene of pTAMob (Table S3.1). *E. coli* Top10 carrying pESAC13_*aby*

and pTAMob was used as parental strain for conjugations carried out with *E. coli* BAP1 (Pfeifer *et al.*, 2001) as recipient strain. Cultures of *E. coli* Top10 harbouring pESAC13_*aby* and pTAMob were grown at 37 °C overnight in LB containing 20 µg/ml gentamycin and 50 µg/ml kanamycin and *E. coli* BAP1 was grown at 37 °C overnight in LB. Cultures were then washed three times with LB without antibiotics and pellets were resuspended in 100 µl of LB. The pellets of both cells were combined and drops containing 5 µl of this mixture were placed onto LB plates without antibiotics to be incubated overnight at 37 °C. On the next day, biomass from the grown spots was streaked onto LB plates containing 50 µg/ml kanamycin. After overnight incubation at 37 °C, exconjugants were checked, by colony PCR, first, for the presence of the *sfp* gene present in *E. coli* BAP1 and absent in *E. coli* Top10. Then, the *E. coli* BAP1 exconjugants were checked for the presence pESAC13_*aby* by colony PCR of five different genes present in the *aby* BGC. All the primers used are listed in Table S3.1.

E. coli Top10 containing pESAC13_*aby* and pTAMob (Strand *et al.*, 2014) was used for mating with *S. lividans* K4-144, *S. albus* J1074 and *S. albus* S4 Δ5, *S. coelicolor* M1145 and *S. coelicolor* M1152, *Streptomyces* sp. FR-008 LQ3 and *Salinispora tropica* CNB-4401 as described before (Kieser *et al.*, 2000). When exconjugants appeared, these were transferred to new MS agar plates supplemented with 10 mM $MgCl_2$, 20 mg nalidixic acid and 1.5 mg thiostrepton. Genomic DNA was extracted (GenElute Bacterial Genomic DNA Kit; Sigma-Aldrich) and PCRs on *abyU* and the *tsr* thiostrepton resistance gene were carried out to verify the presence of pESAC13_*aby*. For the strains that were not successfully transformed using this method, two other conjugation procedures were tested. First, the classic intergeneric conjugation protocol between *E. coli* ET12567/pUZ8002 and methylation-restrictive *Streptomyces* was used as described before (Kieser *et al.*, 2000) but with some variations that include using recipient cells in the form of spores and/or mycelial fragments, different ratios of recipient and/or donor cells and different concentrations of $MgCl_2$ and/or $CaCl_2$ ranging from 10 to 100 mM. Finally, biparental mating using the helper strain *E. coli* pR9604 and *E. coli* Top10 pESAC13_*aby* as donor was also carried out (Jones *et al.*, 2013). Briefly, overnight cultures of the helper and donor strains were harvested and washed twice with LB, pellets were combined for mating and incubated overnight in 5 µl spots in LB plates. Biomass was streaked and cultured overnight in LB plates containing 50 µg/ml kanamycin and 50 µg/ml carbenicillin. The resulting *E. coli* pR9604 pESAC13_*aby* was used for intergeneric conjugations as before (Kieser *et al.*, 2000).

3.2.3 Abyssomicin production assays in E. coli *BAP1 and* Streptomyces *species.*

Abyssomicin production assays in *E. coli* BAP1 were carried out in liquid media as follows. Fermentations of *E. coli* BAP1 and *E. coli* BAP1 pESAC13_*aby* at 16 °C and 37 °C were carried out in A1 media (2% soluble starch, 1% glucose, 1% malt extract, 0.5% corn flour, 0.4% yeast extract, 0.2% peptone, 3% sea salt, in deionized water, pH 7.3) and SGG media (Gottardi *et al*., 2011) containing 50 µg/ml kanamycin. 250 ml flasks containing 50 ml of media were used for 5-days fermentations at 200 rpm where each day 5 ml of culture were extracted for subsequent analysis. Samples were adjusted to pH 4 using HCl, mixed 1:1 with ethyl acetate and incubated at room temperature and 250 rpm for 2 hours. Tubes were centrifuged at 4,000 rpm for 15 minutes and the upper organic phase was transferred into test tubes for evaporation at 22 °C and 10 psi (TurboVap® LV, Biotage). Pellets were resuspended in 200 µl of methanol and stored at -20 °C. Frozen extracts were diluted 1:9 in ddH_2O and centrifuged twice at 13,200 rpm for 15 minutes in preparation for chemical analysis. All supernatants were analysed in a mass spectrometer ESI microTOF-Q II (Bruker Daltonics). Separation of the compounds was performed by a Raptor™ ARC-18 (2.7 µ, 150 x 2.1 mm, Restex) chromatographic column protected by a guard column. The mobile phase was ultrapure water (solvent A) and acetonitrile (solvent B), both with 0.1% formic acid (v/v). The elution gradient was: 0–55 min 20-100% B; 55-59 min 100% B; 59-60 min 100-20% B, 60-75 min 20% B followed by column washing and reconditioning (1 minute). The flow rate was 0.2 mL/min. The column oven was at 35 °C and the injection volume was 3 µL. The UV analyses were performed in the wavelength range of 190–600 nm, with the mass spectrometer operating in positive mode (m/z 50–2800). Neither bioassays of the fermentation extracts nor plug assays from plates were carried out, as *E. coli* BAP1 carries the *aby* BGC in a pESAC backbone that is maintained under the presence of kanamycin.

Similarly, fermentations of *S. albus*, *S. lividans* and *S. coelicolor* transformed and untransformed with pESAC13_*aby* were carried out in SGG and A1 at 30 °C and 200 rpm in 250 ml flasks containing 50 ml of media. For *S. albus* and *S. lividans*, four more production media were used: A1Sal* (Zhang *et al*., 2018), R5A (Kieser *et al*., 2000), SM10 (Malmierca *et al*., 2018) and SM17 (Malmierca *et al*., 2018). Extractions and LC-MS were carried out as described for *E. coli* BAP1. Moreover, the fermentation extracts were used to perform antimicrobial bioactivity assays against *B. subtilis*. CDM agar (0.5% glucose, 0.05% tri-sodium citrate•$2H_2O$, 0.3% KH_2PO_4, 0.7% K_2HPO_4, 0.01% $MgSO_4$•$7H_2O$, 0.1% $(NH_4)_2SO_4$, 1.5% agar, in deionized water, pH 7.0) (Riedlinger *et al.*, 2004) plates with and without 5 mM PABA

were prepared and inoculated following a 1:100 ratio with *B. subtilis* cultures adjusted to an OD_{600}=1.00. 0.6 mm (diameter) filter paper disks were cut using a paper drill and put on top of the agar plates. Frozen extracts were diluted 1:9 in ddH_2O and 10 µl were put onto the filter paper disks. Extracts of *M. maris* AB-18-032 were used as positive control, antimicrobial susceptibility disks of nalidixic acid (Oxoid™, ThermoFisher) were used as negative control to test the effect of PABA and filter disks with 10 µl of 10% methanol were used as negative control. After overnight incubation at 37 ºC, the ZOI generated by each extract was measured.

S. albus, *S. lividans* and *S. coelicolor* transformed and untransformed with pESAC13_*aby* were grown in SGG agar, A1 agar, A1Sal* agar, R5A agar, SM10 agar and SM17 agar plates at 30 ºC for 15 days. 0.6 cm (diameter) agar plugs from each strain grown in each production media were placed into *B. subtilis* inoculated CDM agar plates. Plugs of *M. maris* AB-18-032 were used as positive control and antimicrobial susceptibility disks of nalidixic acid (Oxoid™, ThermoFisher) were used as negative control to test the effect of PABA. After overnight incubation at 37 ºC, the ZOI of each plug was measured.

3.2.4 Ribosome engineering of the Streptomyces *heterologous hosts.*

Spores of *S. albus*, *S. lividans* and *S. coelicolor* wild-type and harbouring pESAC13_*aby* were spread on Tryptic Soy Agar (TSA) containing three different concentrations of streptomycin or rifampin (5, 10 and 30 µg/ml). After cultivation for 15 days at 30 ºC, spontaneous streptomycin-resistant (*str*) and rifampin-resistant (*rif*) mutants appeared. Isolated colonies were grown in SGG agar for 15 days at 30 ºC. In order to test for enhanced antimicrobial activity, 0.6 cm (diameter) agar plugs from each plate were placed onto Mueller Hinton (MH) agar plates with and without 5 mM PABA inoculated with *Bacillus subtilis* DSM 10 at OD_{600}=1.00 following a 1:100 ratio. After overnight incubation at 37 ºC, the ZOI in mm of each plug was measured.

3.2.5 CRISPR-Cas9 mediated promoter knock-in in S. albus *and* S. lividans.

In order to drive abyssomicin production in *Streptomyces*, the strong constitutive *Streptomyces* promoter *kasOp** was introduced via CRISPR upstream the natural promoter of the activator AbyI. For that, a *kasOp** cassette with adapter arms containing restriction sites was synthesised, digested with XbaI, purified and quantified. The CRISPR vector pCRISPomyces-2 was digested with XbaI, dephosphorylated, purified, quantified and ligated to the *kasOp** cassette. The resulting vector, pCM2, was transformed into *E. coli* DH5α and the obtained clones were checked by PCR and sequencing. Then, 2 kb homology arms were cloned in two steps. First, a

2 kb downstream homology arm was amplified from pESAC13_*aby* to be inserted into the SpeI site of pCM2 by Gibson assembly (New England Biolabs, E2611S). Then, the plasmid from correct transformants was digested with HindIII and a 2 kb upstream homology arm amplified from pESAC13_*aby* was cloned by Gibson assembly. The resulting vector was transformed into *E. coli* DH5α and verified by sequencing before transforming the conjugative strain *E. coli* ET12567. *S. albus* pESAC13_*aby* and *S. lividans* pESAC13_*aby* exconjugants containing pCRISPomyces-2_*kasOp** were obtained using the classic intergeneric conjugation between *E. coli* ET12567 and methylation-restrictive *Streptomyces* (Kieser *et al.*, 2000). Exconjugants were verified by PCR after carrying out genomic extractions. All the primers used for cloning and verification can be found in Table S3.1. In order to induce plasmid loss, single colonies were subsequently streaked and grown in TSA plates at 37 °C until single colonies appeared for 10-15 days. Abyssomicin production by *S. albus* pESAC13_*aby_kasOp** and *S. lividans* pESAC13_*aby_kasOp** was assessed in plates and in liquid cultures using different production media as described before.

3.2.6 Gene expression analysis.

Samples for RNA extractions were obtained from 24 hours liquid cultures as follows. For each strain, 3 ml of culture were divided into three Eppendorf tubes and placed immediately in ice. One millilitre of RNAprotect Bacteria Reagent (Qiagen Cat No./ID: 76506) was added to all samples, followed by vortex mixing for 5 seconds. Samples were left at room temperature for 5 minutes, centrifuged for 5 minutes at 5,000 x *g* and the supernatant was decanted completely using filter paper to absorb any liquid. Pellets were put on ice, followed by freezing and storage at -80 °C until the extractions were carried out. For the lysis and homogenisation of the mycelium, the pellets were defrosted and resuspended by pipetting in 150 µl of 30 mg/ml lysozyme in TE buffer. Samples were incubated at room temperature for 10 minutes with mixing by inversion every two minutes. Six hundred microlitres of a RLT-β-ME (2-Mercaptoethanol, for molecular biology, minimum 98% GC/titration. Sigma M3148-25 ml) solution was added to each tube followed by vortex mixing. The mixture was then transferred to a Fastprep tube (Deltalab: 409007.N, 409115/2, #03961-1-104) that had been pre-cooled in ice. Tubes were shaken in a FastPrep-24 at level 6 for 20 seconds, placed in ice for 1 minute, shaken for 20 seconds and placed in ice for 1 minute. To pre-purify the lysate, two PLG tubes (5PRIME Phase Lock Gel Heavy 2ml - 200. Quanta BioSciences cat# 2302830) per sample were pre-packaged by centrifugation at maximum speed for 30 seconds; 750 µl of homogenised sample was transferred to PLG tubes followed by addition of 375 µl of Aqua Phenol at pH 5.0

and 375 µl of CIA (24 ml chloroform and 1 ml of isoamyl alcohol). The mixture was mixed vigorously for 30 seconds and centrifuged at room temperature for 5 minutes at 13,000 x *g*. The upper phase was decanted into a new PLG tube and a further 750 µl of Aqua Phenol-CIA was added. The mixture was mixed vigorously for 30 seconds and centrifuged at room temperature for 5 minutes at 13,000 x *g*. The upper phase was decanted into a new 2 ml tube and mixed by inversion with 395 µl of absolute ethanol. RNA was then purified using the RNeasy Mini Kit (Qiagen ID:74104) following the manufacturer's instructions.

Prior to carrying out RT-PCRs, the extracted RNA was digested with DNaseI using the TURBO DNA-free Kit (Ambion #AM1907) as follows: 50 µl of extracted RNA were mixed gently with 5 µl of 10X DNase I buffer and 1 µl of rDNase I and incubated in a thermocycler at 37 °C for 30 minutes. One microlitre of rDNase I was added again and the samples were further incubated at 37 °C for 30 minutes. After incubation, 12 µl of DNase Inactivation Reagent were added and samples were incubated at room temperature for 3 minutes, mixing the samples by inversion every minute. Samples were centrifuged at 10,000 x *g* for 2 minutes and the supernatant was transferred into a new tube. In order to assess the quality and quantity of RNA, 7 µl of RNA were run in an agarose gel (1.5% gel, 90 V, 1 hour) and 1 µl of RNA was used for nanodrop quantification. Samples were stored at -80 °C.

In order to analyse gene expression, RT-PCR primers were designed for several genes in the *M. maris aby* BGC. Primers were also designed for the housekeeping gene *hrdB*. Primers were designed to amplify approximately 500 bp as close as possible to the 5' end of the gene of interest using OligoPerfect software (Thermofisher) and setting up the amplicon length between 300-700 bp, ideally 500 bp and Tm = 60 °C. For *abyA1* and *abyD* two different primer pairs were tested. All the primers can be found in Table S3.1. All primer pairs were tested by PCR using DreamTaq and genomic DNA as template. The phage artificial chromosome containing the *aby* BGC was used as positive control and H_2O was used for the negative control. PCR products were run in a 2% agarose gel at 90 V for 30 minutes.

For samples for which the amount of RNA was not enough to perform all the RT-PCR reactions, cDNA was synthesised using the BioRad iScript cDNA Synthesis kit following the manufacturer's instructions to be used as template for the RT-PCR reaction. RT-PCRs were carried out using the One Step RT-PCR SuperScript III system with Platinum Taq DNA polymerase following the manufacturer's instructions. Note that in this kit there were two different polymerases. One amplifies template cDNA whereas the other is a standard

polymerase that needs dsDNA as template and it is used as control to check that there is no dsDNA in the sample. PCR products were run in a 2% agarose gel at 90 V for 30 minutes.

3.2.7 Genome sequencing, assembly and annotation.

S. lividans K4-114 containing pESAC13_*aby_kasOp** was first sequenced using Illumina MiSeq (MicrobesNG). The resulting paired end reads had already been trimmed using FastP by the sequencing provider, so SPAdes was directly used using spades.py command with the --careful option to minimise the number of mismatches in the contigs, t4 to indicate the number of threads and the -k flag at 21, 33, 55, 77 (Bankevich *et al.*, 2012). QUAST was used with the SPAdes output file contigs.fasta to retrieve the assembly statistics (Gurevich *et al.*, 2013). Bandage was used to visualise the connections of the assembled contigs using the SPAdes output assembly_graph_with_scaffolds.gfa (Wick *et al.*, 2015).

In order to align the sequenced genome against a reference, the expected strain sequence was built *in silico* using the genome sequence of *S. lividans* TK24 (CP009124) and inserting pESAC13_*aby_kasOp** into the previously described *attB* site in the chromosome of *S. lividans* TK24 (Combes *et al.*, 2002). All the trimmed Illumina reads were mapped against pESAC13_*aby_kasOp** and *S. lividans* TK24 pESAC13_*aby_kasOp** using Geneious Prime 2020.1.

S. lividans K4-114 pESAC13_*aby_kasOp** was also sequenced using Nanopore technology. In order to merge all the reads generated by Nanopore sequencing, all the fastq files containing the reads were merged into one using cat *.fastq > combined.fastq. To trim the reads Filtlong was used filtering the read by minimum length 500 bp and keep percent 90. That reduced the reads from 1800 Mb to 1620 Mb. The resulting output from Filtlong was used as input for SPAdes. All the trimmed Nanopore reads were mapped against pESAC13_*aby_kasOp** and *S. lividans* TK24 pESAC13_*aby_kasOp** using Geneious Prime 2020.1.

All the Illumina and Nanopore reads were merged and aligned to pESAC13_*aby_kasOp** and *S. lividans* TK24 pESAC13_*aby_kasOp** using Geneious Prime 2020.1.

3.3 Results.

3.3.1 Chassis selection.

In order to select the heterologous hosts with higher chances of success for the expression and production of abyssomicins, a literature review comprising the last ten years (2006-2016) of research papers on the heterologous expression of polyketide biosynthetic gene clusters was carried out (Table S3.2). This search was specific for bacterial polyketide natural products produced in bacterial heterologous hosts, and thus, polyketide natural products from fungi, plants or marine sponges for example, were not considered. Equally, polyketide natural products heterologously produced in fungi were not analysed.

During the period comprising 2006-2016, an average of seven studies were published per year on the heterologous production of bacterial polyketide natural products (Figure 3.1). This might seem a small number, however, it should be highlighted that studies on the heterologous production of non-polyketide natural products, non-bacterial polyketide natural product and studies using non-bacterial heterologous hosts were not taken into account. Moreover, in order to avoid repetition, different studies regarding the same natural product produced in the same heterologous host were counted just once. For example, numerous papers were published around the heterologous production of erythromycin in *E. coli* BAP1, in this case only the first paper published was considered. On the contrary, when a natural product was produced in a first heterologous host and years later in a second, both of them were considered.

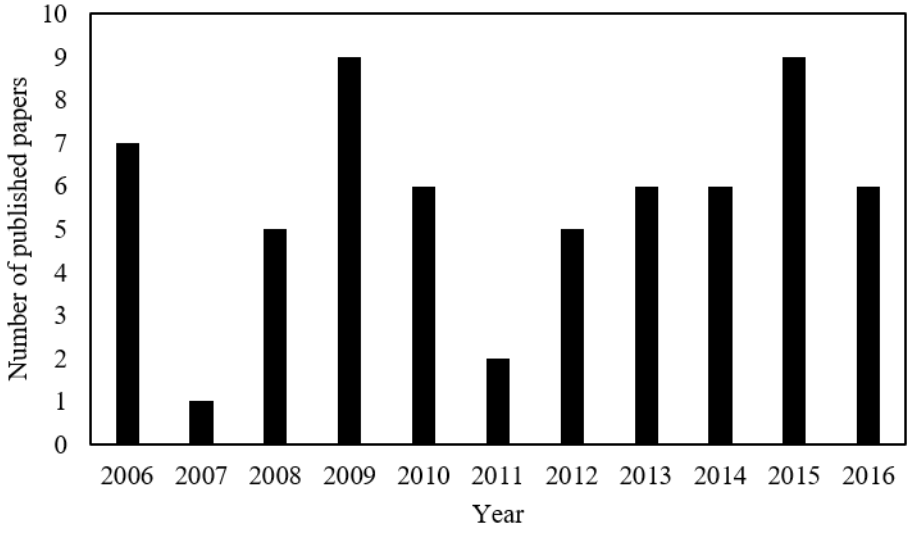

Figure 3.1. Number of papers published between 2006-2016 on the production of bacterial polyketide natural products in bacterial heterologous hosts.

From this literature review, it was possible to find that *S. lividans*, *S. coelicolor*, *E. coli* and *S. albus* were the preferred microorganisms for the heterologous production of bacterial polyketide natural products (Figure 3.2A). More specifically, *S. lividans* K4-114 and *S. lividans* TK23 were the most frequently selected *S. lividans* strains (Figure 3.2B), *S. coelicolor* CH999 and *S. coelicolor* M512 were the most used *S. coelicolor* strains (Figure 3.2C) and *E. coli* BAP1 followed by *E. coli* BL21 (DE3) were the preferred *E. coli* hosts (Figure 3.2D). Interestingly, most of the articles reporting *S. albus* as heterologous host did not mention a particular strain and those that did always reported the wild-type strain *S. albus* J1074 (Figure 3.2E).

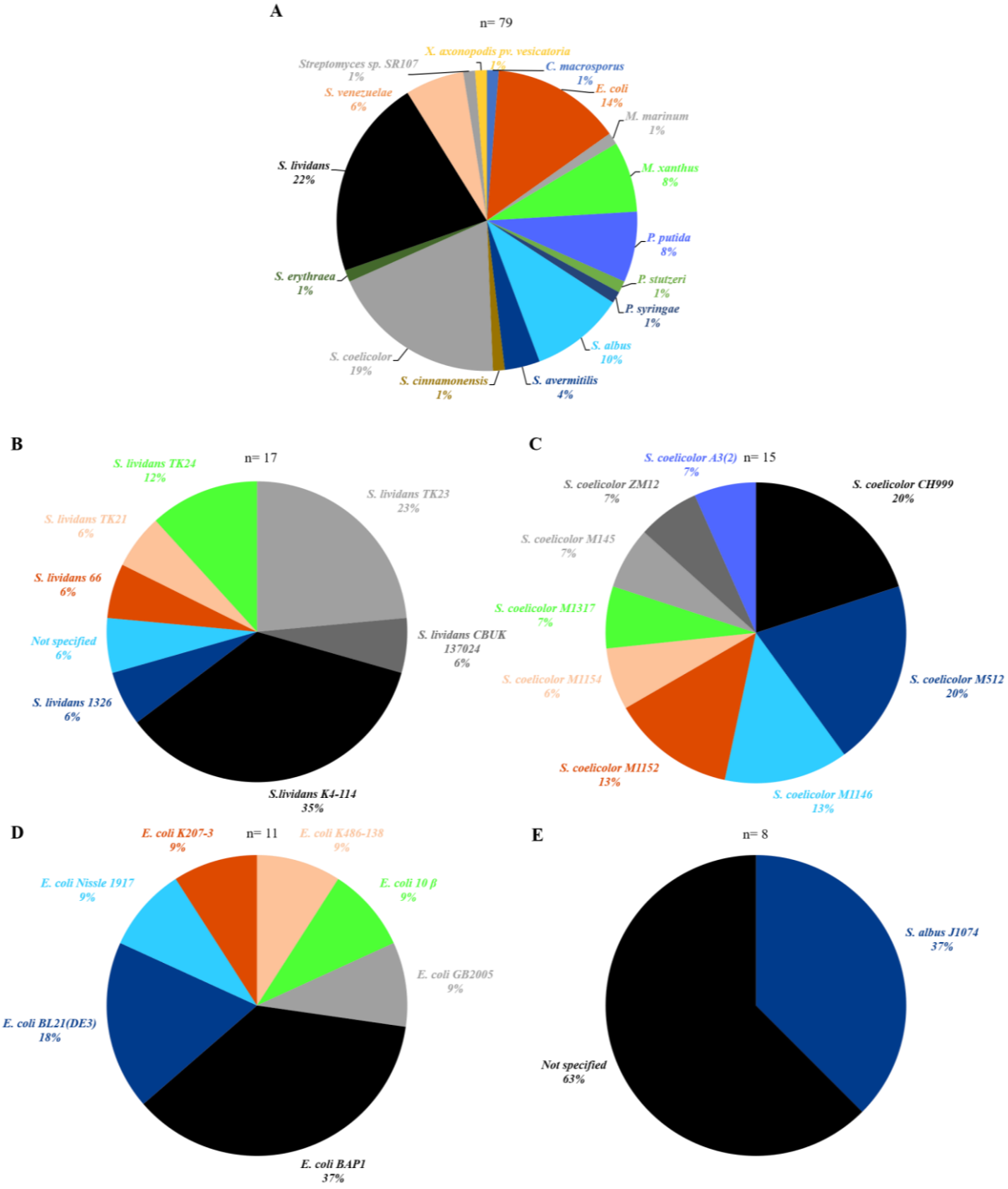

Figure 3.2. A) Heterologous hosts used for the production of bacterial polyketide natural products during 2006-2016 as reported in the literature. **B)** *S. lividans* strains used in those studies. **C)** *S. coelicolor* strains used in those studies. **D)** *E. coli* strains used in those studies. **E)** *S. albus* strains used in those studies.

In order to further analyse the viability of using a *Streptomyces* host to express the *M. maris* AB-18-032 *aby* BGC, the general codon usage of *M. maris* AB-18-032, *S. lividans* TK24, *S. coelicolor* A3(2), *S. albus* J1074 and the *aby* BGC was analysed (Figure 3.3). Generally, this analysis showed that the codon usage followed a similar pattern in all the strains and the *aby* BGC. Only a few codons were shown to be more abundant in the *aby* BGC than in the possible

heterologous hosts. This would be the case of the codon ACC (Thr) that appears more frequently in the *aby* BGC than in *S. lividans* and *S. coelicolor* and the codon CTG (Leu) that is especially abundant in the *aby* BGC even in comparison with the general codon usage of *M. maris*. To a lesser extent, other less frequently used codons such as CCA (Pro), CAA (Gln), CGA (Arg) and TTG (Leu) also seemed to be more predominant in *M. maris* and the *aby* BGC than in the *Streptomyces* genomes.

The codon usage of the abyssomicin BGC present in *S. koyangensis* SCSIO 5802 was also included in this analysis, as heterologous expression of the *abm* BGC was reported in *S. coelicolor* M1152 (Song *et al.*, 2017; Tu *et al.*, 2018). The codon usage in the *abm* BGC was similar to the codon usage in the *aby* BGC, with the exception of the codons AAG (Lys), AGG (Arg), CTC (Leu), GAG (Glu) and TCC (Ser) that were more abundant in the *abm* and CTG (Leu) and CAC (His) that were more abundant in the *aby* (Figure 3.3). Since the main objective behind including the codon usage of the *abm* BGC was to compare it against the codon usage of the *aby* BGC and the codon usage of the *abm* BGC followed the general codon usage of the *Streptomyces* heterologous hosts, the codon usage of *S. koyangensis* SCSIO 5802 was not included in this analysis.

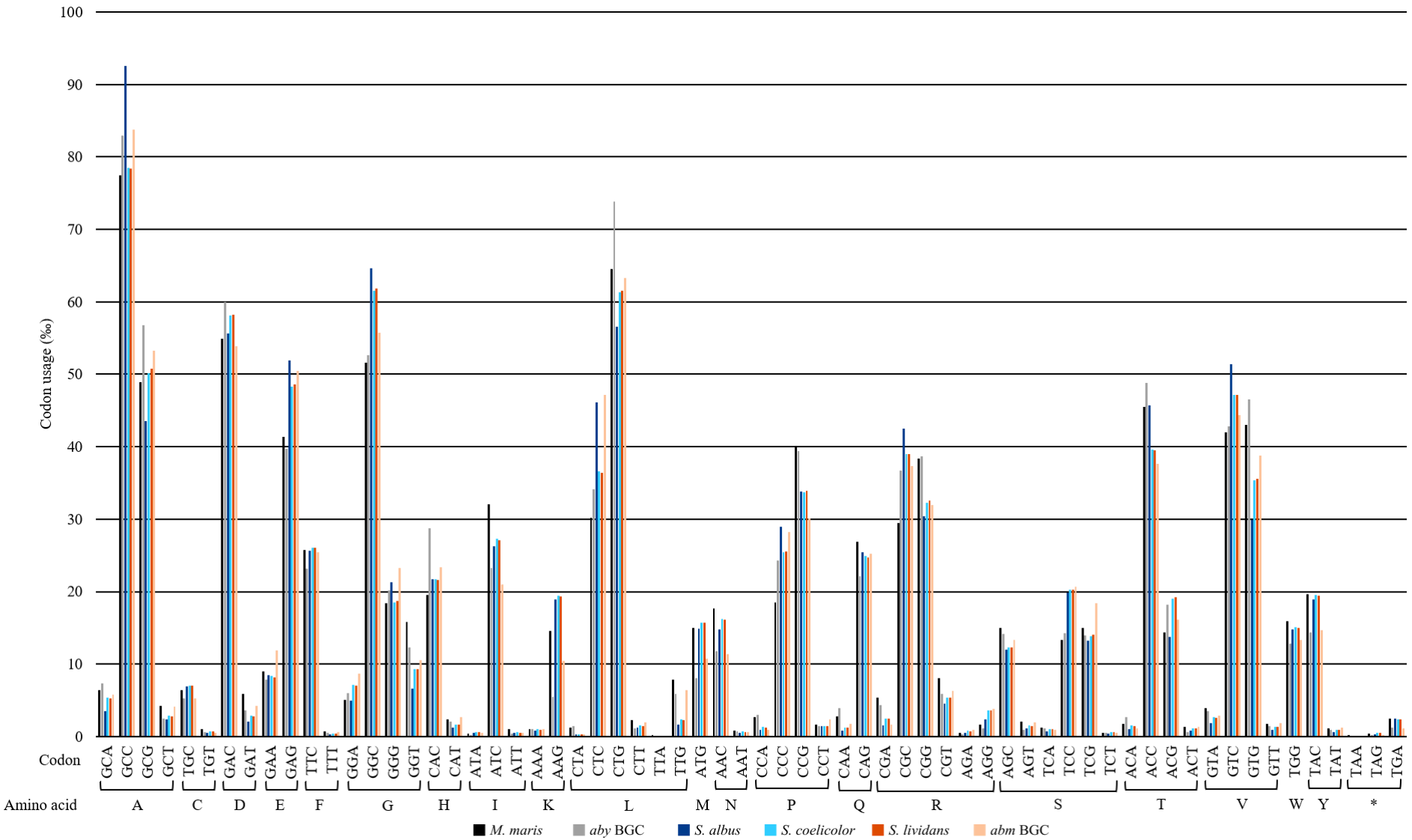

Figure 3.3. Codon usage in *M. maris* AB-18-032, the *aby* BGC, *S. albus* J1074, *S. coelicolor* A3(2), *S. lividans* TK24 and the *abm* BGC.

3.3.2 Transformation of the abyssomicin BGC in E. coli *BAP1 and* Streptomyces *species.*

In order to transform *E. coli* BAP1 with the abyssomicin BGC containing vector pESAC13_*aby*, the mobilisation plasmid pTAMob was used (Strand *et al.*, 2014). This vector was transferred by electroporation into *E. coli* Top10 pESAC13_*aby* (Figure S3.3), that was then used as parental strain for conjugations with *E. coli* BAP1 (Pfeifer *et al.*, 2001). After conjugation, in order to differentiate both *E. coli* strains, PCRs were carried out to amplify the *sfp* gene present in *E. coli* BAP1 and absent in *E. coli* Top10. This PCR yielded unspecific bands despite trying to optimise the reaction by gradient PCR (Figure S3.4). Despite this, it was possible to differentiate correct *E. coli* BAP1 exconjugants (Figure S3.5) that were later confirmed for the presence pESAC13_*aby* (Figure S3.6).

S. lividans K4-144 and *S. albus* J1074 were successfully transformed with the integrative vector pESAC13_*aby* by conjugation with *E. coli* Top10 pESAC13_*aby* pTAMob (Figure S3.7). *S. coelicolor* M1145, *S. coelicolor* M1152 and *Streptomyces* sp. FR-008 LQ3 exconjugants containing pESAC13_*aby* were obtained using the classic intergeneric conjugation between *E. coli* ET12567 and methylation-restrictive *Streptomyces* (Kieser *et al.*, 2000) (Figure S3.8). None of the conjugation protocols used yielded correct exconjugants of *S. albus* S4 Δ5 nor *Salinispora tropica.*

3.3.3 Abyssomicin production assays in E. coli *BAP1 and* Streptomyces *species.*

In order to check if *E. coli* BAP1 was capable of producing abyssomicins when carrying pESAC13_*aby*, fermentations were carried out at 16 °C (Figure S3.9 and S2.10) and 37 °C (Figure S3.11 and S2.12) in A1 and SGG containing 50 µg/ml kanamycin and extracts were analysed by LC-MS. None of the already known abyssomicins (Table S3.3) were detected in any of the tested conditions, despite the decision of carrying out fermentations at 16 ºC was taken since growth at this temperature favours the heterologous expression of complex and large proteins such has PKSs (Rosano & Ceccarelli, 2014). Antimicrobial disk assays were not carried out with the extracts as the kanamycin present in the media would give a ZOI greater than any potential abyssomicin produced by fermentation. For that same reason, antimicrobial plug assays were not performed either.

Similarly, fermentations were carried out for *S. albus* J1074 and *S. albus* J1074 harbouring pESAC13_*aby*, *S. lividans* K4-114 and *S. lividans* K4-114 harbouring pESAC13_*aby* and *S. coelicolor* M1152 and *S. coelicolor* M1152 harbouring pESAC13_*aby* in A1 and SGG. For *S.*

albus and *S. lividans* fermentations were also carried out in other four production media (A1Sal*, SM10, SM17 and R5A), but due to time and equipment limitations production in these media was not tested for the remaining hosts. From these fermentations, samples were taken from day 1 to day 5 and the organic extracts were used for antimicrobial plate assays against *B. subtilis*. *S. albus* J1074 extracts revealed background antimicrobial activity when cultured in A1, R5A, SM10 and SM17 (Figure 3.4A). In comparison, *S. albus* pESAC13_*aby* only showed enhanced antimicrobial activity when cultured in SGG, however, the ZOI observed had the same size despite the presence of PABA (Figure 3.4B). For *S. lividans* K4-114, no background antimicrobial activity was detected against *B. subtilis* in any fermentation media and a ZOI of 2 mm that reduced to 1.5 mm in the presence of PABA was observed for *S. lividans* pESAC13_*aby* (Figure 3.4C and 3.4D).

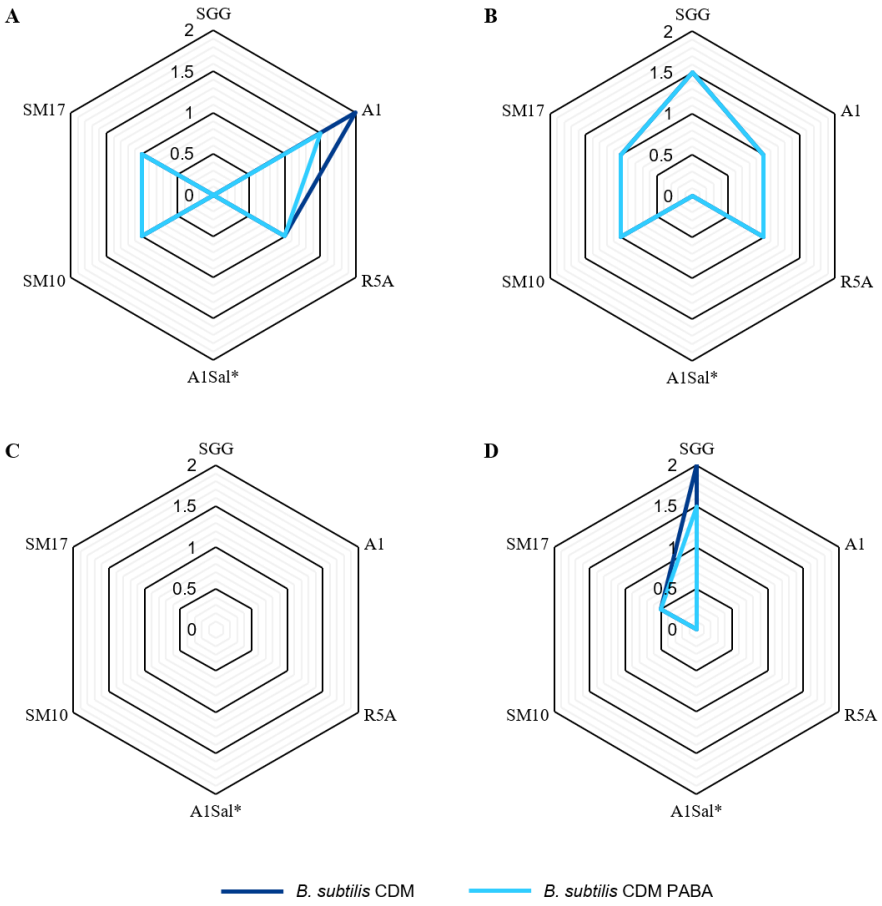

Figure 3.4. Zones of inhibition (mm) generated by the fermentation extracts of **A)** *S. albus* J1074, **B)** *S. albus* pESAC13_*aby*, **C)** *S. lividans* K4-114 and **D)** *S. lividans* pESAC13_*aby* when grown in SGG, A1, R5A, A1Sal*, SM10 and SM17 media for five days. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

In the case of *S. coelicolor* M1152, fermentation extracts of M1152 grown in SGG and A1 showed ZOI of 2 mm (Figure 3.5A). Extracts from *S. coelicolor* pESAC13_*aby* generated ZOI

of 1 mm and 2 mm when obtained from SGG and A1 cultures respectively and the addition of PABA did not reduce these (Figure 3.5B).

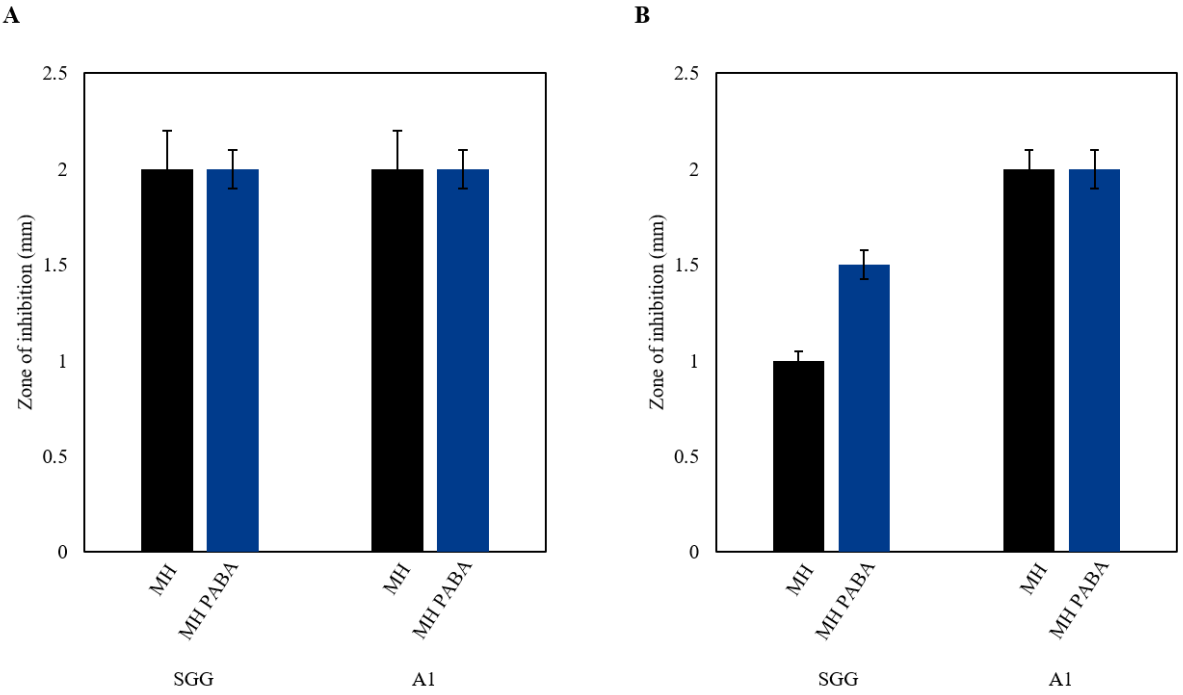

Figure 3.5. A) Zones of inhibition generated by *S. coelicolor* M1152 fermentation extracts when grown in SGG and A1 media for five days against *B. subtilis*. **B)** Zones of inhibition generated by *S. coelicolor* M1152 harbouring pESAC13_*aby* fermentation extracts when grown in SGG and A1 media for five days against *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

Those same extracts of *S. albus*, *S. lividans* and *S. coelicolor* were run through an LC-MS, and as for *E. coli*, none of the already known abyssomicins (Table S3.3) were detected in any of the tested conditions (Figure 3.11 and 3.12).

In order to test whether abyssomicin production in the heterologous hosts would be restricted to these growing on solid media, *S. albus*, *S. lividans* and *S. coelicolor* were cultured on plates of the six production media for 15 days. After that, plugs from these plates were used in an antimicrobial susceptibility test against *B. subtilis*. No differences were appreciated between *S. albus* J1074 and *S. albus* pESAC13_*aby*, except when grown in SGG, although PABA did not restore growth in any case (Figure 3.6). In the case of *S. lividans*, none of the plugs generated a ZOI. Also, no differences were appreciated between the ZOI generated by *S. coelicolor* M1152 and *S. coelicolor* pESAC13_*aby* that remained at 0.5 mm when the strains were grown in SGG and 1 mm when grown in A1 (Figure 3.7A and 3.7B).

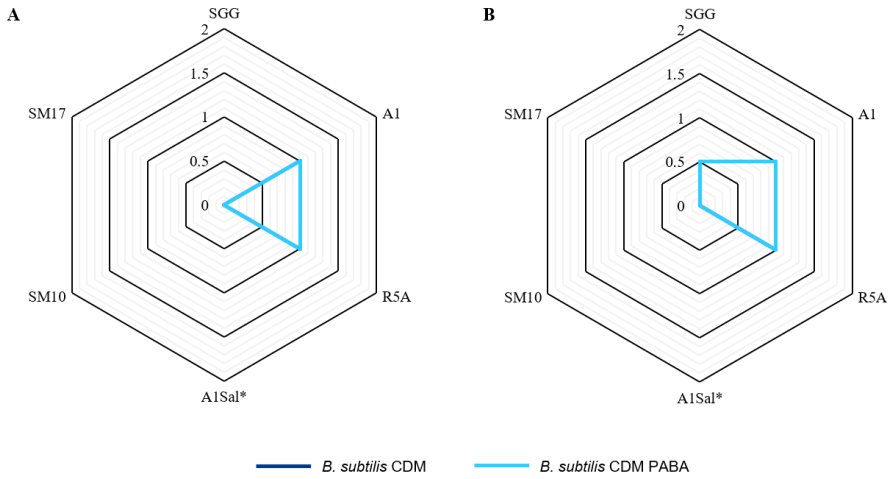

Figure 3.6. Zones of inhibition (mm) generated by the plugs of **A)** *S. albus* J1074 and **B)** *S. albus* pESAC13_*aby* grown in SGG, A1, R5A, A1Sal*, SM10 and SM17 for 15 days tested against *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

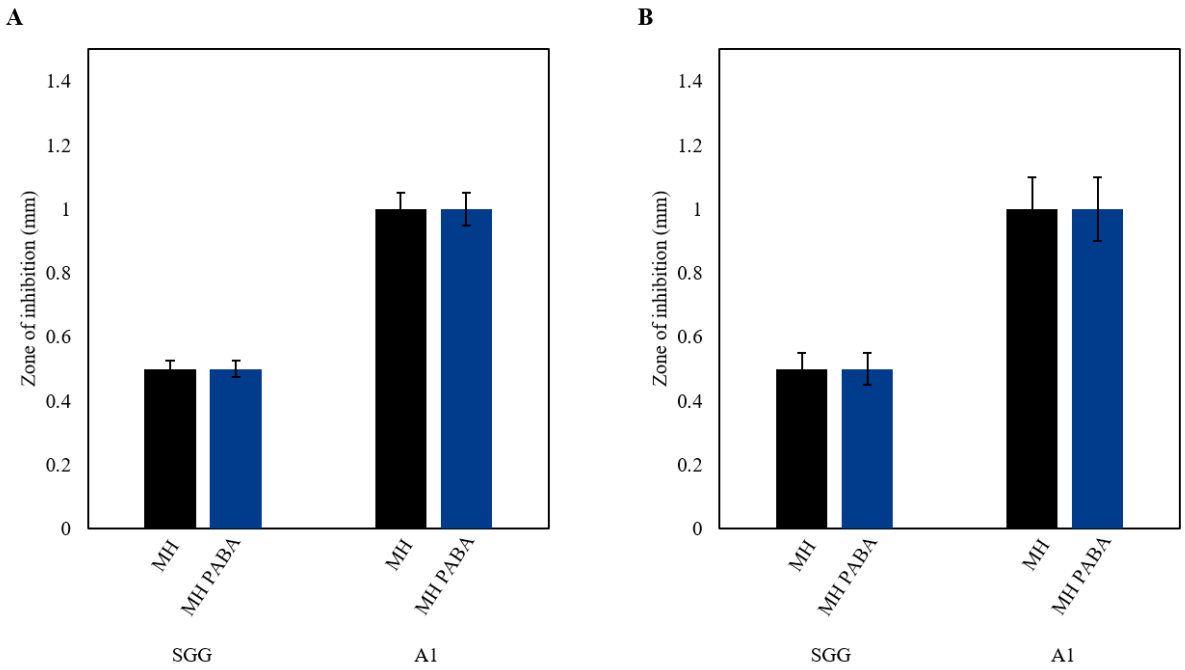

Figure 3.7. Zones of inhibition (mm) generated by the plugs of **A)** *S. coelicolor* M1152 and **B)** *S. coelicolor* M1152 harbouring pESAC13_*aby* grown in SGG and A1 for 15 days tested against *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

Since the antimicrobial activity exhibited by the plugs followed similar trends to those produced by the extracts from liquid cultures, organic extractions from the plates were not carried out.

3.3.4 Ribosome engineering of the Streptomyces *heterologous hosts.*

In order to test whether it would be possible to activate abyssomicin production in the heterologous hosts, ribosome engineering was the first approach followed. *S. albus* J1074, *S. lividans* K4-114 and *S. coelicolor* M1152 and harbouring pESAC13_*aby* ribosome engineering mutants were obtained from culturing these strains under the presence of different concentrations of streptomycin or rifampin (5, 10 and 30 µg/ml). Then, plugs from isolated mutants grown in SGG agar for 15 days were used to determine any potential antimicrobial activity against *B. subtilis*. For *S. lividans* K4-114 (SLwt), five mutants with enhanced antimicrobial activity were obtained for the wild-type strain that generated ZOI of 0.5 – 1 mm (Figure 3.8). None of the *S. lividans* pESAC13_*aby* mutants screened displayed any antimicrobial activity. In the case of *S. coelicolor* M1152, six mutants were obtained from the wild-type strain (SCwt) and two from *S. coelicolor* harbouring pESAC13_*aby* (SCpac) (Figure 3.8). One of the mutants of *S. coelicolor* pESAC13_*aby* generated a ZOI 2.5 mm larger than the wild-type *S. coelicolor* pESAC13_*aby* (Figure 3.7). However, since the same improvement in antimicrobial activity was observed for one mutant of the wild-type *S. coelicolor* M1152 and the presence of PABA in the media seemed to have no effect (Figure 3.8), no further analysis were carried out around these strains. Regarding *S. albus* J1074 and *S. albus* pESAC13_*aby*, none of the screened ribosome engineering mutants displayed higher antibiotic activity than the wild-type strains (Figure 3.6).

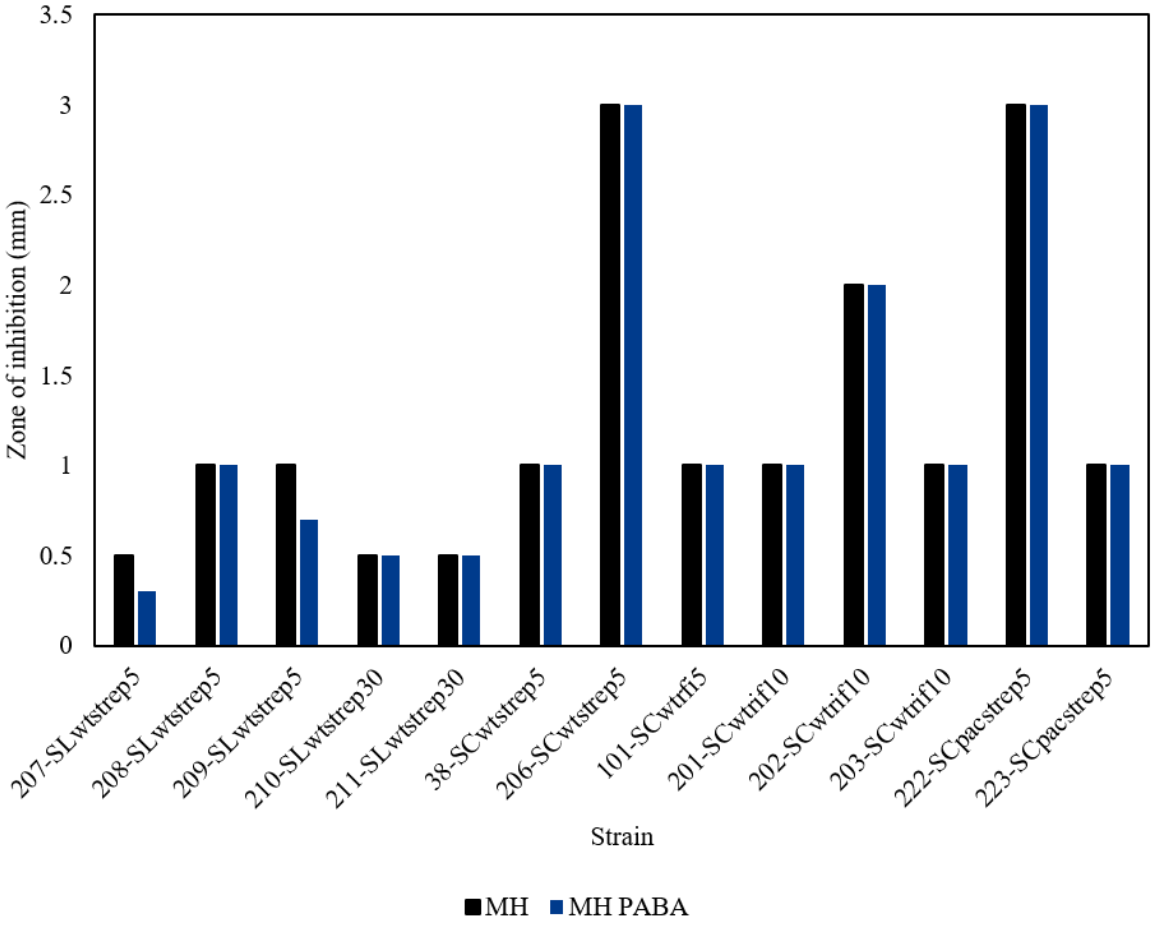

Figure 3.8. Zones of inhibition (mm) generated by the plugs of the ribosome engineering mutants obtained for *S. coelicolor* M1152 (SC wt), *S. coelicolor* pESAC13_*aby* (SCpac) and *S. lividans* K4-114 (SLwt) when grown in SGG for 15 days and tested against *B. subtilis*. Due to technical limitations only one replicate per strain was used for this assay.

3.3.5 CRISPR-Cas9 mediated promoter knock-in.

Previous work revealed that the abyssomicin BGC present in the genome of *M. maris* AB-18-032 contains four transcriptional regulators, including a LuxR transcriptional regulator (*abyH*), two putative activators from the SARP family (*abyR* and *abyI*) and a putative TetR-like repressor (*abyC*) (Gottardi *et al.*, 2011). That same work demonstrated that in-frame deletion of *abyI* completely abolished abyssomicin production while in-frame deletions of *abyR* and *abyC* resulted in a reduction of atrop-abyssomicin C production to less than 10% of the wild-type level. The influence of the in-frame deletion of *abyH* was not studied. Since knocking-out the putative repressor *abyC* was likely going to result in the reduction or abolition of abyssomicin C production, we decided to focus our second approach to drive abyssomicin production in the heterologous hosts on the introduction of the strong synthetic *Streptomyces* promoter *kasOp** to drive the expression of the putative transcriptional activator AbyI. In order

to do that, *kasOp** was cloned into the CRISPR vector pCRISPomyces-2 together with 2 kb homology arms to repair the Cas9 cut upstream *abyI* and introduce the *kasOp** between *abyI* and its natural promoter *abyIp* (Figure 3.9).

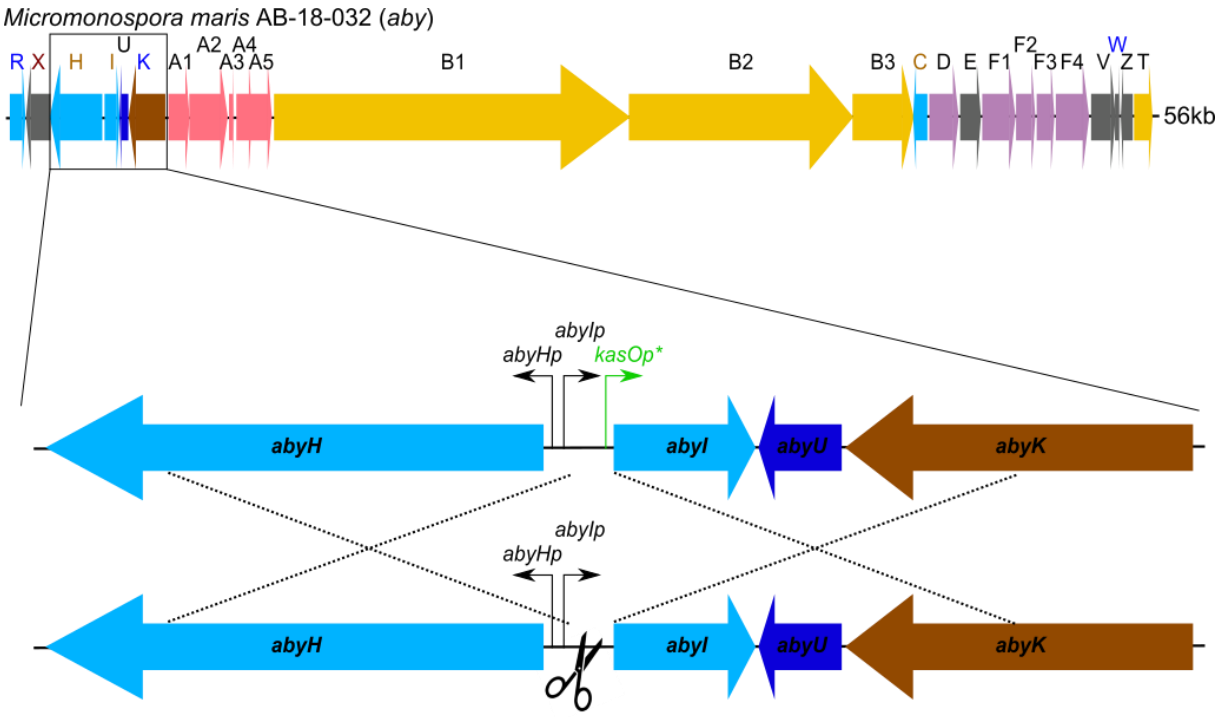

Figure 3.9. CRISPR-Cas9 mediated promoter knock-in strategy. The *kasOp** constitutive synthetic promoter is placed downstream of the wild-type promoter of *abyI* to increase expression of AbyI.

S. albus J1074 and *S. lividans* K4-114 harbouring pESAC13_*aby* were successfully transformed with the CRISPR editing vector pCRISPomyces-2_*kasOp** via intergeneric conjugation with *E. coli* ET12567 (Figure S3.13). This method, however, failed to generate exconjugants for *S. coelicolor* pESAC13_*aby*. Other researchers also reported that the pCRISPomyces-2 system was unsuitable for this species (Alberti & Corre, 2019). Successful *S. lividans* and *S. albus* exconjugants were cultured continuously at 37 °C for 10-15 days to induce plasmid loss and then, abyssomicin production was assessed in liquid media as described before. *S. albus* pESAC13_*aby_kasOp** fermentation extracts in A1 and R5A media displayed a ZOI of 1 mm (Figure 3.10C), whereas *S. albus* pESAC13_*aby* generated, in addition to that, ZOI when cultured in SGG, SM10 and SM17 (Figure 3.10B). For *S. lividans*, none of the extracts of the wild-type K4-114 displayed any background antimicrobial activity against *B. subtilis* (Figure 3.10D), *S. lividans* pESAC13_*aby* extracts generated a ZOI of 2 mm when cultured in SGG that reduced to 1.5 mm in the presence of PABA (Figure 3.10E), and *S. lividans* pESAC13_*aby_kasOp** generated ZOI of 2 mm when cultured in SGG and A1 that also reduced to 1.5 mm ZOI in the presence of PABA (Figure 3.10F). All these organic extracts were run in

the LC-MS (Figure 3.11 and 3.12), however, none of the peaks corresponded to any already known abyssomicin (Table S3.3).

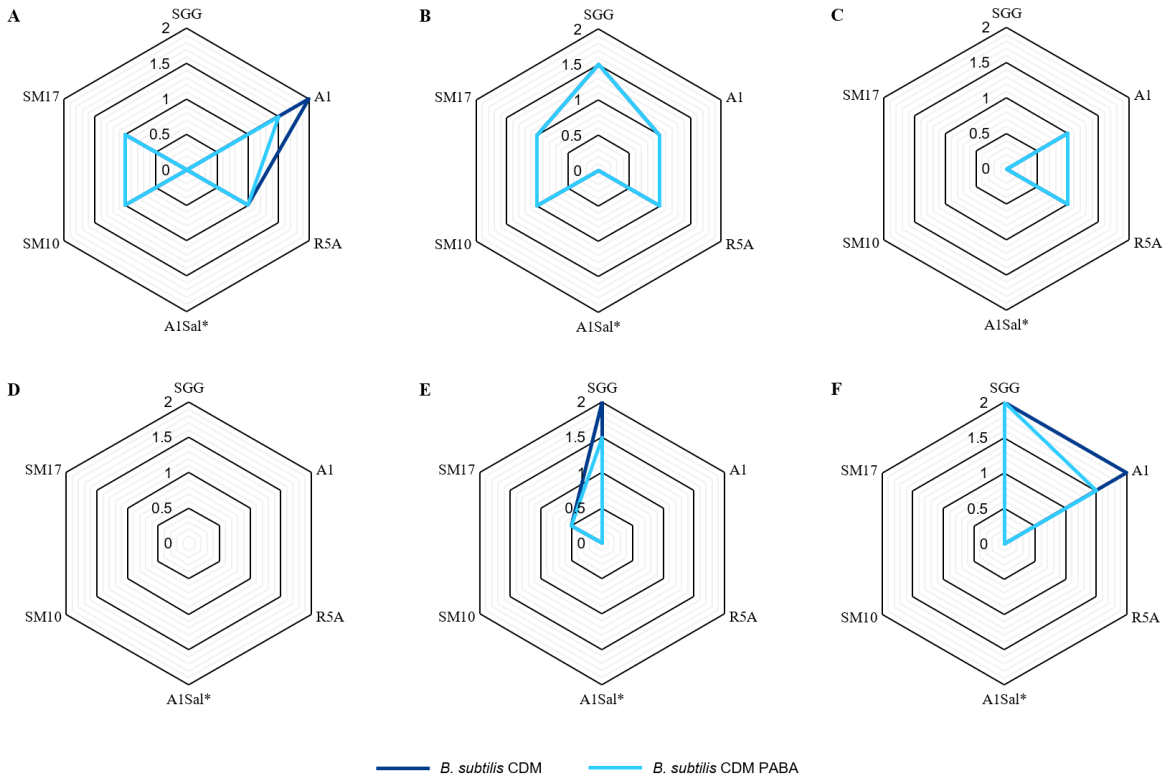

Figure 3.10. Zones of inhibition (mm) generated by the fermentation extracts of **A)** *S. albus* J1074, **B)** *S. albus* pESAC13_*aby*, **C)** *S. albus* pESAC13_*aby_kasOp**, **D)** *S. lividans* K4-114, **E)** *S. lividans* pESAC13_*aby* and **F)** *S. lividans* pESAC13_*aby_kasOp** when grown in SGG, A1, R5A, A1Sal*, SM10 and SM17 media for five days tested against *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

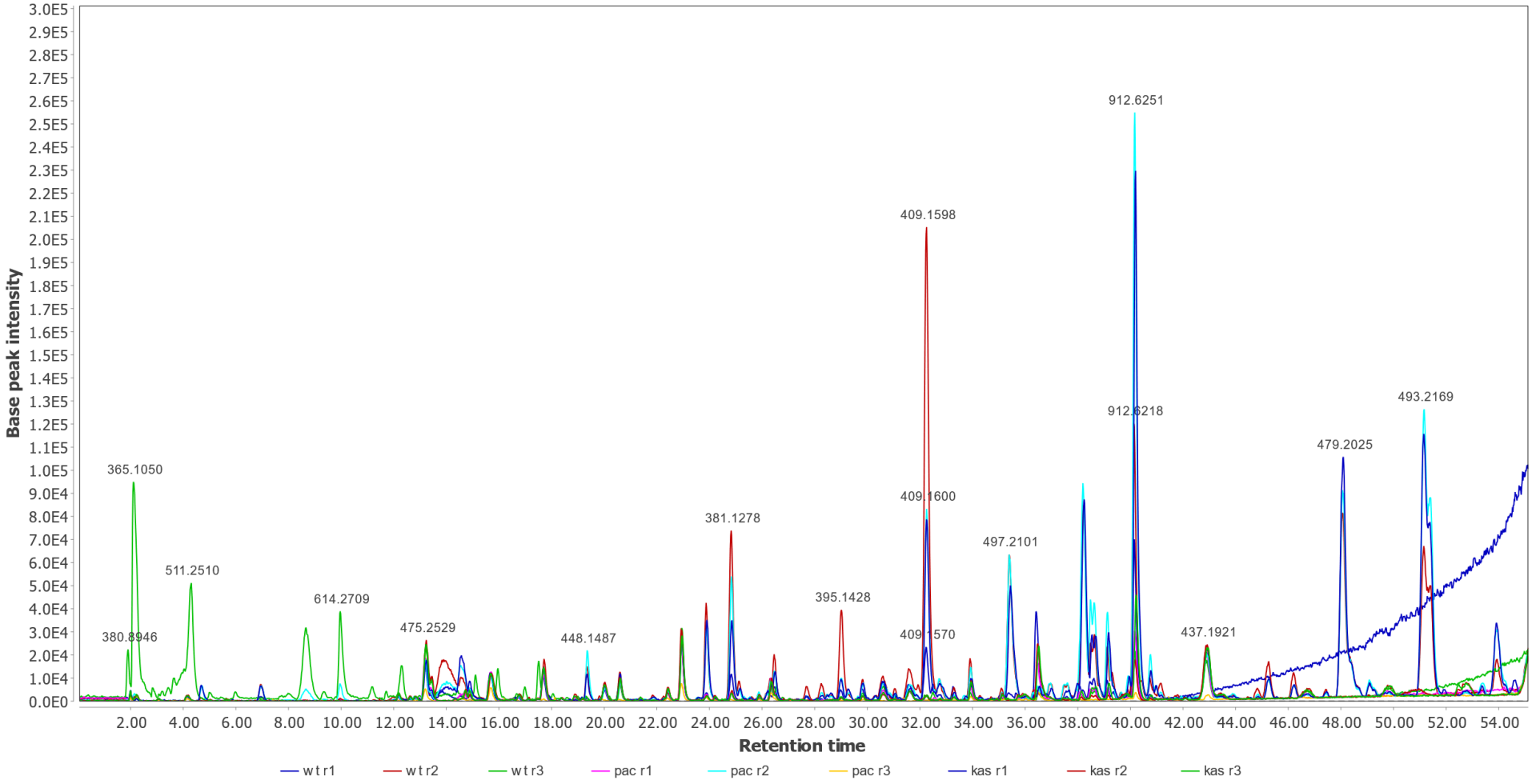

Figure 3.11. LC-MS profile of the fermentation extracts of *S. albus* J1074 (wt) with and without pESAC13_*aby* (pac) and pESAC13_*aby_kasOp** (kas) grown in triplicate at 30 °C in SGG media for 3 days.

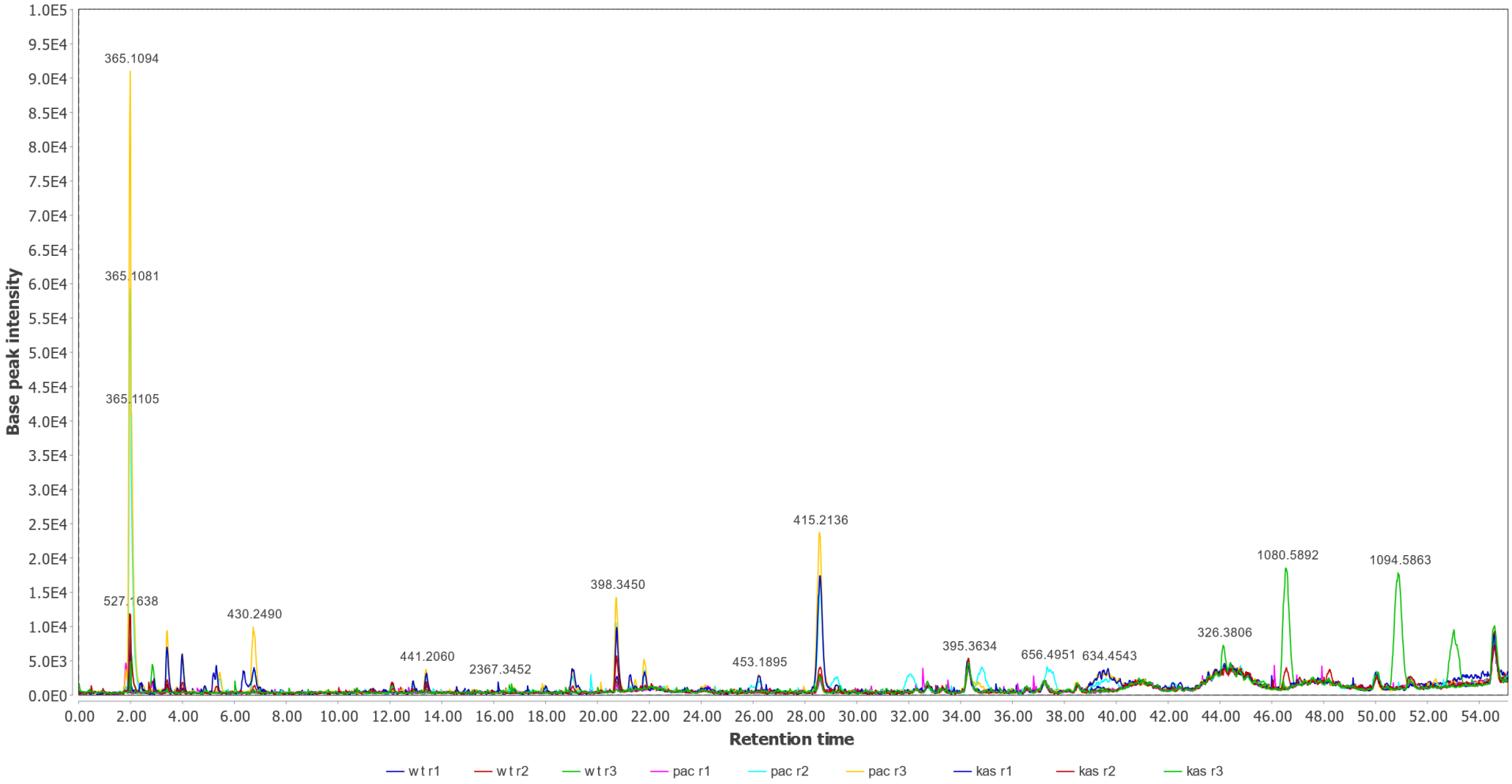

Figure 3.12. LC-MS profile of the fermentation extracts of *S. lividans* K4-114 (wt) with and without pESAC13_*aby* (pac) and pESAC13_*aby_kasOp** (kas) grown in triplicate at 30 °C in A1 media for 3 days.

The potential of these strains to produce compounds with antimicrobial activity against *B. subtilis* was also analysed by culturing the strains in solid media. *S. albus* J1074 plugs from A1 and R5A plates generated a ZOI of 1 mm (Figure 3.13A), similar to *S. albus* pESAC13_*aby* that, in addition to that, only displayed a ZOI of 0.5 mm in SGG (Figure 3.13B). *S. albus* pESAC13_*aby_kasOp** plug from A1 media did generate a ZOI of 2 mm, however, no differences were appreciated with the addition of PABA (Figure 3.13C). None of the plugs of *S. lividans* generated ZOI.

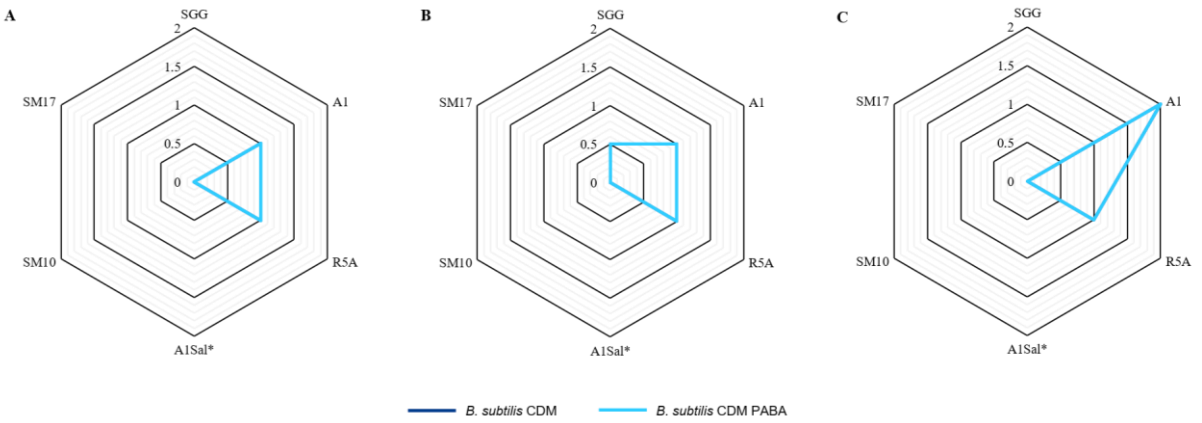

Figure 3.13. Zones of inhibition (mm) generated by the plugs of **A)** *S. albus* J1074, **B)** *S. albus* pESAC13_*aby* and **C)** *S. albus* pESAC13_*aby_kasOp** when grown in SGG, A1, R5A, A1Sal*, SM10 and SM17 media for 15 days tested against *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

Since the plug assay results were not significatively different from the bioassay of the extracts, organic extractions were not carried out from the solid media.

3.3.6 Gene expression analysis.

In order to investigate whether the heterologous hosts were capable of expressing the genes present in the *aby* BGC, RNA extractions and RT-PCRs were carried out. RNA extractions were performed after culturing the strains in SGG for 24 hours and the RT-PCRs revealed that for every tested gene there was gene expression in the strain harbouring pESAC13_*aby* and pESAC13_*aby_kasOp** (Figure 3.14). Due to technical limitations, RT-PCR were not carried out for every gene within the *aby* BGC, but since most genes seem to be expressed in operons, at least one gene from each transcriptional units was amplified, for example, from the operon formed by *abyA1-abyA5*, only *abyA1* was amplified.

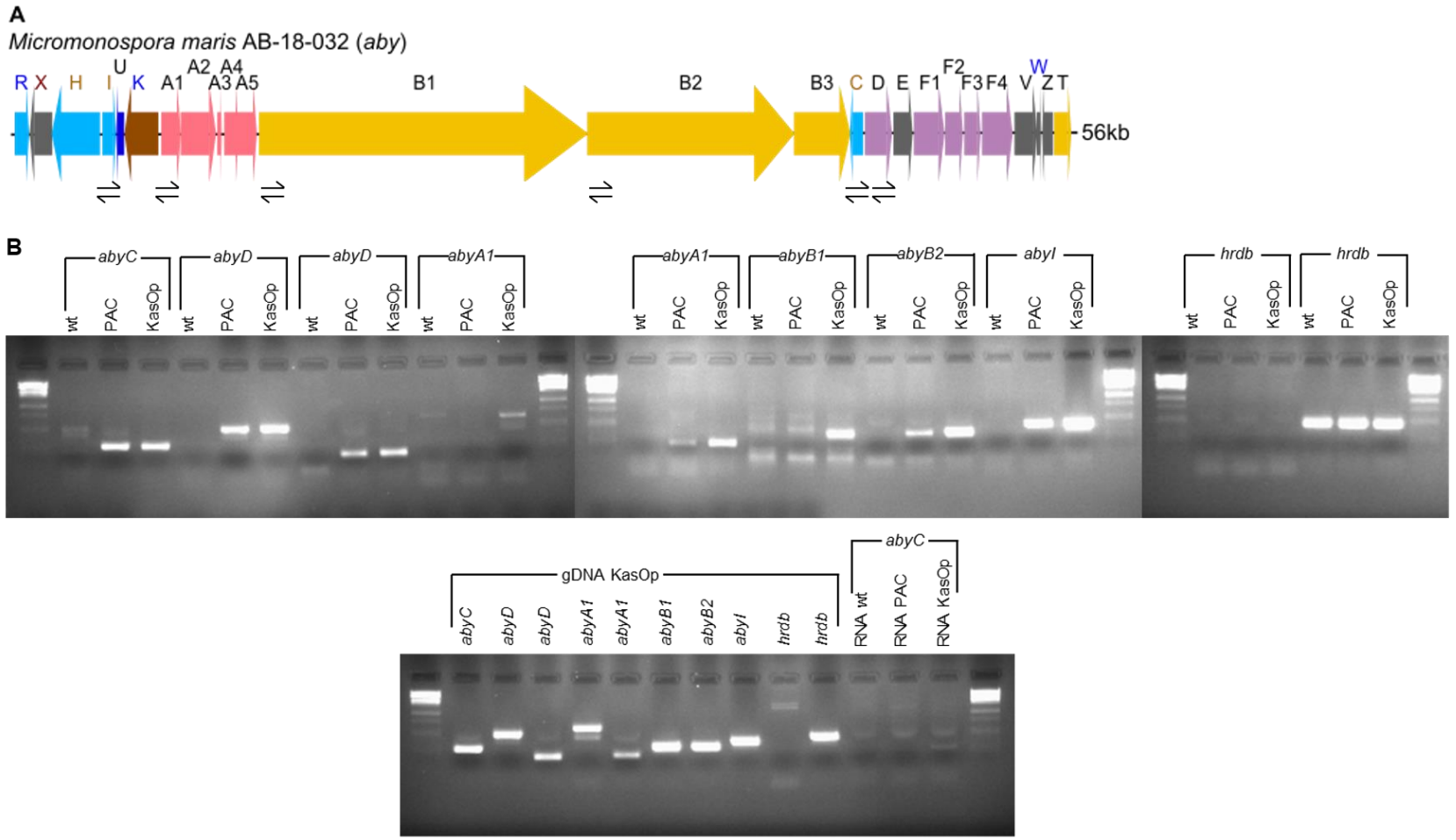

Figure 3.14. A) Location of the RT-PCR primer pairs within the abyssomicin BGC of *M. maris* AB-18-032. **B)** RT-PCRs of different genes along the *aby* BGC carried out using extracted RNA from *S. lividans* K4-114 (wt), *S. lividans* pESAC13_*aby* (PAC) and *S. lividans* pESAC13_*aby_kasOp** (KasOp). Genomic DNA (gDNA) of *S. lividans* pESAC13_*aby_kasOp** was used as positive control for each RT-PCR, and *hrdb* was used as positive control for being a housekeeping gene. For *abyD* and *abyA1* two different primer sets were tested.

Moreover, in most cases, and although RT-PCR is not a quantitative method, it was possible to see a difference in band bright between the strain harbouring pESAC13_*aby* and the strain harbouring pESAC13_*aby_kasOp** (Figure 3.14), indicating that overexpressing the putative pathway activator AbyI was a good approach to increase the expression of the *aby* BGC genes.

3.3.7 Genome sequencing, assembly and annotation.

In order to identify potential gene rearrangements or point mutations that could affect abyssomicin biosynthesis, the genome of *S. lividans* pESAC13_*aby_kasOp** was first sequenced using Illumina MiSeq. SPAdes was used to assemble the trimmed reads generated by the provider and then QUAST was used with the SPAdes output file contigs.fasta to retrieve the assembly statistics (Table S.3.4) (Bankevich *et al.*, 2012; Gurevich *et al.*, 2013). Bandage was used to visualise the connections of the assembled contigs using the SPAdes output assembly_graph_with_scaffolds.gfa (Figure S3.14 and S3.15) (Wick *et al.*, 2015). The trimmed Illumina reads were mapped first against pESAC_13_*aby_kasOp** (Figure 3.15). These covered the entire vector and the *aby* BGC with a mean coverage of 108X (Table S3.5) and it was possible to identify several mutations along the *aby* BGC mainly located in four conflict sites within *abyB1* and one conflict site upstream *abyI* where *kasOp** had been inserted (Table S3.6). At these five points, the coverage reduces drastically, so in order to know the precise sequence it would be necessary to PCR amplify these regions and Sanger sequence them.

The trimmed Illumina reads were also mapped against the entire genome of *S. lividans* TK24 pESAC_13_*aby_kasOp** (Figure 3.16). From this alignment it was possible to see that the vector containing the *aby* BGC had inserted into the previously described *attB* site in the chromosome of *S. lividans* TK24 (Combes *et al.*, 2002). The mean coverage along the alignment was 58X with a pairwise identity of 99.4% (Table S3.7).

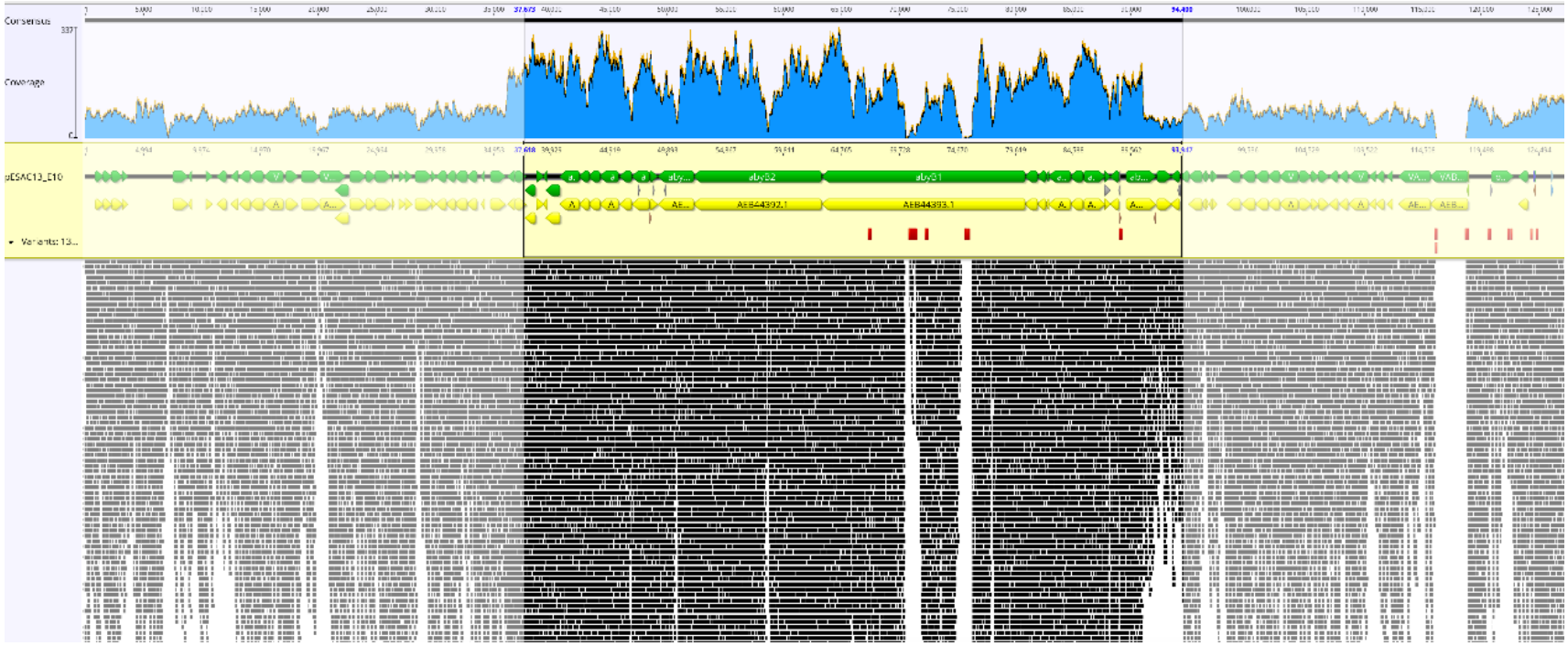

Figure 3.15. Alignment of the Illumina MiSeq reads against pESAC13_*aby_kasOp**. The selected region corresponds to the *aby* BGC and *abyB1* is the largest gene within the *aby* BGC displaying four red marks.

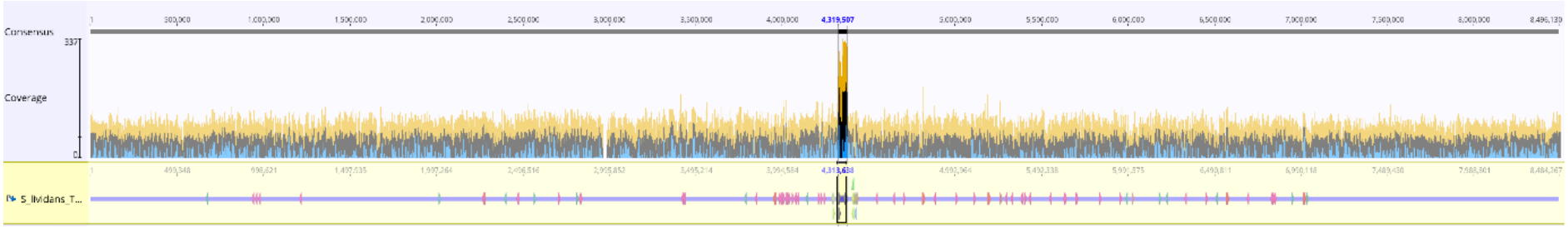

Figure 3.16. Alignment of the Illumina MiSeq reads against *S. lividans* pESAC13_*aby_kasOp**. The selected region highlights the insertion in the genome of pESAC13_*aby_kasOp**.

The genome of *S. lividans* pESAC13_*aby_kasOp** was also sequenced by Nanopore sequencing, however, the number of reads that passed the quality check was very low, resulting in a very low coverage (Figure 3.17). Despite this, it was sufficient to check for the presence of pESAC13_*aby_kasOp**. Merging both the Illumina and Nanopore sequencing results together confirmed the need to Sanger sequence the conflicting regions mentioned before.

Figure 3.17. Alignment of the Nanopore reads against pESAC13_*aby_kasOp** and close look at the region upstream *abyI* where the *kasOp** was inserted via CRISPR.

3.4 Discussion.

3.4.1 Heterologous production of abyssomicins and current challenges in heterologous production projects.

For decades, heterologous production has been a suitable approach for the production of high value products naturally produced by other non-model organisms that are recalcitrant to conventional laboratory methods. More than 80% of the small molecules with medical applications derive from or are inspired by natural products synthesised by *Actinobacteria*, although researchers have also reported the discovery of natural products from, for example, myxobacteria and cyanobacteria (Cook & Pfleger, 2019). In order to facilitate the biosynthesis of these new compounds, phylogenetically similar heterologous hosts are generally the most promising candidates. For that reason, with the aim of producing abyssomicins in a heterologous host for which genetic tools have already been developed, the last ten years (2006-2016) of existing literature on heterologous production of bacterial polyketide natural products were reviewed. That provided an interesting insight into the currently available heterologous hosts, the hosts that have been used more successfully and strains from which genera have been utilised to express BGCs from other different genera (Table S3.2, Figure 3.1 and Figure 3.2).

As it has been reported before, another key aspect to consider when expressing genes in a heterologous host is the effects produced by codon usage biases. Codon usage not only affects gene expression and mRNA levels, it also has an important impact on translation and protein folding, as even synonymous codon replacements can change protein structure and function (Angov *et al.*, 2008; Zhoua *et al.*, 2016). Since strains of *S. lividans*, *S. coelicolor* and *S. albus* seemed to be the most promising hosts according to the literature survey and the three of them had been used previously for the heterologous production of *Micromonospora* secondary metabolites (Lombó *et al.*, 2006; Yang *et al.*, 2015; Braesel & Eustáquio, 2019), the genomes of the type strains *S. lividans* TK24, *S. coelicolor* A3(2) and *S. albus* J1074 were used to analyse the codon usage of these strains and compare it to the codon usage of *M. maris* AB-18-032 and more specifically to the codon usage in the *aby* BGC (Figure 3.3). The results showed, in general, a similar codon usage pattern between strains, however, it was possible to identify some codons that were more abundant within the *aby* BGC than in *Streptomyces*. Indeed, in this analysis, the codon usage of the abyssomicin BGC present in *S. koyangensis* SCSIO 5802 that had previously been expressed in *S. coelicolor* M1152 (Song *et al.*, 2017; Tu *et al.*, 2018) was included and the codon usage in the *abm* BGC was similar to the codon usage in the *aby*

BGC with minor exceptions (Figure 3.3). Based on the results from this analysis and the observations from the literature review carried out before, the following hosts were selected for the expression of the *aby* BGC: *S. albus* J1074, *S. lividans* K4-114, *S. coelicolor* M1145, *S. coelicolor* M1152 and *E. coli* BAP1. Moreover, while this thesis was underway, various new potentially relevant hosts for the heterologous production of bacterial secondary metabolites were developed in other laboratories. Among those, three were chosen to be included in this study on the heterologous production of abyssomicins: *Streptomyces* sp. FR-008 LQ3 (Liu *et al*., 2016), *S. tropica* CNB-4401 (Zhang *et al*., 2018) and *Streptomyces albus* S4 Δ5 (Fazal *et al*., 2019). Among those, *S. tropica* CNB-4401, would be the host phylogenetically closest to *M. maris* AB-18-032.

The *aby* BGC was successfully transformed into *S. albus* J1074 and the engineered super hosts *S. lividans* K4-114, *S. coelicolor* M1145, *S. coelicolor* M1152, *Streptomyces* sp. FR-008 LQ3 and *E. coli* BAP1. Interestingly, pESAC13_*aby* was not amenable to transformation in *S. albus* S4 Δ5, an *S. albus* strain in which five native BGCs had been deleted (Fazal *et al.*, 2019) nor the engineered *S. tropica* CNB-4401 (Zhang *et al.*, 2018), that would be more closely related to *Micromonospora* than *Streptomyces* hosts. Due to time and workload limitations it was not possible to keep repeating these transformations for longer than two months. With the transformed strains, production assays were carried out in the two media in which abyssomicins had been reported previously and various solid and liquid media commonly used for the production of secondary metabolites. First, fermentations were carried out in liquid media, and the organic extracts obtained were used in antimicrobial plate diffusion assays. Despite *S. albus* pESAC13_*aby* showed enhanced antimicrobial activity when cultured in SGG in comparison to *S. albus* J1074, and *S. lividans* pESAC13_*aby* generated ZOI of 2 mm that reduced to 1.5 mm in the presence of PABA (Figure 3.4), none of the already known abyssomicins were found when the extracts were analysed through LC-MS. As it has been seen before that liquid and solid-state fermentations affect *Streptomyces* differentiation and the production of secondary metabolites (English *et al.*, 2017; Manteca & Yagüe, 2018), plug assays from plates were carried out later to study whether the strains could be producing abyssomicins in solid media, but the results revealed that was unlikely (Figure 3.6 and 3.7).

Multiple hypotheses could be formulated when looking for an explanation to these results. In first place, it would be possible that the *M. maris aby* BGC would be silent, however, the *aby* BGC was active in *M. maris* at the moment of generating the phage artificial chromosome

library that captured the *aby* BGC into pESAC13. Unless the activation of the *aby* BGC in *M. maris* relies in one or more regulatory genes present outside the *aby* BGC and thus absent in pESAC13_*aby*, this should be active in the heterologous hosts. Another hypothesis would be that the heterologous hosts were not able to recognise some of the promoter regions in the *aby* BGC and thus transcription of some genes would fail. This would be very likely to happen in *E. coli* BAP1 for example, as most of the times, when this host is used for the production of natural products from *Actinobacteria*, native promoters are replaced by *E. coli* promoters such as the inducible T7 promoter (Zhang *et al.*, 2010; Antosch *et al.*, 2014). In other cases, for example, for the production of oxytetracyclines in *E. coli* BAP1, production was achieved only after over-expressing alternative sigma factors that activated transcription (Stevens *et al.*, 2013). Indeed, a similar approach is often followed when using *Streptomyces* as heterologous hosts. For example, strong *Streptomyces* promoters such as *ermE*p* and *kasOp** have been used to activate and drive production of oxytetracylines (Yin *et al.*, 2016), spinosad (Tan *et al.*, 2017) and toyocamycin (Xu *et al.*, 2017). Other hypothesis that could explain why abyssomicin production in the heterologous hosts was not detected could comprise, in first instance, the fact that if there was any production, this could be too low to be detected (Baltz, 2010). Second, it could be possible that due to challenges encountered by the heterologous hosts during transcription or translation, these could be producing intermediates, modified intermediates or chemically modified abyssomicins that could not be identified. However, in this latter case, there would have been new peaks in the LC-MS extracts from the strains harbouring pESAC13_*aby* that were absent in the extracts of the wild-type strains. Another factor to take into account would be the availability of the required precursors, as other native BGCs from the heterologous host could be consuming the same building blocks required for abyssomicin biosynthesis. Finally, failed expression of the resistance gene or the ABC exporter system could lead to the repression of abyssomicin biosynthesis by the host organism to avoid abyssomicin accumulation to lethal levels.

In an attempt to activate abyssomicin production in the heterologous hosts, two different approaches were followed. First, ribosome engineering was applied to the *Streptomyces* heterologous hosts as well as to the wild-type strains. The isolated mutants were grown in the abyssomicin production media and then were tested for their ability to produce antimicrobial compounds against *B. subtilis*. Although over 250 drug-resistant mutants were screened, none of the *S. lividans* pESAC13_*aby* nor *S. albus* pESAC13_*aby* mutants screened displayed higher antimicrobial activity than the wild-type strains (Figure 3.6 and 3.8). In the case of *S.*

coelicolor, one mutant of *S. coelicolor* pESAC13_*aby* generated a ZOI bigger than the wild-type *S. coelicolor* pESAC13_*aby* (Figure 3.7), but, since the same improvement in antimicrobial activity was observed for one mutant of the wild-type *S. coelicolor* and the presence of PABA in the media seemed to have no effect (Figure 3.8), no further analysis were carried out around this strain. Due to time and workload limitations no more rounds of ribosome engineering were attempted.

Since ribosome engineering does not allow the targeted activation of selected BGCs (Shentu *et al.*, 2015; Ochi, 2017), the activation of the *aby* BGC in the heterologous hosts via CRISPR was pursued. Using the *Streptomyces* CRISPR system based on the plasmid pCRISPomyces-2 (Cobb *et al.*, 2015), the strong synthetic *Streptomyces* promoter *kasOp** was introduced upstream *abyI*, the main activator in the *aby* BGC (Figure 3.9), in *S. albus* pESAC13_*aby* and *S. lividans* pESAC13_*aby*. After that, liquid and solid-state fermentations were carried out, although again, it was not possible to identify any of the known abyssomicins by LC-MS (Figure 3.11 and 3.12). This approach has been reported to be fruitful in several heterologous production projects before, where the heterologous hosts even reached or improved the production yield of the native producer (Bai *et al.*, 2015; Yin *et al.*, 2016; Li *et al.*, 2018), for that reason, further investigations were carried out and the gene expression of the *aby* BGC in the heterologous hosts was analysed. RNA extractions and RT-PCRs were carried out in *S. lividans* harbouring pESAC13_*aby* and pESAC13_*aby_kasOp**, and all of the tested genes showed active transcription (Figure 3.14), indicating that translation errors may be the explanation for a lack of abyssomicin biosynthesis in the heterologous hosts.

Since carrying out a proteomics analysis was initially outside the scope of this project, the genome of one of the heterologous hosts after the introduction of pESAC13_*aby_kasOp** was re-sequenced instead to identify any potential mutation that could be stopping translation prematurely or that could be leading to the production of unidentified intermediates. *S. lividans* pESAC13_*aby_kasOp** was sequenced using Illumina and Nanopore technologies and the results revealed four conflict sites within *abyB1* and one conflict site upstream *abyI* where *kasOp** had been inserted (Table S3.6). This last site should not be problematic, as although it seems that the CRISPR-mediated insertion of *kasOp** truncated the natural promoter of *abyI*, it was possible to demonstrate active gene expression (Figure 3.14). The four points in *abyB1* where the coverage reduced drastically, seem to indicate the presence of some point mutations as well as larger deletions (Table S3.6). Pending re-sequencing of those sites, it seems likely

that these mutations could be the reason why none of the heterologous hosts produced abyssomicins. It would be interesting to see whether these mutations occurred once in *S. lividans* or while pESAC13_*aby* was in *E. coli* prior to conjugations. Sequencing the other heterologous hosts as well as pESAC13_*aby* would aid in understanding this. These results highlight the importance of sequencing the heterologous hosts after transformation, as the high GC content and the repetitive nature of the PKS genes make them hotspots of genome rearrangements and mutations (Dunn & Anderson, 2019; Olarte *et al.*, 2019).

3.5 Conclusions and future perspectives.

The aim of this chapter was to investigate the heterologous expression of the abyssomicin BGC from *M. maris* AB-18-032. The six selected hosts are frequently used for the heterologous production of bacterial PKSs and their codon usage is similar to that of *M. maris* and the *aby* BGC. The *aby* BGC was successfully conjugated into all the different hosts. These strains were cultured under different conditions but no abyssomicin production was detected. To try to drive abyssomicin biosynthesis two of the heterologous hosts were engineered using CRISPR, however, production was not achieved either despite active gene expression. Finally, sequencing the engineered *aby* BGC in *S. lividans* confirmed several mutations in the polyketide synthase. This highlights the importance of strain sequencing after engineering.

Future work could focus at confirming the potential mutations found in *abyB1* and finding their origin. Moreover, since it has been demonstrated that editing of the *aby* cluster in the heterologous host is possible using the pCRISPomyces-2 system, it should be possible to correct any mutations to re-establish abyssomicin production. Indeed, having a heterologous abyssomicin producer would allow, then, to carry out modifications along the BGC that could result in new abyssomicin structures with bioactivities potentially interesting in the biopharmaceutical industry. For example, the KS domains in modules 2 and 6 of the PKS could be modified to accept methylmalonyl-CoA and malonyl-CoA respectively to produce abyssomicin 2 in *M. maris* instead of abyssomicin C (Figure 3.18A). Other modifications could include activating/inactivating DH, KR and/or ER domains to produce abyssomicins with varying saturation/oxidation patterns (Figure 3.18B). And finally, chain release could be controlled by the position of the terminal acyl carrier protein domain of AbyB3.

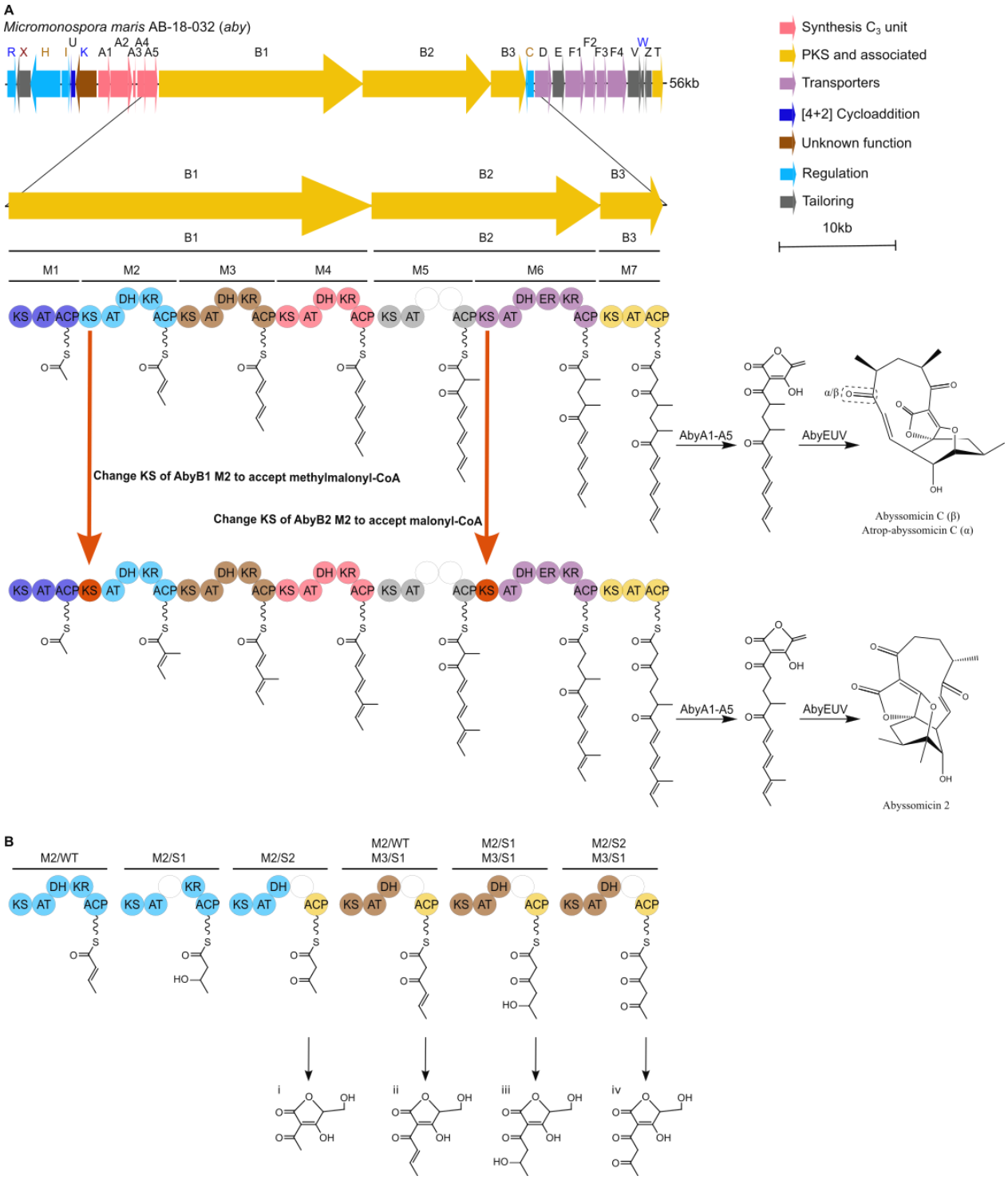

Figure 3.18. A) Schematic representation of abyssomicin biosynthesis in *M. maris* and how it could be engineered to produce abyssomicin 2 instead of abyssomicin C. **B)** Examples of modules that could be inactivated within AbyB1 to synthetise products with different saturation/oxidation patterns.

In conclusion, this work is an example of the unpredictability of biological systems. Despite this project was designed always making informed decisions to maximise the chances of

success, the scientific community is still far from fully controlling biological systems. In order to achieve that, more basic research is needed.

3.6 References.

Ahmed, Y., Rebets, Y., Estévez, M. R., Zapp, J., Myronovskyi, M. & Luzhetskyy, A. (2020). Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb Cell Fact* **19**, 1–16.

Alberti, F. & Corre, C. (2019). Editing *Streptomycete* genomes in the CRISPR/Cas9 age. *Nat Prod Rep* **36**, 1237–1248.

Alberti, F., Foster, G. D. & Bailey, A. M. (2017). Natural products from filamentous fungi and production by heterologous expression. *Appl Microbiol Biotechnol* **101**, 493–500.

Alberti, F., Leng, D., Wilkening, I., Song, L., Tosin, M. & Corre, C. (2018). Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *bioRxiv* 265645.

Angov, E., Hillier, C. J., Kincaid, R. L. & Lyon, J. A. (2008). Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS One* **3**.

Antosch, J., Schaefers, F. & Gulder, T. A. M. (2014). Heterologous reconstitution of ikarugamycin biosynthesis in *E. coli*. *Angew Chemie Int Ed* **53**, 3011–3014.

Aso, Y., Nagao, J. I., Koga, H., Okuda, K. I., Kanemasa, Y., Sashihara, T., Nakayama, J. & Sonomoto, K. (2004). Heterologous expression and functional analysis of the gene cluster for the biosynthesis of and immunity to the lantibiotic, nukacin ISK-1. *J Biosci Bioeng* **98**, 429–436.

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., Miao, J., Lou, C., Zhang, L. & Demain, A. L. (2015). Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc Natl Acad Sci U S A* **112**, 12181–12186.

Baltz, R. H. (2010). *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J Ind Microbiol Biotechnol* **37**, 759–772.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S. & other authors. **(2012).** SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455–477.

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y. & van Wezel, G. P. (2016). Taxonomy, physiology, and natural products of *Actinobacteria*. *Microbiol Mol Biol Rev* **80**, 1–43.

Billingsley, J. M., DeNicola, A. B. & Tang, Y. (2016). Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. *Curr Opin Biotechnol* **42**, 74–83.

Braesel, J. & Eustáquio, A. S. (2019). Heterologous expression of a putative Clpc chaperone gene leads to induction of a host metabolite. *Artic J Braz Chem Soc* **30**, 499–508.

Carro, L., Castro, J. F., Razmilic, V., Nouioui, I., Pan, C., Igual, J. M., Jaspars, M., Goodfellow, M., Bull, A. T. & other authors. **(2019).** Uncovering the potential of novel *Micromonosporae* isolated from an extreme hyper-arid Atacama Desert soil. *Sci Rep* **9**.

Choi, K. R., Cho, J. S., Cho, I. J., Park, D. & Lee, S. Y. (2018). Markerless gene knockout and integration to express heterologous biosynthetic gene clusters in *Pseudomonas putida*. *Metab Eng* **47**, 463–474.

Cobb, R. E., Wang, Y. & Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth Biol* 2–7.

Combes, P., Till, R., Bee, S. & Smith, M. C. M. (2002). The *Streptomyces* genome contains multiple pseudo-*attB* sites for the φC31-encoded site-specific recombination system. *J Bacteriol* **184**, 5746–5752.

Cook, T. B. & Pfleger, B. F. (2019). Leveraging synthetic biology for producing bioactive polyketides and non-ribosomal peptides in bacterial heterologous hosts. *Med chem comm* **10**, 668-681.

Davison, E. K. & Brimble, M. A. (2019). Natural product derived privileged scaffolds in drug discovery. *Curr Opin Chem Biol* **52**, 1-8.

Dunn, M. J. & Anderson, M. Z. (2019). To repeat or not to repeat: Repetitive sequences regulate genome stability in *Candida albicans*. *Genes (Basel)* **11**, 866.

English, A. L., Boufridi, A., Quinn, R. J. & Kurtböke, D. I. (2017). Evaluation of fermentation conditions triggering increased antibacterial activity from a near-shore marine intertidal environment-associated *Streptomyces* species. *Synth Syst Biotechnol* **2**, 28–38.

Fazal, A., Thankachan, D., Harris, E. & Seipke, R. F. (2019). A chromatogram-simplified *Streptomyces albus* host for heterologous production of natural products. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* **113**, 511–520.

Fu, J., Wenzel, S. C., Perlova, O., Wang, J., Gross, F., Tang, Z., Yin, Y., Stewart, A. F., Mü Ller, R. & Zhang, Y. (2008). Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res* **36**, 113.

Gomez-Escribano, J. P. & Bibb, M. J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **4**, 207–215.

Gottardi, E. M., Krawczyk, J. M., Von Suchodoletz, H., Schadt, S., Mühlenweg, A., Uguru, G. C., Pelzer, S., Fiedler, H. P., Bibb, M. J. & other authors. **(2011).** Abyssomicin biosynthesis: Formation of an unusual polyketide, antibiotic-feeding studies and genetic analysis. *ChemBioChem* **12**, 1401–1410.

Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075.

Huo, L., Hug, J. J., Fu, C., Bian, X., Zhang, Y. & Müller, R. (2019). Heterologous expression of bacterial natural product biosynthetic pathways. *Nat Prod Rep* **36**, 1412-1436.

Je, W. P., Hong, J. S. J., Parajuli, N., Won, S. J., Sung, R. P., Lim, S. K., Jae, K. S. & Yeo, J. Y. (2008). Genetic dissection of the biosynthetic route to gentamicin A2 by heterologous expression of its minimal gene set. *Proc Natl Acad Sci U S A* **105**, 8399–8404.

Jones, A. C., Gust, B., Kulik, A., Heide, L., Buttner, M. J. & Bibb, M. J. (2013). Phage P1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. *PLoS One* **8**.

Julien, B. & Shah, S. (2002). Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*. *Antimicrob Agents Chemother* **46**, 2772–2778.

Kallscheuer, N., Kage, H., Milke, L., Nett, M. & Marienhagen, J. (2019). Competence of *Corynebacterium glutamicum* as a host for the production of type I polyketides. *bioRxiv* 622399.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. a. **(2000).** Practical *Streptomyces* Genetics. *John Innes Found Norwich*, UK ISBN 0-7084-0623-8.

Li, S., Wang, J., Xiang, W., Yang, K., Li, Z. & Wang, W. (2018). An autoregulated fine-tuning strategy for titer improvement of secondary metabolites using native promoters in *Streptomyces*. *ACS Synth Biol* **7**, 522–530.

Liu, Q., Xiao, L., Zhou, Y., Deng, K., Tan, G., Han, Y., Liu, X., Deng, Z. & Liu, T. (2016). Development of *Streptomyces* sp. FR-008 as an emerging chassis. *Synth Syst Biotechnol* **1**, 207–214.

Lombó, F., Velasco, A., Castro, A., De La Calle, F., Braña, A. F., Sánchez-Puelles, J. M., Méndez, C. & Salas, J. A. (2006). Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two *Streptomyces* species. *ChemBioChem* **7**, 366–376.

Malmierca, M. G., González-Montes, L., Pérez-Victoria, I., Sialer, C., Braña, A. F., Salcedo, R. G., Martín, J., Reyes, F., Méndez, C. & other authors. **(2018).** Searching for

glycosylated natural products in actinomycetes and identification of novel macrolactams and angucyclines. *Front Microbiol* **9**.

Manteca, Á. & Yagüe, P. (2018). *Streptomyces* differentiation in liquid cultures as a trigger of secondary metabolism. *Antibiotics* **2**, 41.

Mutka, S. C., Carney, J. R., Liu, Y. & Kennedy, J. (2006). Heterologous production of epothilone C and D in *Escherichia coli* †. *Biochemistry* **45**, 1321–1330.

Myronovskyi, M. & Luzhetskyy, A. (2019). Heterologous production of small molecules in the optimized: *Streptomyces* hosts. *Nat Prod Rep* **36**, 1281-1294.

Myronovskyi, M., Rosenkränzer, B., Nadmid, S., Pujic, P., Normand, P. & Luzhetskyy, A. (2018). Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* **49**, 316–324.

Nah, H. J., Pyeon, H. R., Kang, S. H., Choi, S. S. & Kim, E. S. (2017). Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. *Front Microbiol* **8**, 394.

Ochi, K. (2017). Insights into microbial cryptic gene activation and strain improvement: Principle, application and technical aspects. *J Antibiot (Tokyo)* **70**, 25-40.

Ochi, K. & Hosaka, T. (2013). New strategies for drug discovery: Activation of silent or weakly expressed microbial gene clusters. *Appl Microbiol Biotechnol* **97**, 87–98.

Olarte, R. A., Menke, J., Zhang, Y., Sullivan, S., Slot, J. C., Huang, Y., Badalamenti, J. P., Quandt, A. C., Spatafora, J. W. & Bushley, K. E. (2019). Chromosome rearrangements shape the diversification of secondary metabolism in the cyclosporin producing fungus *Tolypocladium inflatum*. *BMC Genomics* **20**, 120.

Peirú, S., Rodríguez, E., Tran, C. Q., Carney, J. R. & Gramajo, H. (2007). Characterization of the heterodimeric MegBIIa:MegBIIb Aldo-Keto reductase involved in the biosynthesis of L-mycarose from *Micromonospora megalomicea*. *Biochemistry* **46**, 8100–8109.

Penn, J., Li, X., Whiting, A., Latif, M., Gibson, T., Silva, C. J., Brian, P., Davies, J., Miao, V. & other authors. **(2006).** Heterologous production of daptomycin in *Streptomyces lividans*. *J Ind Microbiol Biotechnol*, 121–128.

Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E. & Khosla, C. (2001). Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790–1792.

Porter, J. L., Tobias, N. J., Pidot, S. J., Falgner, S., Tuck, K. L., Vettiger, A., Hong, H., Leadlay, P. F. & Stinear, T. P. (2013). The cell wall-associated mycolactone polyketide synthases are necessary but not sufficient for mycolactone biosynthesis. *PLoS One* **8**.

Riedlinger, J., Reicke, A., Zähner, H., Krismer, B., Bull, A. T., Maldonado, L. A., Ward, A. C., Goodfellow, M., Bister, B. & other authors. **(2004).** Abyssomicins, inhibitors of the para-aminobenzoic acid pathway produced by the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **57**, 271–279.

Ro, D. K., Paradise, E. M., Quellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S. & other authors. **(2006).** Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943.

Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front Microbiol* **5**, 172.

Roulet, J., Taton, A., Golden, J. W., Arabolaza, A., Burkart, M. D. & Gramajo, H. (2018). Development of a cyanobacterial heterologous polyketide production platform. *Metab Eng* **49**, 94–104.

Shentu, X., Liu, N., Tang, G., Tanaka, Y., Ochi, K., Xu, J. & Yu, X. (2015). Improved antibiotic production and silent gene activation in *Streptomyces diastatochromogenes* by ribosome engineering. *J Antibiot (Tokyo)* **69**, 1–5.

Song, Y., Li, Q., Qin, F., Sun, C., Liang, H., Wei, X., Wong, N. K., Ye, L., Zhang, Y. & other authors. **(2017).** Neoabyssomicins A–C, polycyclic macrolactones from the deep-sea derived *Streptomyces koyangensis* SCSIO 5802. *Tetrahedron* **73**, 5366–5372.

Stevens, D. C., Conway, K. R., Pearce, N., Villegas-Peñaranda, L. R., Garza, A. G. & Boddy, C. N. (2013). Alternative sigma factor over-expression enables heterologous expression of a type II polyketide biosynthetic pathway in *Escherichia coli*. *PLoS One* **8**.

Strand, T. A., Lale, R., Degnes, K. F., Lando, M. & Valla, S. (2014). A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One* **9**, 1–6.

Tan, G. Y., Deng, K., Liu, X., Tao, H., Chang, Y., Chen, J., Chen, K., Sheng, Z., Deng, Z. & Liu, T. (2017). Heterologous biosynthesis of spinosad: An omics-guided large polyketide synthase gene cluster reconstitution in *Streptomyces*. *ACS Synth Biol* **6**, 995–1005.

Tu, J., Li, S., Chen, J., Song, Y., Fu, S., Ju, J. & Li, Q. (2018). Characterization and heterologous expression of the neoabyssomicin/abyssomicin biosynthetic gene cluster from *Streptomyces koyangensis* SCSIO 5802. *Microb Cell Fact* **17**, 1–14.

Tu, Q., Yin, J., Fu, J., Herrmann, J., Li, Y., Yin, Y., Stewart, A. F., Müller, R. & Zhang, Y. (2016). Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Sci Rep* **6**, 24648.

Wang, P., Kim, W., Pickens, L. B., Gao, X. & Tang, Y. (2012). Heterologous expression and manipulation of three tetracycline biosynthetic pathways. *Angew Chemie Int Ed* **51**, 11136–11140.

Wang, Q., Xie, F., Tong, Y., Habisch, R., Yang, B., Zhang, L., Müller, R. & Fu, C. (2020). Dual-function chromogenic screening-based CRISPR/Cas9 genome editing system for actinomycetes. *Appl Microbiol Biotechnol* **104**, 225–239.

Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. (2015). Bandage: Interactive visualization of *de novo* genome assemblies. *Bioinformatics* **31**, 3350–3352.

Xu, X., Wang, J., Bechthold, A., Ma, Z. & Yu, X. (2017). Selection of an efficient promoter and its application in toyocamycin production improvement in *Streptomyces diastatochromogenes* 1628. *World J Microbiol Biotechnol* **33**, 1–8.

Yan, F., Auerbach, D., Chai, Y., Keller, L., Tu, Q., Hüttel, S., Glemser, A., Grab, H. A., Bach, T. & other authors. **(2018).** Biosynthesis and heterologous production of vioprolides: Rational biosynthetic engineering and unprecedented 4-methylazetidinecarboxylic acid formation. *Angew Chemie Int Ed* **57**, 8754–8759.

Yang, C., Huang, C., Zhang, W., Zhu, Y. & Zhang, C. (2015). Heterologous expression of fluostatin gene cluster leads to a bioactive heterodimer. *Org Lett* **17**, 5324–5327.

Yin, S., Li, Z., Wang, X., Wang, H., Jia, X., Ai, G., Bai, Z., Shi, M., Yuan, F. & other authors. **(2016).** Heterologous expression of oxytetracycline biosynthetic gene cluster in *Streptomyces venezuelae* WVR2006 to improve production level and to alter fermentation process. *Appl Microbiol Biotechnol* **100**, 10563–10572.

Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. (2010). Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem Biol* **17**, 1232–1240.

Zhang, J. J., Moore, B. S. & Tang, X. (2018). Engineering *Salinispora tropica* for heterologous expression of natural product biosynthetic gene clusters. *Appl Microbiol Biotechnol* **102**, 8437–8446.

Zhoua, Z., Danga, Y., Zhou, M., Li, L., Yu, C. H., Fu, J., Chen, S. & Liu, Y. (2016). Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc Natl Acad Sci U S A* **113**, E6117–E6125.

Zhu, S., Duan, Y. & Huang, Y. (2019). The application of ribosome engineering to natural product discovery and yield improvement in *Streptomyces*. *Antibiotics* **3**, 133.

Ziermann, R. & Betlach, M. C. (1999). Recombinant polyketide synthesis in *Streptomyces*: Engineering of improved host strains. *Biotechniques* **26**, 106–110.

Chapter 4. Improving the production and detection of abyssomicin C.

Abstract

Bacteria from the phylum *Actinobacteria* are considered the most prolific source of bioactive natural products, indeed more than 65% of the commercially available antibiotics were originally isolated from these microorganisms. As described in Chapter 3, the use of heterologous hosts is an approach frequently used to exploit the biosynthetic capabilities of those *Actinobacteria* strains that remain recalcitrant to standard molecular microbiology methods and genetic engineering. In most cases, the exploitation of native strains as bio-factories for drug production still relies on the use of diverse mutagenic techniques that do not require molecular manipulation of the strain to increase the production of the target compound. *Micromonospora maris* AB-18-032 was the first microorganism in which abyssomicin biosynthesis was described. Its main fermentation product is the antimicrobial atrop-abyssomicin C, however, as with most natural products, production yields in the wild-type strain are low. In this work, the application of ribosome engineering to generate *M. maris* drug-resistant mutants with up to 3.4-fold enhanced abyssomicin C production is presented. This not only facilitates its detection but could also be useful for future scaling-up projects. Moreover, here, the influence that media pH plays on antimicrobial plate diffusion assays was studied. Antibiotic susceptibility tests are basic in antimicrobial drug discovery projects that start with the screening of hundreds of microbial isolates to identify the most promising strains. During this process, strains are often disregarded because they do not show any antimicrobial activity, suggesting that any BGCs present are silent under the fermentation conditions used. However, the influence that the pH of the screening media exerts on the detection of antimicrobials is rarely reported. The data presented here suggests that simple changes to pH may significantly improve detection of antimicrobial compounds in both agar-based and solvent extract whole-cell antimicrobial screens.

4.1 Introduction.

4.1.1 Ribosome engineering.

The sequencing revolution, at the beginning of the 21st century, unveiled that the biosynthetic potential hidden in bacterial genomes, specifically from the phylum *Actinobacteria* was huge (Bentley *et al.*, 2002; Ikeda *et al.*, 2003). However, most BGCs are silent under standard

laboratory conditions or the production yield of interesting secondary metabolites is too low (Reen *et al.*, 2015). Ribosome engineering is a technique in which strains are cultured at sub-lethal concentrations of antibiotics that target the ribosomal machinery in order to obtain drug-resistant mutants (Shima *et al.*, 1996; Ochi *et al.*, 2004). These drug-resistant mutants show mutations in ribosomal genes. For example, streptomycin resistant mutants develop mutations in the 30S ribosomal S12 protein (*rpsL* gene) whereas rifampicin resistant mutants develop those mutations in the RNA polymerase β-subunit (*rpoB* gene). Apart from streptomycin or rifampicin, other antibiotics, such as gentamicin, paromomycin, geneticin, fusidic acid, thiostrepton, and lincomycin, have also been successfully applied to screen for mutants with increased levels of secondary metabolite production (Zhu *et al.*, 2019).

While the full molecular details involved in increased secondary metabolite production in such mutants is still under investigation, two main paradigms have been studied for this: the rifamycin and streptomycin resistance mutations. Rifamycin is an antibiotic that targets the RNA polymerase and the bacterial alarmone guanosine 5′-disphosphate 3′-disphosphate (ppGpp) binds to the RNA polymerase regulating gene expression (Artsimovitch *et al.*, 2004; Saito *et al.*, 2006). Rifampin resistance mutations may mimic the ppGpp-bound form of the RNA polymerase, that has enhanced affinity to the promoter region of genes that regulate the synthesis of secondary metabolites, activating the expression of BGCs (Lai *et al.*, 2002; Xu *et al.*, 2002; Ochi, 2007; Hosaka *et al.*, 2009).

On the other hand, streptomycin resistance is normally originated from mutations at either the 30S ribosomal protein S12 (*rpsL* gene), the 16S ribosomal RNA (16S rDNA) and/or S-adenosyl methionine (SAM)-dependent 16S rRNA methyltransferase (*rsmG* gene) (Ochi, 2017). It has been demonstrated that *rpsL*-mutant ribosomes are more stable than wild-type ribosomes and show increased expression of the translation factor ribosome recycling factor enhancing protein synthesis (Okamoto-Hosoya *et al.*, 2003; Hosaka *et al.*, 2006). Enhanced protein synthesis results in antibiotic overproduction and cryptic gene activation. In contrast, *rsmG* mutants overexpress SAM synthetase, an enzyme that stimulates antibiotic production (Nishimura *et al.*, 2007; Zhao *et al.*, 2010).

During the past two decades, ribosome engineering has been widely used to activate or increase the production of secondary metabolites in many bacterial species. Through ribosome engineering, 180-fold production of actinorhodin in *S. coelicolor* was achieved (Wang *et al.*,

2008), as well as, 41-fold production of violacein in *E. coli* (Ahmetagic & Pemberton, 2011), 40-fold production of chloramphenicol in *S. coelicolor* (Gomez-Escribano & Bibb, 2011) and 109-fold production of tiancimycin D in *S.* sp. CB03234 (Zhuang *et al.*, 2019) among others. Apart from antimicrobials, this approach has been successfully utilised for enzyme (Kubo *et al.*, 2013; Liu *et al.*, 2013), ethanol (Suzuki *et al.*, 2015) or butanol (Chen *et al.*, 2012; Gao *et al.*, 2012) production. A recent study also used ribosome engineering to modulate primary metabolism, introducing mutations into *Propionibacterium shermanii* to increase vitamin B12 production (Tanaka *et al.*, 2017).

Ribosome engineering is a cheap, fast and easy approach compared to traditional and current strain improvement methods that involve mutagenesis (Bose, 2016), genome shuffling (Zhang *et al.*, 2002b; Gong *et al.*, 2009; Magocha *et al.*, 2018), heterologous expression (Xu & Wright, 2019) and metabolic engineering (Jeschek *et al.*, 2017). Most of these methods, require the understanding and establishment of a genetic system and specific molecular biology techniques for the targeted organisms (Zhu *et al.*, 2019). The application of ribosome engineering is straightforward and involves basic microbiology techniques for the isolation and screening of the drug-resistant mutants, which makes it highly accessible to researchers with different levels of experience and resources.

4.1.2 Bioscreening for antimicrobial natural products: state of the art.

Since the discovery of the first antibiotics almost a century ago, a series of methods have been developed to facilitate the discovery of novel secondary metabolites with antimicrobial activity. To date, mainly two different experimental strategies for the discovery of new natural products with antibacterial activity have dominated the field: biological and chemical screenings. Biological screenings are based on plating techniques in which an indicator or test strain is inoculated into/onto agar and an agar plug or culture extract from a grown culture of the isolate to be studied is placed on the agar, and then, 16-20 hours later, the clearance zone around the plug/extract is measured. Then, if antimicrobial activity is found, the active compound needs to be isolated and identified by chemical methods. On the other side, chemical screening approaches use physicochemical data such as mass spectrometry (MS) and UV spectra to identify new natural products. Culture extracts are analysed by liquid chromatography (LC) and the UV and MS spectra of the different peaks are compared against

existing natural products libraries to determine the novelty of the compounds (Leeds *et al.*, 2006).

Within biological screenings, different methods have been gradually adapted to target compounds of specific families or with specific modes of action and range from simple plate diffusion assays to more expensive and technical screenings (Table 4.1). The simplest form of plate diffusion assay also called whole-cell empiric screening involves growing strains in solid or liquid media to then test their plugs or extracts against a test strain for antimicrobial activity that appears in the form of a zone of inhibition (ZOI). This methodology does not assess whether the antimicrobial activity comes from one or more compounds, it does not provide information regarding the mechanism of action (MOA) nor discriminates known from novel antibiotics (Singh *et al.*, 2011). In order to obtain information regarding the mode of action, whole-cell reporter assays were developed later, in which the indicator strain/s used are reporter strains that have been engineered to help identify the MOA of antibacterial compounds. An interesting example was the development of a panel of *B. subtilis* reporter strains that were built by studying the transcriptional profile of *B. subtilis* after treatment with various antibiotics with different MOA. The promoters of those genes that were specifically upregulated in response to an antibiotic with a particular MOA were cloned in front of a luciferase reporter gene to yield five constructs amenable for pathway-specific high-throughput screening (Hutter *et al.*, 2004). Apart from allowing the identification of the MOA of specific secondary metabolites, these strains provide additional sensitivity and allow for the identification of bioactive compounds at concentrations below that which are antibacterial (Leeds *et al.*, 2006).

Other types of bioassays include screening using a drug-sensitive and single drug-resistant pair and cell-free screenings (Table 4.1). In the sensitive-resistant pair assays, the bioactivity of plugs or extracts is tested against a strain that is sensitive and a strain that is resistant to an antibiotic (both from the same species). This way, the compounds that target the binding site where resistance is mapped would be more active against the antibiotic-sensitive strain thus creating a larger ZOI than in the plate containing the antibiotic-resistant strain (Chait *et al.*, 2010). This method has shown special potential for the discovery of compounds from the same classes but with structural variations. In addition to that, cell-free antimicrobial assays were also developed during the last decade. These are not generally used to carry out primary screenings because most of the compounds that initially show antimicrobial activity then cannot reach their intracellular targets due to poor penetration or active efflux (Leeds *et al.*,

2006; Marchiaro *et al.*, 2008; Balouiri *et al.*, 2016). Despite this, carefully selected cell free assays can be powerful as secondary screenings when studying compounds with specific targets.

Table 4.1. Screening approaches used in antimicrobial drug discovery. Modified from (Singh *et al.*, 2011).

Approach	Cellular activity	Mode of action	Advantages	Drawbacks
Whole-cell empiric screening	✓	x	Coverage for all targets that affect bacterial growth	High rediscovery rate
Whole-cell reporter assays	✓	✓	Additional sensitivity	Interference of assay readout. False positives
Sensitive-resistant pair screening	✓	✓	Rapid discovery of molecules that bind to the same binding site and new analogues	Low sensitivity. Requires a drug-resistant strain
Cell-free enzyme	x	✓	Precise assays for specific targets	Most compounds then lack whole-cell antibacterial activity

Several more bioscreening approaches have been developed based on broth dilutions, microfluidics, ATP bioluminiscence or the use of automated instruments among others (Jorgensen & Ferraro, 2009; Khan *et al.*, 2019). However, antimicrobial plate diffusion assays are still the most prevalent approach followed in antimicrobial drug discovery laboratories, as they are easy to carry out and there is no need for specialised staff, materials are inexpensive and commonly available in microbiology laboratories and allow quick data acquisition and easy results analysis. Indeed, so many variations to this methodology have been published since the 1950s that the Clinical Laboratory Standards Institute reviews and updates the original procedure every three years to ensure standardisation, uniformity and reproducibility (CLSI, 2018).

4.1.3 Abyssomicin C mode of action and detection by plate diffusion assay.

Abyssomicin C and its atrop- isomer produced by *M. maris* AB-18-032 are effective antimicrobial compounds against strains such as *B. subtilis*, MRSA and *M. tuberculosis* (Riedlinger *et al.*, 2004; Freundlich *et al.*, 2010). Abyssomicin and atrop-abyssomicin C are antifolate compounds. They inhibit the synthesis of PABA within the chorismate pathway by

irreversibly binding to the aminodeoxychorismate synthase (ADCS) via Michael addition to a cysteine residue (Figure 4.1) (Keller *et al.*, 2007a, b).

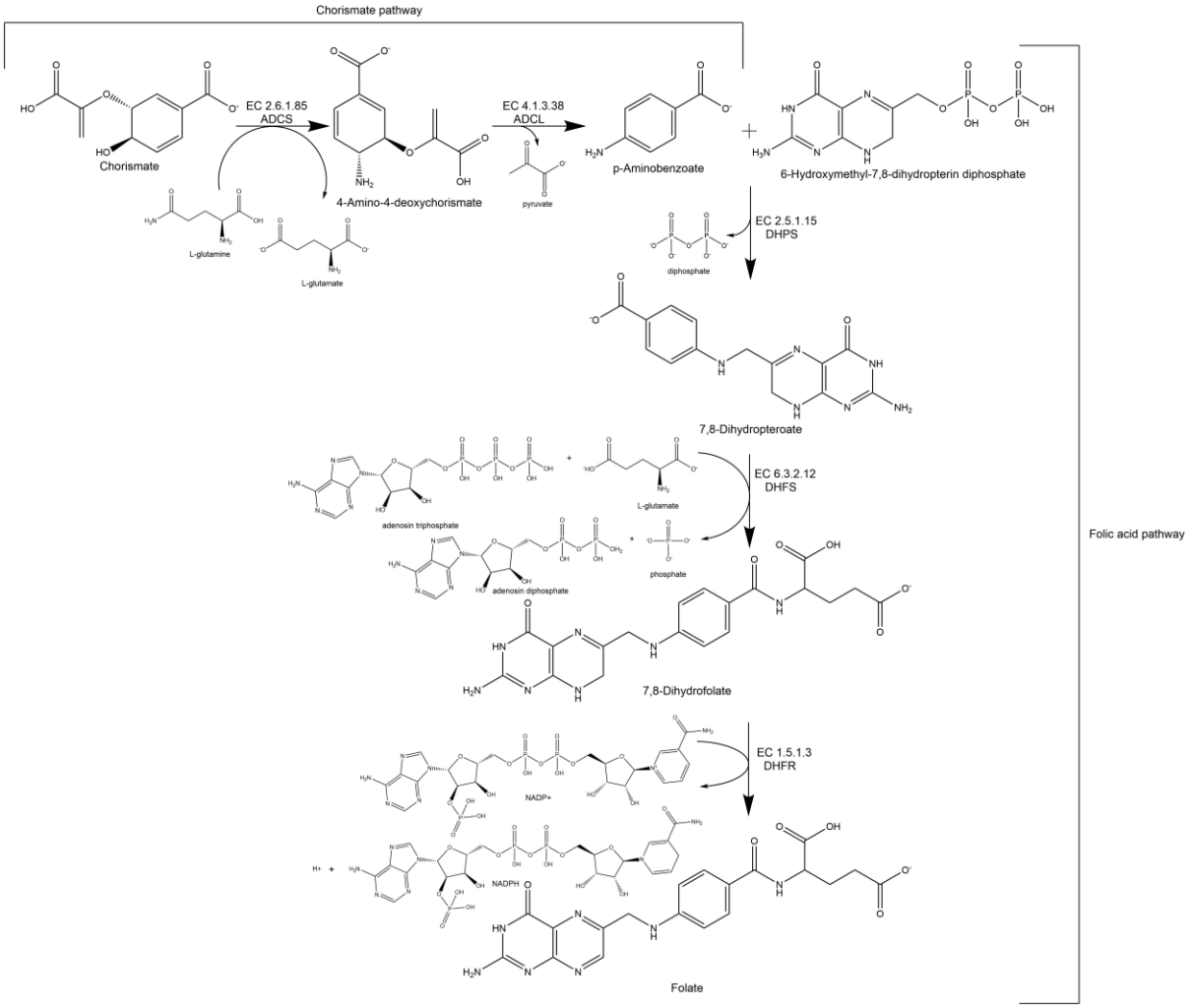

Figure 4.1. Schematic representation of the chorismate and folic acid pathways.

During PABA biosynthesis, chorismate and glutamine are aminated by ADCS, yielding glutamate and 4-amino-4-deoxychorismate (ADC). Then, PABA is generated from the aromatisation of ADC by 4-amino-4-deoxychorismate lyase (ADCL) with the loss of a pyruvate molecule (Figure 4.1). Interestingly, in many bacteria, ADC biosynthesis requires two separate enzymes, a glutamine amidotransferase (GAT) and the ADC synthase, whereas plants and lower eukaryotes display a bifunctional glutamine amidotransferase–aminodeoxychorismate synthase (GAT–ADCS) (Sadaka *et al.*, 2018). Previous work demonstrated that the oxabicyclooxane ring system of abyssomicin C and atrop-abyssomicin C mimics the chorismate transition state (Figure 4.2A) and binds to the Cys263 residue of the ADCS (Figure 4.2B) (Nicolaou *et al.*, 2009; Bihelovic & Saicic, 2012; Saicic *et al.*, 2014).

This mechanism is attributed to the α,β-unsaturated ketone unique to abyssomicin and atrop-abyssomicin C that is absent in the inactive abyssomicins (Sadaka *et al.*, 2018).

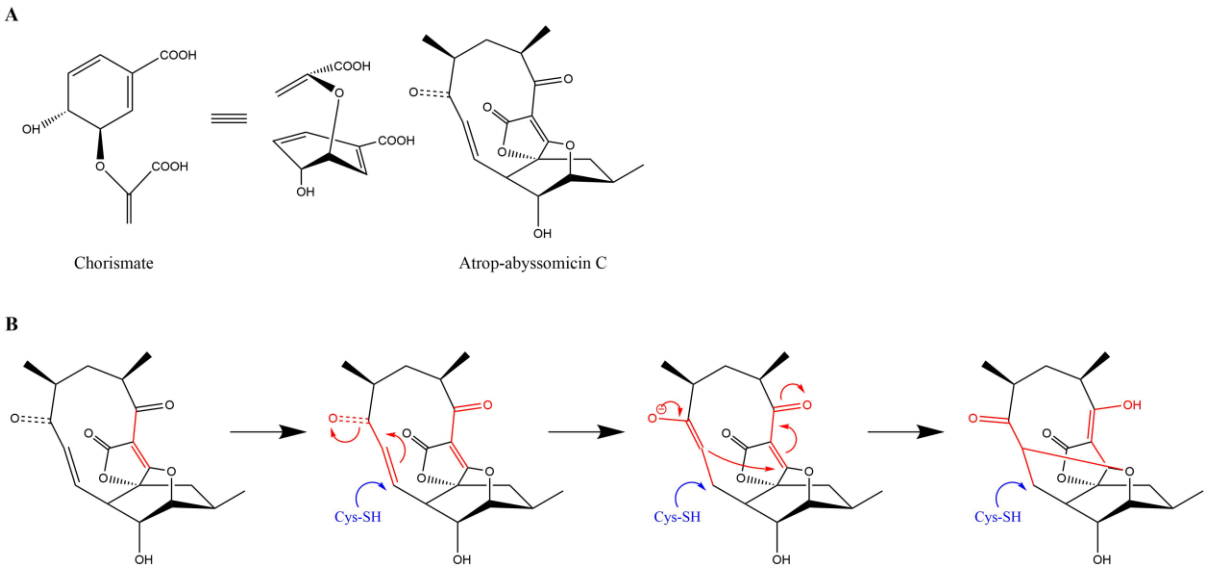

Figure 4.2. A) Structures of chorismate and atrop-abyssomicin C. **B)** Michael addition reaction by which atrop-abyssomicin C irreversibly binds to the Cys263 of ADCS. Michael acceptors are highlighted in red. Adapted from (Sadaka *et al.*, 2018).

M. maris is a slow growing marine isolate whose production capabilities under laboratory conditions yield a maximum of 58 mg/l abyssomicin C after 96 hours in 10 l stirred batch fermenters (Riedlinger *et al.*, 2004). Antimicrobial diffusion assays on plates are the most convenient method when carrying out quick routine antimicrobial assays, and although only abyssomicins C, J, 2 and atrop-abyssomicin C display antibiotic activity from the 38 abyssomicins characterised to date, their unique mode of action makes them easily identifiable when produced in enough amounts. Abyssomicin C was discovered during the screening of a library of *Actinobacteria* when trying to find antimicrobial secondary metabolites targeting the shikimate pathway and more specifically the biosynthesis of aromatic amino acids and PABA, the precursor of folic acid. An agar plate diffusion assay modified as an antagonism test was designed for this purpose, consisting in a growth inhibition reversal test against *B. subtilis* in a minimal medium growing in parallel the *B. subtilis* strain in the same agar medium but supplemented with 5 mM PABA. Extracts applied on a filter disk were tested on both agar plates and the extracts containing abyssomicin C caused an inhibition zone exclusively on the agar plate without PABA (Riedlinger *et al.*, 2004).

Despite the availability of this specific agar plate antimicrobial assay, the detection of abyssomicin C production by *M. maris* was found challenging at the beginning of this work. The first antimicrobial plate assays carried out in media containing PABA to restore growth of the test strain resulted in ZOI larger than those generated by the same plugs on plates without PABA. The addition of PABA into the media was being done after the media had been adjusted to pH 7.3 and autoclaved, thus decreasing the pH of the media. As a result of that, an investigation was started to find the conditions that eased the detection of small amounts of abyssomicin C in antimicrobial plate assays using four media with different compositions at pH 5, pH 7 and pH 9 with and without PABA to support the growth of the test strain.

In this chapter, first the application of ribosome engineering to increase abyssomicin C production in *M. maris* AB-18-032 is presented. A library of ribosome engineering mutants was generated and screened for enhanced antimicrobial activity. For the best producer mutants, *rpsL* and *rpoB* genes were sequenced and, after quantifying their production by analytical methods, the genome of the best producer was also entirely sequenced. Moreover, an attempt to optimise the detection conditions of abyssomicin C when screening through plate diffusion assays is described. This includes the study of how extracellular pH could be modulating the transport of metabolites across the bacterial cell membrane, subsequently influencing sensitivity to antimicrobial compounds.

4.2 Materials and methods.

4.2.1 Ribosome engineering.

M. maris AB-18-032 spores (~10^7 to 10^8) were spread on SGG agar (Gottardi *et al.*, 2011) containing separately three different concentrations of streptomycin, rifampin or chloramphenicol (5, 10 or 30 µg/ml). After cultivation for 20 days at 30 ºC, spontaneous drug-resistant mutants appeared. Isolated colonies were grown in SGG agar for 15 days at 30 ºC. In order to test for enhanced antimicrobial activity, 0.6 cm (diameter) agar plugs from each plate were placed onto MH agar plates with and without 5 mM PABA inoculated with *Bacillus subtilis* DSM 10 at OD_{600} = 1.00 following a 1:100 ratio. Due to the high number of ribosome engineering mutants tested, a single plug from each ribosome engineering mutant was used in the antimicrobial plate assays. Since abyssomicin C and atrop-abyssomicin C block the synthesis of PABA, when this compound is present in the media it rescues the growth of the test strain. That way, when a ZOI is caused by the effect of these abyssomicins, the plugs in the plates with PABA display a reduced ZOI in comparison to the plates without PABA. After overnight incubation at 37 ºC, the ZOI in mm of each plug was measured. Some drug-resistant mutants exhibiting greater antimicrobial activity than the *M. maris* wild-type strain were selected to be exposed to a second antibiotic (streptomycin, rifampin or chloramphenicol depending on the first exposure). Antimicrobial activity was then determined for the double resistant mutants and isolates showing the highest antimicrobial activity were used for further analysis. Genomic DNA was extracted (GenElute Bacterial Genomic DNA Kit; Sigma-Aldrich) and primers designed to amplify *rpsL*, *rpoB* and *gidB* genes (Table S4.1). Clean PCR products were sequenced by Sanger.

4.2.2 Quantifying abyssomicin production.

In order to quantify the abyssomicin production yield of *M. maris* AB-18-032 and the ribosome engineering mutants, the strains were grown in 50 ml of SGG in 250 ml baffled Erlenmeyer flasks at 30 °C and 200 rpm. Samples consisting of 5 ml of culture were taken after 24, 48, 72, 96 and 120 hours from inoculation and organic extractions were performed as described in Chapter 2. The identification of the compounds was based on mass spectrometry (accurate mass and ion fragmentation pathway, Figure 4.3). Abyssomicin C yield was calculated for each strain only for the day that showed higher production using a calibrated standard curve and

measuring the area under the curve. Due to limitations in equipment availability only one replicate per sample was used for the quantification.

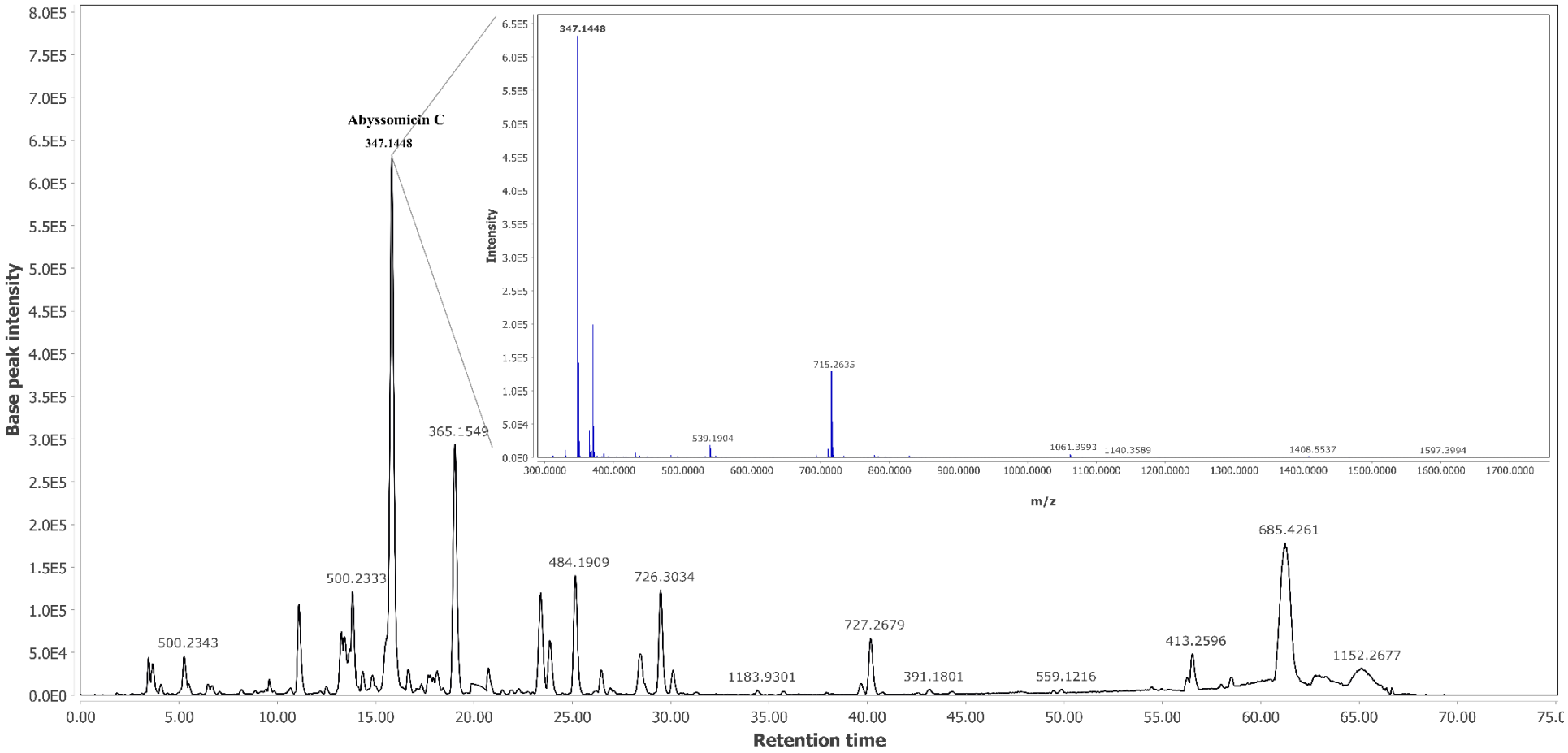

Figure 4.3. LC-MS model chromatogram of abyssomicin C from a fermentation extract of *M. maris* AB-18-032 grown in SGG at 30 ºC and 200 rpm for 3 days.

4.2.3 Genome sequencing, assembly and annotation.

The genomic DNA of *M. maris* AB-18-032 and the mutant exhibiting the highest abyssomicin C production was sequenced using Illumina HiSeq PE150 in order to identify any mutations, in addition to those frequently identified in *rpsL*, *rpoB* and *gidB* genes that could also led to a hyper-productive phenotype. The resulting paired end reads were trimmed using the pre-sets for Illumina adapters (Kmer length = 27), by quality (minimum quality = 20) , based on paired read overhangs (minimum overlap = 20) and by short reads (minimum length = 20 bp) using BBDuk. The remaining reads were assembled and mapped to the NCBI reference genome sequence of *M. maris* AB-18-032 (NC_015434.1) using Geneious Prime 2020.1. Mapping statistics can be found in Table S4.3 and Table S4.4. Mutations were called when the majority of the mapped reads presented a different nucleotide to that of the wild-type strain, that is the minimum variant frequency was higher than 0.5, and when the P-value was below 10^{-6}, since higher P-values flagged many other positions with lower coverage and sequencing quality due to the well known limitations of sequencing a genome rich in CG and repetitive regions.

4.2.4 Media screening to facilitate abyssomicin C detection by plate diffusion assay.

In order to ease the detection of abyssomicin C via antimicrobial plate diffusion assays, a screening of different media to support the growth of the antimicrobial indicator strain for antimicrobial susceptibility testing was carried out. The following media were prepared according to the manufacturer's instructions and adjusted to pH 5, pH 7 and pH 9 prior to sterilisation: Luria Bertani (LB) agar Miller (Sigma-Aldrich), MH agar (Sigma-Aldrich), Nutrient agar (NA) (Sigma-Aldrich) and CDM agar (Riedlinger *et al.*, 2004). All media were prepared with and without 5 mM PABA. Antimicrobial plate diffusion assays were carried out against *B. subtilis* DSM 10. For that, a flask containing 50 ml LB broth was inoculated with a single colony and grown overnight at 37 °C and 200 rpm. The culture's optical density was adjusted to OD_{600} = 1.00 and 250 µl were mixed with 25 ml of warm media. After solidifying, three plugs from *M. maris* AB-18-032 and each of the ribosome engineering mutants grown in SGG for 14 days at 30 °C were placed on top of each plate. Plates were incubated overnight at 37 °C and the ZOI that appeared were measured.

4.2.5 Analysis of the impact of pH on the sensitivity of plate diffusion assays.

In order to investigate if antimicrobial plate diffusion assays could be optimised to enhance their sensitivity, the influence of the media composition and the pH was investigated. Antimicrobial plate diffusion assays against *B. subtilis* DSM 10, *S. aureus* DSM 2569 and *E. coli* DSM1103 were carried out in LB agar, MH agar, NA and CDM agar at pH 5, pH 7 and pH 9 for five different commercial antibiotics with different structures and modes of action. Plates were prepared as described before and filter paper disks were placed on top of each plate. These were soaked with 5 µl of several different serial dilutions of the following antibiotic concentrated stocks: nalidixic acid (25 mg/ml), carbenicillin (50 mg/ml), thiostrepton (50 mg/ml), rifampicin (50 mg/ml) and tetracycline (10 mg/ml). Five replicates were used for each antibiotic. Plates were incubated overnight at 37 °C and the ZOI that appeared were measured. The detection limit for each antibiotic was established as the concentration of antibiotic that generates a visible ZOI of 0.5 mm.

In order to study how each media supports the growth of the antimicrobial susceptibility test strain, growth curves of *B. subtilis*, *S. aureus* and *E. coli* in LB, MH, Nutrient Broth (NB) (Sigma-Aldrich) and CDM at pH 5, pH 7 and pH 9 were carried out using a CLARIOstar Plus microplate reader (BMG LABTECH). A flask containing 50 ml LB broth was inoculated with a colony of *B. subtilis*, *S. aureus* or *E. coli* and grown overnight at 37 °C and 200 rpm. This culture was used to inoculate the media above (1:100 dilution) and growth was monitored in each media in triplicate in a 96-well plate at 37 °C with 300 rpm shaking between measurements and 500 rpm shaking for 30 seconds before each measure was taken.

Then, with the aim of analysing whether the results obtained from the antimicrobial disk assays could be caused by the change in agar density at each different pH, growth curves of *B. subtilis*, *S. aureus* and *E. coli* were carried out in MH at pH 5, pH 7 and pH 9 containing nalidixic acid, carbenicillin, thiostrepton, rifampicin and tetracycline. Each of these were carried out in triplicate. Various concentrations of each antibiotic against each test strain were tested. The most descriptive results were obtained using the concentrations in Table 4.2.

Table 4.2. Concentrations of antibiotics used for the growth curves.

	Nalidixic acid (µg/ml)	Carbenicillin (µg/ml)	Rifampin (µg/ml)	Tetracycline (µg/ml)	Thiostrepton (µg/ml)
B. subtilis	2.5	50	0.05	0.1	0.05
S. aureus	25	0.05	0.005	1	500
E. coli	$2.5x10^{-7}$	0.5	0.5	0.01	0.05

Then, in order to study how pH affects antibiotic transport across the cell membrane, a LC-MS based assay was designed based on previous work for the directed detection of antimicrobial transport across membrane (Zhou *et al.*, 2015). Calibrated standard curves were prepared for each compound with seven points across the concentration range of 50 µg/ml to 50 pg/ml. The serial dilutions of the standards were injected in a mass spectrometer ESI microTOF-Q II (Bruker Daltonics). Separation of the compounds was performed by a Raptor™ ARC-18 (2.7 µ, 150 x 2.1 mm, Restex) chromatographic column protected by a guard column. The mobile phase was ultrapure water (solvent A) and acetonitrile (solvent B), both with 0.1% formic acid (v/v). The elution gradient used was: 0–1 min 5-15% B; 1-3 min 15-50% B; 3-4.20 min 50-70% B; 4.20-4.40 min 70-80% B; 4.40-6.30 min 80-98% B; 6.30-6.50 min 98-5% B and 6.50-7 min 5% B followed by column washing and reconditioning (1 minute). The flow rate was 0.2 mL/min. The column oven was at 35 °C and the injection volume was 3 µL. The UV analyses were performed in the wavelength range of 190–600 nm, with the mass spectrometer operating in positive mode (m/z 50–2800).

For the antibiotic uptake assay, *E. coli* was grown overnight in MH at pH 7 at 37 °C and 200 rpm. Then, the culture was diluted 1:50 in fresh MH at pH 7 and grown to mid-log phase (OD_{600} = 0.4-0.6). Cells were centrifuged for 10 minutes at 5,000 x *g* and the pellet was washed with an equal volume of minimal medium (MM) at pH 5, pH 7 or pH 9. For *E. coli* the MM used was M9 supplemented with 0.2% glucose. Cells were centrifuged for 10 minutes at 5,000 x *g* and the pellets were resuspended in MM at pH 5, pH 7 or pH 9 to bring cells to an OD_{600} of 4.0. For the uptake assay, 0.5 ml of resuspended cells were mixed with 0.5 ml of MM at pH 5, pH 7 or pH 9 containing a single antibiotic (25 µg nalidixic acid, 50 µg carbenicillin 50 µg thiostrepton, 50 µg rifampicin or 10 µg tetracycline). For each pH, a blank without bacteria was prepared. Tubes were then incubated at 37 °C and 200 rpm for 30 minutes and, after that, centrifuged at 3,700 x *g* for 10 minutes (Figure 4.4). Finally, the supernatant was transferred to an HPLC vial and antibiotics were quantified by LC-MS as described before. The area under the curve was calculated to determine the amount of antibiotic that had cross the cell membrane.

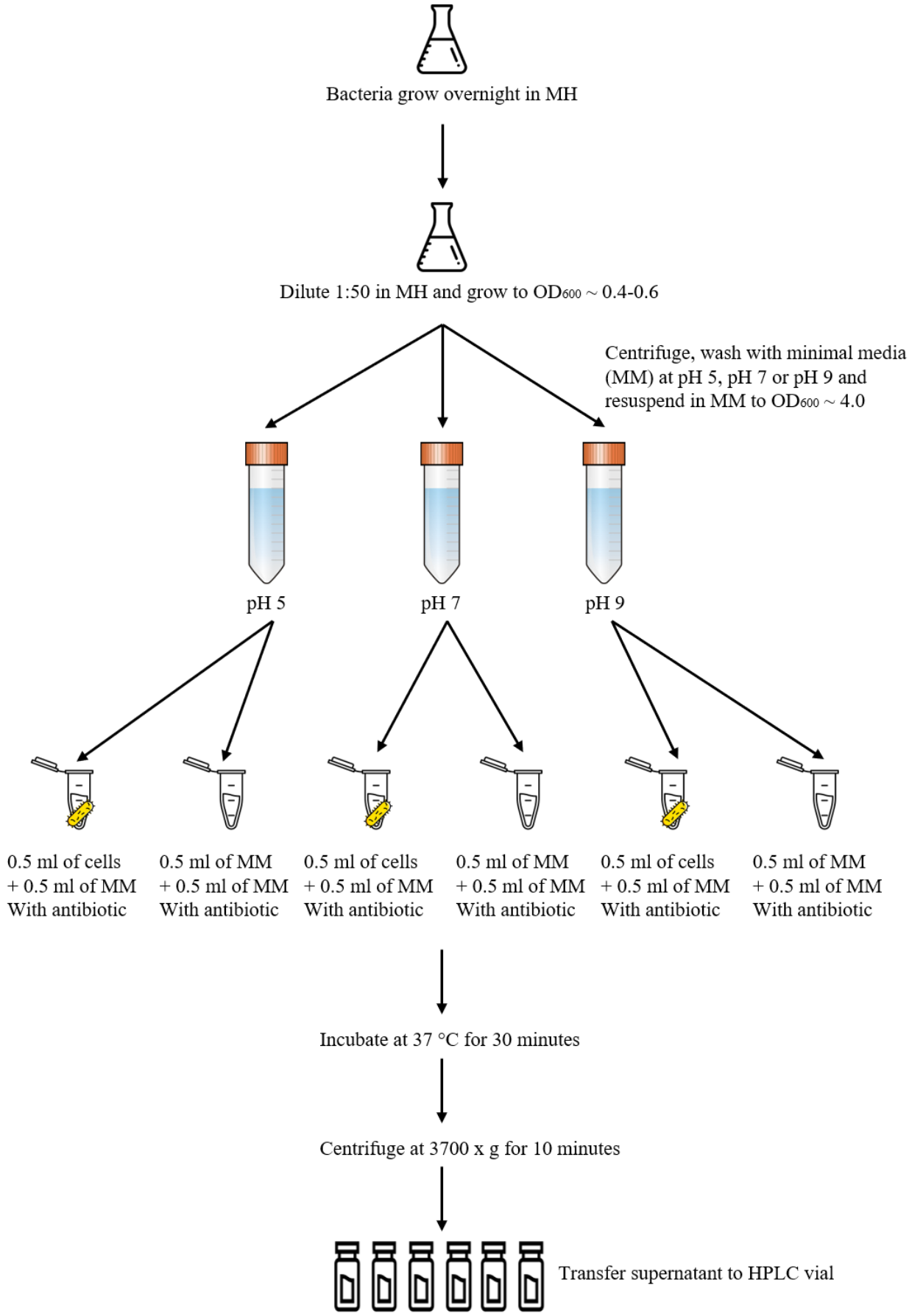

Figure 4.4. Schematic representation of the LC-MS based antimicrobial uptake assay.

Finally, in order to study if it would be possible to apply the knowledge generated by those previous tests to improve the antimicrobial strain screening assay traditionally used in antimicrobial drug discovery, ten *Actinobacteria* from agar plates and organic extracts were screened. These strains, designated here as isolates 1-10 (Table 4.3), were grown in GYM (per litre: glucose 4.0 g, yeast extract 4.0 g, malt extract 10.0 g, $CaCO_3$ 2.0 g, agar 12.0 g; pH 7.2) plates for 14 days at 30 °C. Plugs (6 mm diameter) from these plates were used in antimicrobial diffusion assays against *B. subtilis* and *S. aureus* in LB agar, MH agar, NA and CDM agar plates at pH 5, pH 7 and pH 9. The same strains were grown in 250 ml baffled flasks containing 50 ml of GYM for 5 days at 30 °C and 200 rpm. Then organic extractions were carried out with ethyl acetate as described before and 5 µl of the extracts were used to soak the filter paper disks that were placed on top of the plates. Plates were incubated overnight at 37 °C and the ZOI that appeared were measured. Due to time limitations this experiment was carried out only once with a single replicate.

Table 4.3. *Actinobacteria* strains used to challenge the indicator strains in the antimicrobial plate diffusion assays.

Strain number	Phylogenetic affiliation	Origin
1	*Streptomyces rengensis* DSM 40551	DSMZ
2	*Streptomyces incarnatus* strain NRRL 8089	Dr. Tamura Lab (Okayama University)
3	*Streptomyces* sp. AF1	Dr. Montero-Calasanz Lab (Newcastle University)
4	*M. maris* ribosome engineering mutant mm21	This thesis
5	*Streptomyces lividans* K4-114	
6	*Streptomyces lividans* pESAC13_*aby*	This thesis
7	*Streptomyces lividans* pESAC13_*aby_kasOp**	This thesis
8	*Streptomyces albus* J1074	
9	*Streptomyces albus* pESAC13_*aby*	This thesis
10	*Streptomyces albus* pESAC13_*aby_kasOp**	This thesis

4.3 Results.

4.3.1 Ribosome engineering.

In order to enhance abyssomicin production in *M. maris*, a system for which genetic tools for efficient genome editing have not been developed yet, ribosome engineering approach was used. Over 140 single and double drug-resistant mutants were isolated from plates containing sublethal concentrations of streptomycin, rifampin and/or chloramphenicol. Among these spontaneous mutants, a small number displayed phenotypical differences in comparison to the wild-type strain, including lose or reduction of pigmentation and changes in the shape of the edges of the colonies. Changes in growth rate were also appreciated for some mutants, including both strains with faster and slower growth than the parental strain. The antimicrobial production potential of these strains was studied via plug diffusion assay in minimal media plates with and without PABA (Riedlinger *et al.*, 2004). Nine mutants generated ZOI larger than those produced by *M. maris* AB-18-032 and, in all cases, the ZOI in the plates containing PABA were at least 2 mm smaller (Figure 4.5).

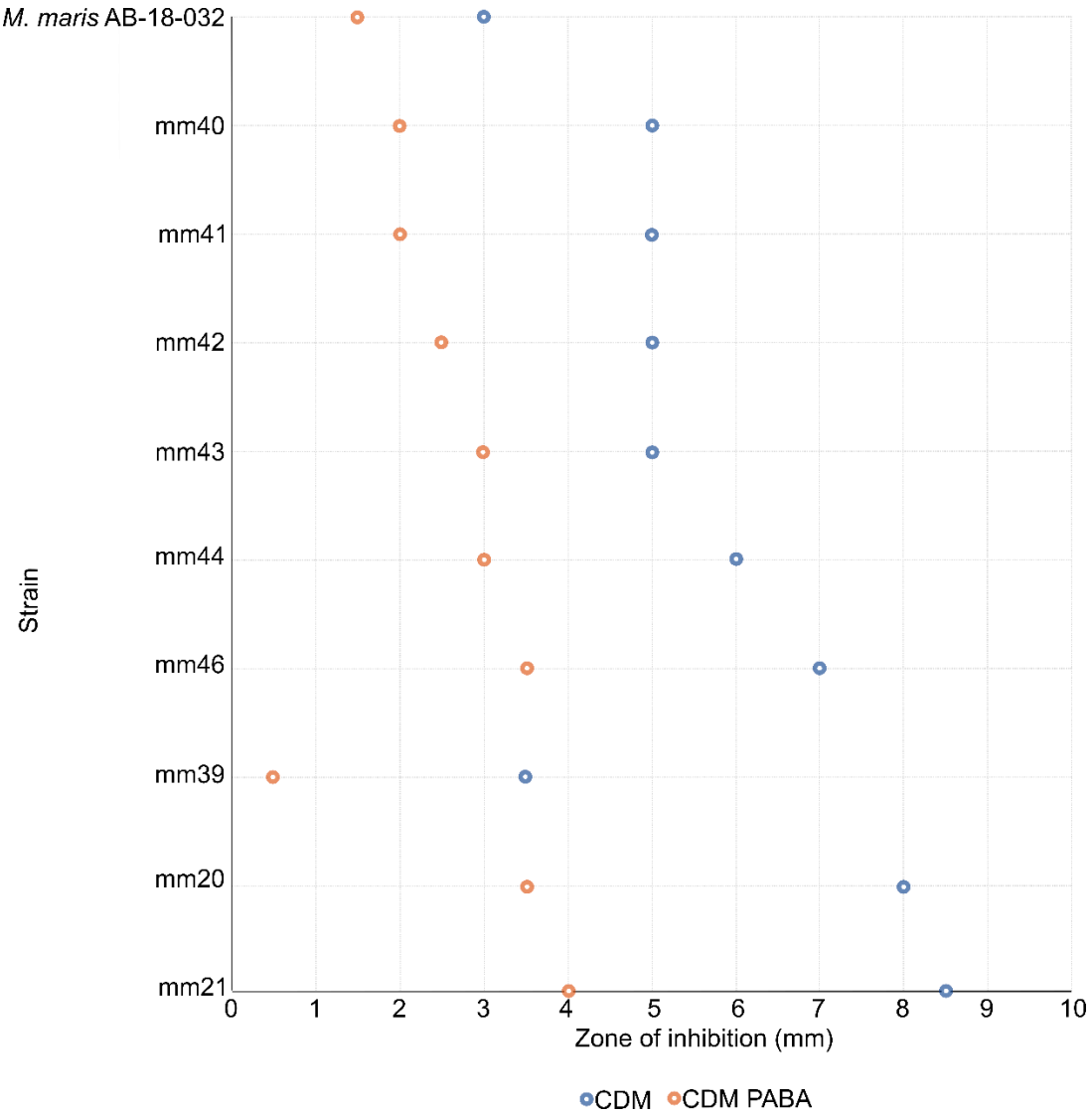

Figure 4.5. Zones of inhibition generated by the sequenced ribosome engineering mutants of *M. maris* AB-18-032 by plug diffusion assay against *B. subtilis*. Due to the high number of ribosome engineering mutants tested, a single plug from each ribosome engineering mutant was used in the antimicrobial plate assays.

The mutants exhibiting higher antimicrobial activity than the wild-type strain were subjected to further analysis. For the streptomycin resistant mutants, the *rpsL* gene was sequenced, as well as the *rpoB* gene for the rifampin resistant mutants. For the chloramphenicol resistant mutants *rpsL*, *rpoB* and *gidB* were sequenced. Three streptomycin resistant mutants, all obtained under different concentrations of streptomycin, displayed an amino acid substitution at the position 88 in RpsL, one of them being K88E and two of them K88R (Table 4.4). Two mutants obtained under the presence of 30 µg/ml of streptomycin displayed three identical consecutive nucleotide substitutions that resulted in the amino acid substitution P91Y. Only one rifampin and one chloramphenicol resistant mutant with enhanced antimicrobial activity in comparison to *M. maris* AB-18-032 were obtained. The rifampin resistant mutant displayed

a H419R substitution, whereas no mutations were found in *rpsL*, *rpoB* nor *gidB* genes for the chloramphenicol resistant mutant, suggesting spontaneous mutations could have taken place in a different component within the ribosomal machinery. Two double drug-resistant mutants were obtained after exposing the chloramphenicol resistant mutant mm43 to 30 µg/ml of rifampin and the amino acid substitutions H419Y and A424T were found (Table 4.4).

Table 4.4. Mutations found in the sequenced ribosomal genes of the ribosome engineering mutants of *M. maris* AB-18-032.

Strain	Parent strain	Antibiotic	Concentration (µg/ml)	Nucleotide substitution (*gene*)	Amino acid substitution
mm40	*M. maris* AB-18-032	streptomycin	5	A262G (*rpsL*)	K88E
mm41		streptomycin	10	A263G (*rpsL*)	K88R
mm42		streptomycin	30	A263G (*rpsL*)	K88R
mm44		streptomycin	30	C271T (*rpsL*) C272A (*rpsL*) C273T (*rpsL*)	P91Y
mm46		streptomycin	30	C271T (*rpsL*) C272A (*rpsL*) C273T (*rpsL*)	P91Y
mm39	mm43	rifampin	10	A1256G (*rpoB*)	H419R
mm43		chloramphenicol	5	-	-
mm20		rifampin	30	C1255T (*rpoB*)	H419Y
mm21		rifampin	30	G1270A (*rpoB*)	A424T

Then, in order to confirm enhanced abyssomicin C production by analytical methods, fermentations were carried out in SGG media and samples were extracted during five consecutive days in order to identify the day at which abyssomicin C yield was higher for each strain. In the case of *M. maris* AB-18-032, the single mutants mm41 and mm42, and the double mutant mm21, abyssomicin C production was highest on day three, with all abyssomicin C converted into abyssomicin D by day five (Figure S4.1, S.4.3, S.4.4, S.4.6). For the single mutant mm40 and the double mutant mm20, the highest abyssomicin C production was found after five days of fermentation (Figure S4.2, S.4.5). Overall, the maximum abyssomicin C production yield of each strain was: 186 mg/l for *M. maris* AB-18-032, 608 mg/l for mm40, 632 mg/l for mm41, 283 mg/l for mm42, 117 mg/l for mm20 and 296 mg/l for mm21 (Figure 4.6).

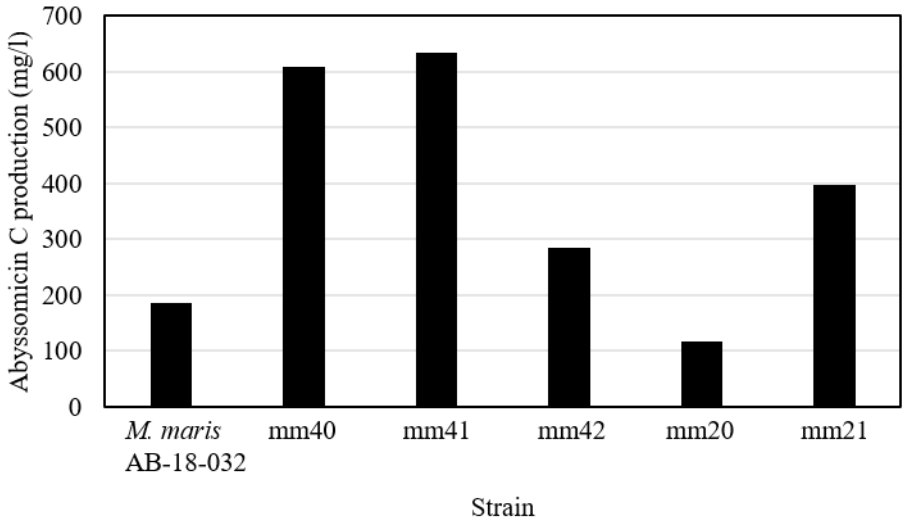

Figure 4.6. Abyssomicin C production yield (mg/l) of *M. maris* AB-18-032 and its ribosome engineering mutants. Due to limitations in equipment availability only one replicate per sample was used for the quantification.

Interestingly, mm40 and mm41 displayed a similar abyssomicin C yield, although they presented a different amino acid substitution in RpsL, and mm42, also a streptomycin resistant mutant that presented the same point mutation as mm41 produced half its yield (Table 4.4, Figure 4.6). On the other side, the double mutants mm20 and mm21 that had shown the largest ZOI by plug diffusion assay produced less abyssomicin C than the single mutants mm40 and mm41 (Figure 4.6).

Finally, in order to investigate whether the best abyssomicin C producer, mm41, harboured mutations other than in *rpsL*, *rpoB* or *gidB* genes that could also contribute to the enhanced abyssomicin C production, the genomes of mm41 and its parent strain, *M. maris* AB-18-032, were sequenced. Two frame-shifts in two different transposases were identified, as well as, three amino acid substitutions in a DNA translocase FtsK, five amino acid substitutions in a D-alanyl-D-alanine carboxypeptidase, one point mutation in the promoter sequence of the MurR/RpirR family transcriptional regulator of the *ngcEFG* operon involved in N-acetylglucosamine import and, interestingly, one point mutation in the intergenic region between AbyU and AbyI where potentially both terminators should be (Table 4.5). A complete table that includes the silent mutations and mutations in intergenic regions can be found in Table S4.2.

Table 4.5. Mutations found in the genome of the ribosome engineering mutant mm41.

Change	Location	CDS	Amino Acid Change	Protein Effect	Coverage	Variant Frequency	P-Value
(C)2 -> (C)3	307511	IS110 family transposase CDS		Frame Shift	26	76.90%	$2.20E^{-39}$
T -> A	1634033		S -> T		761	50.10%	0
CATC -> GCTG	1634047	DNA translocase FtsK CDS	SI -> SL	Substitution	836 -> 850	53.8% -> 55.3%	0
CGC -> AAG	1634588		R -> K		776 -> 786	55.6% -> 55.9%	0
G	2485903	IS21 family transposase CDS		Frame Shift	27	85.20%	$6.00E^{-13}$
G -> C	2613622	Intergenic region between AbyU and AbyI			50	54.00%	$3.40E^{-81}$
GC -> AA	3423252		R -> Q		417 -> 418	52.0% -> 52.2%	0
T -> C	3423267	D-alanyl-D-alanine carboxypeptidase CDS	V -> A		393	51.70%	0
TGCG -> ACCC	3423298		GA -> GP	Substitution	420 -> 421	52.3% -> 53.2%	0
GC -> CG	3423319		GL -> GV		423 -> 424	51.1% -> 51.2%	0
A -> G	3423348		Q -> R		398	52.50%	0
A -> C	4678236	Promoter sequence of MurR/RpirR family transcriptional regulator of the *ngcEFG* operon			282	100.00%	0
T -> C	5937058	30S ribosomal protein S12 CDS	K -> R	Substitution	370	100.00%	0

4.3.2 Optimisation of the plug diffusion assay conditions for abyssomicin C detection.

In order to enhance the performance of antimicrobial plug diffusion assays for the detection of small amounts of abyssomicin C, a media screening was carried out. Plates of LB, NA, MH and CDM at pH 5, pH 7 and pH 9 with and without PABA were used to grow the antimicrobial test strain and to rescreen *M. maris* AB-18-032 and its ribosome engineering mutants. *M. maris* AB-18-032 plugs generated ZOI of up to 8 mm when *B. subtilis* was grown in LB at pH 9 and up to 6 mm in MH at pH 9 (Figure 4.7A). A similar pattern was observed for the plugs of mm40, mm41 and mm42, that displayed larger ZOI in LB, NA and MH all at pH 9 (Figure 4.7B-D). Although the plugs from mm20 and mm21 also generated larger ZOI in LB, NA and MH at pH 9, it was possible to observe ZOI of up to 10 cm in CDM at pH 5, that contrasts with the fact that in the other media at pH 5 no ZOI were observed (Figure 4.7E, F). It was observed that in LB, NA and MH at pH 5 and pH 7, *B. subtilis* grew more abundantly closer to the plugs, whereas growth in CDM was uniform at all three pH (data not shown). Abyssomicin C antibacterial activity was rescued in the presence of 5 mM PABA regardless the media and the pH (Figure 4.7).

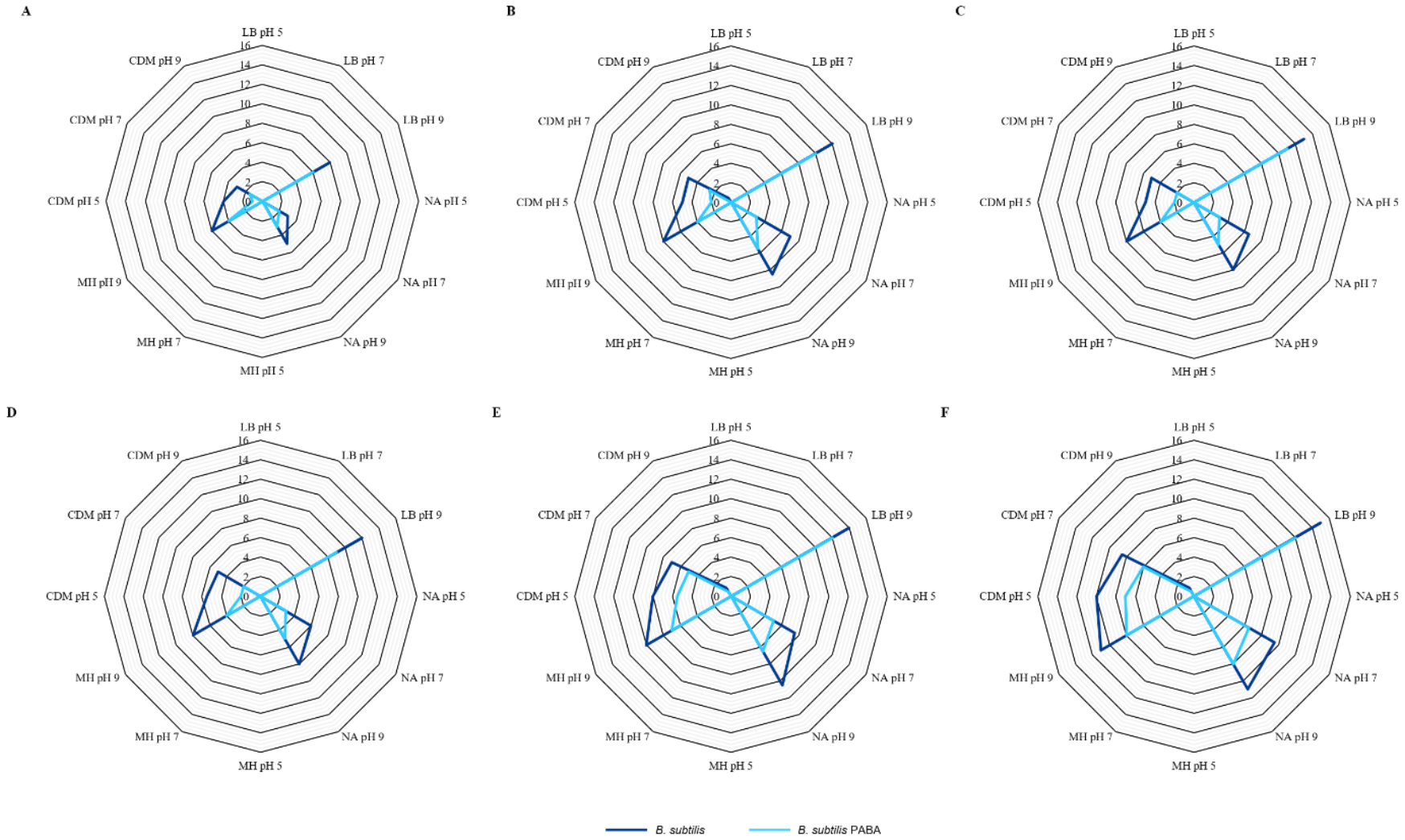

Figure 4.7. Zones of inhibition (mm) generated by plugs of **A)** *M. maris* AB-18-032, **B)** mm40, **C)** mm41, **D)** mm42, **E)** mm20 and **F)** mm21 in plug diffusion assays using different media and different pH to support the growth of the test strain *B. subtilis*. Figure displays the average ZOI calculated from the ZOI displayed by three replicates of each strain.

4.3.3 Analysis of the influence of pH on antimicrobial plate diffusion assays.

In order to investigate how the sensitivity of antimicrobial plate diffusion assays varies depending on the media used to support the growth of the test strain, stocks of five common commercial antibiotics with different structures and modes of action were used to carry out plate diffusion assays in MH, LB, NA and CDM at pH 5, pH 7 and pH 9. Serial dilutions of the concentrated stocks were used to soak filter paper disks and the detection limit for each antibiotic was established as the concentration of antibiotic that generated a visible ZOI of 0.5 mm.

When *B. subtilis* was used as test strain, the detection limit of nalidixic acid was approximately 100 ng in MH, LB and NA at pH 5, whereas when those same media were at pH 7, nalidixic acid was detected only at over 500 ng of nalidixic acid per disk. At pH 9 it was not possible to detect any antimicrobial activity despite increasing the concentration of nalidixic acid used. These differences in detection were not appreciated in CDM, in which at all three pH values the detection limit was around 200 ng of nalidixic acid (Figure 4.8). In the detection of carbenicillin, the four media at pH 5 showed to be more sensitive and up to 20 ng were detected. At pH 7 and pH 9 the detection limit was at 60 ng in LB, NA and CDM and around 120 ng in MH. Rifampicin detection limit was around 5 ng in MH, LB and NA at pH 5 and went up to 30 ng in MH at pH 7 and pH 9 and 15 ng LB and NA at pH 7 and pH 9. Rifampicin detection limit in CDM at pH 5 was 7 ng, whereas 15 ng were needed for detection at pH 7 and pH 9. Tetracycline detection followed a similar behaviour as nalidixic acid, and whereas 20 ng of tetracycline were detectable at pH 5 in MH, LB and NA, ten times more tetracycline were still undetectable at pH 9. In CDM, no significant differences in detection were appreciated despite the different pH values. Finally, thiostrepton detection limit was approximately 5 ng regardless of the media and the pH except at MH pH 9 were no detection was observed below 23 ng (Figure 4.8).

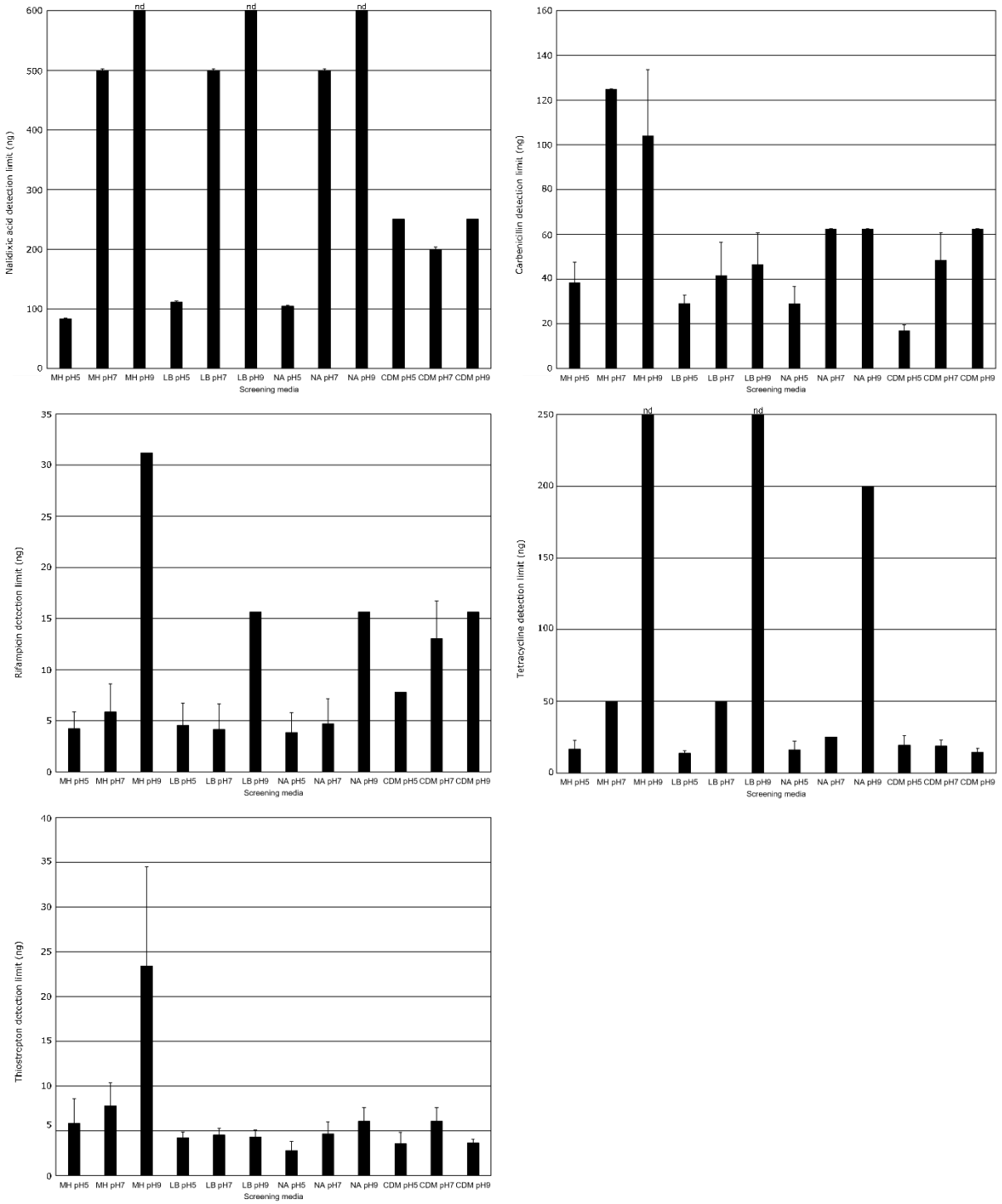

Figure 4.8. *B. subtilis* detection limit of the commercial antibiotics nalidixic acid, carbenicillin, rifampicin, tetracycline and thiostrepton in MH, LB, NA and CDM at pH 5, pH 7 and pH 9. n.d. stands for not detected. Five replicates of each antibiotic were used.

The same experiment was also carried out using *S. aureus* as test strain, with the exception that CDM did not support the growth of this strain so only MH, LB and NA were used. In this case, ZOI of 0.5 mm were generated under the presence of 25 ng of carbenicillin in all media at pH 5, whereas at least 250 ng were needed for it to be detected at pH 9 (Figure 4.9). A similar

detection pattern was observed for the detection of tetracycline. In the case of rifampicin, no differences in detection were observed when growing *S. aureus* at pH 5 and pH 7 but three times more was needed in order for it to be detected at pH 9. Finally, thiostrepton detection was very similar to that observed when using *B. subtilis* as antimicrobial test strain and differences in detection in relation to pH were only appreciated in MH, where at least 160 ng of thiostrepton were needed for the detection at pH 9 whereas 30 ng and 50 ng were sufficient for its detection at pH 5 and pH 7 respectively (Figure 4.9). This test was also carried out for *E. coli* as test strain, however it would need to be repeated with higher concentrations of all antibiotics, as with the concentrations tested no ZOI appeared.

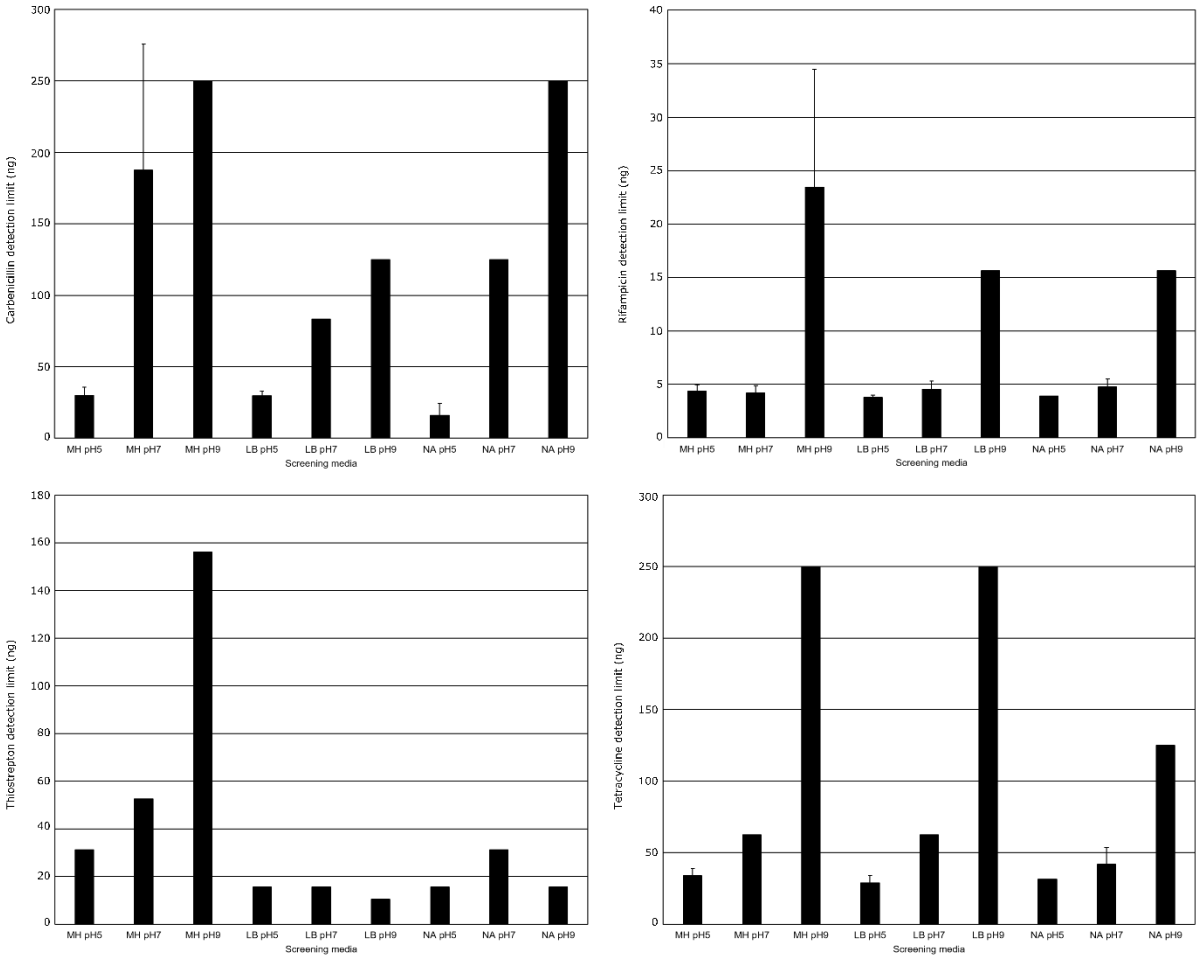

Figure 4.9. *S. aureus* detection limit of the commercial antibiotics carbenicillin, rifampicin, tetracycline and thiostrepton in MH, LB and NA at pH 5, pH 7 and pH 9. Five replicates of each antibiotic were used.

In order to study how each media supports the growth of the test strains and if the discrepancies in sensitivity could be a consequence of differential growth rate and biomass accumulation, growth curves of *B. subtilis*, *S. aureus* and *E. coli* in LB, MH, NB and CDM at pH 5, pH 7 and

pH 9 were carried out. The growth curves of *B. subtilis* showed how the growth of this strain was specially lagged in LB and NB at pH 9 and in MH at pH 5, whereas in CDM the growth was minimal regardless of the pH (Figure 4.10A). In the case of *S. aureus*, differences in growth rate and biomass accumulation were observed in LB and MH although these were more accentuated in NB (Figure 4.10B). As observed before CDM did not support the growth of *S. aureus*. For *E. coli*, the differences in growth were minimal in LB, NB and MH at pH 5, pH 7 and pH 9, whereas in CDM at pH 5 the stationary phase was reached at an OD_{600} = 0.6 and at pH 7 the stationary phase was reach at an OD_{600} = 1.25, similar to pH 9 although the lag phase was 2 hours longer (Figure 4.10C).

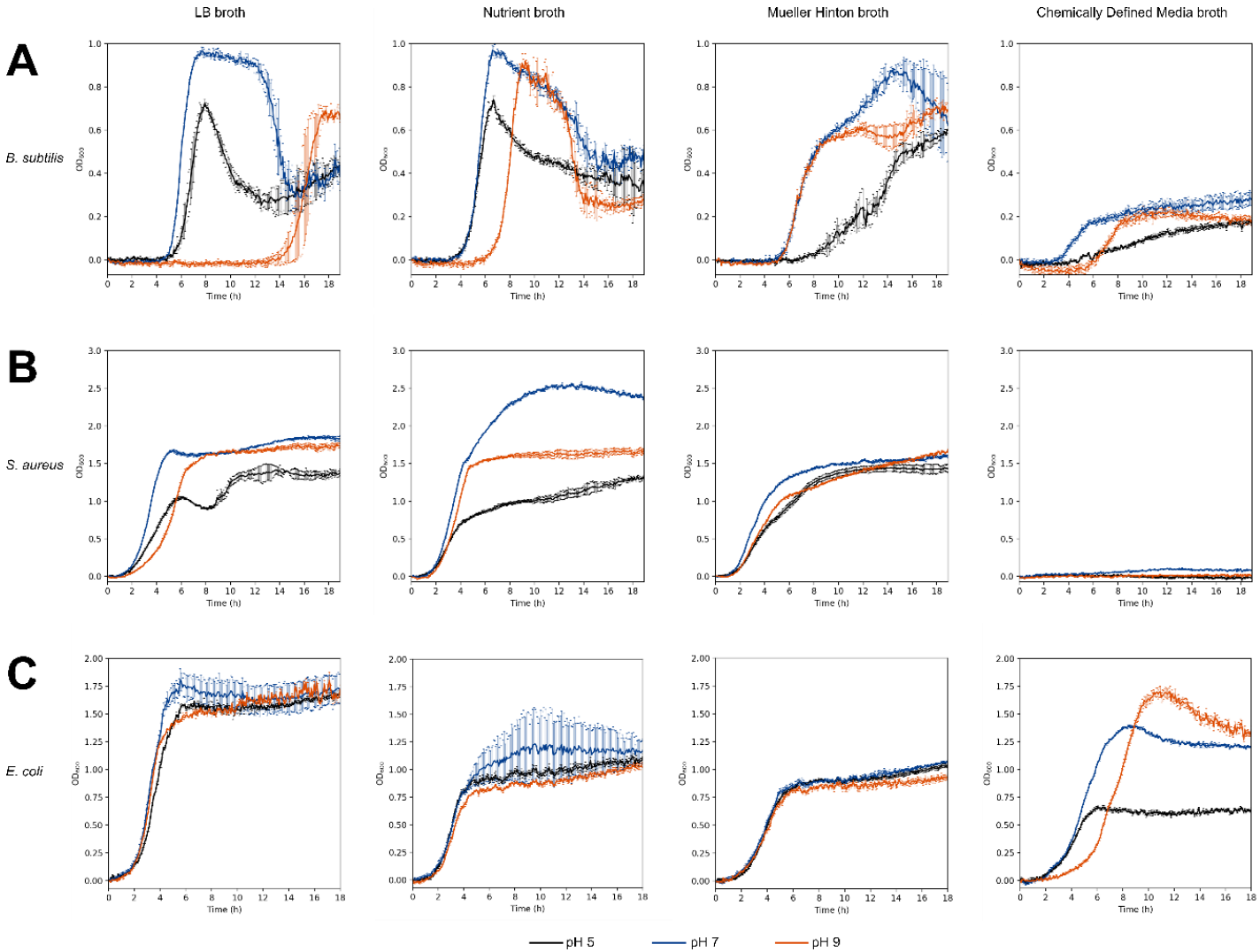

Figure 4.10. Growth curves in LB broth, Nutrient broth, MH broth and CDM broth of **A)** *B. subtilis*, **B)** *S. aureus* and **C)** *E. coli*. Black indicates pH 5, blue indicates pH 7 and orange indicates pH 9. All the growth curves were carried out in triplicate.

With the aim of analysing whether the sensitivity of the antimicrobial disk assays could be related to differential diffusion rates through the agar due to the change in density of the agar at each different pH, growth curves of *B. subtilis*, *S. aureus* and *E. coli* were carried out in MH at pH 5, pH 7 and pH 9 containing the five commercial antibiotics used before in the antimicrobial plate diffusion assays. From this experiment two main different observations were made. Growth of all the strains under the presence of all antibiotics was observed when the media was adjusted to pH 9 (Figure 4.11). For example, *B. subtilis* growth was inhibited in the presence of nalidixic acid when the media was at pH 5 and pH 7, whereas at pH 9 the growth curve was similar to that without antibiotic. This indicates that for that given concentration, strains growing at pH 5 and pH 7 are more sensitive to the antibiotic. These results match the behaviour observed in the agar plates sensitivity test (Figure 4.8). On the other hand, pH is responsible for the ionisation state (the charge and its intramolecular distribution) of small molecules and thus the chemical microspecies or protonation states of a molecule under varying pH conditions. The number of microspecies increases exponentially with the number of functional groups in the molecule and microspecies distribution diagrams determine the concentration of each microspecies at any pH value (Mazák & Noszál, 2016). The microspecies distribution of nalidixic acid (Figure S4.8) indicates that the microspecies present in the media at pH 7 and pH 9 is the same and remains negatively charged, which at pH 9 means that negatively charged molecules cross the membrane with more difficulty. Similar results were observed in the presence of thiostrepton, that inhibited the growth of *B. subtilis* at pH 5 and pH 7 and duplicated the lag phase time at pH 9. Tetracycline at 0.1 µg/ml completely inhibited the growth of *B. subtilis* at pH 5 and pH 7, whereas at pH 9 the growth curve was similar to that without antibiotic. As with nalidixic acid, these results match the results obtained from the agar plate diffusion assays (Figure 4.8) and the microspecies present at each pH (Figure S4.10), as at pH 9, with three negative charges in the molecule the cytoplasmic membrane would be much less permeable. Finally, under the presence of 0.05 µg/ml of rifampin, *B. subtilis* growth is completely inhibited at pH 5, has two times longer lag phase at pH 7 and grows as the control without antibiotic at pH 9. This matches the agar plate diffusion assays and fits with the microspecies present at each pH (Figure S4.9).

Results representing *B. subtilis* sensitivity to carbenicillin followed a different pattern, as *B. subtilis* grew despite the media pH although at different growth rates and with extended lag and log phases. Interestingly, carbenicillin is the only antibiotic tested that inhibits cell wall biosynthesis by acylating the transpeptidase C-terminal domain, and thus, it does not need to

penetrate inside the cell to exhibit its activity (Table S4.5). It was not possible to find a suitable concentration of thiostrepton that would partially inhibit the growth of *E. coli* without precipitating and influencing the plate reader measurements.

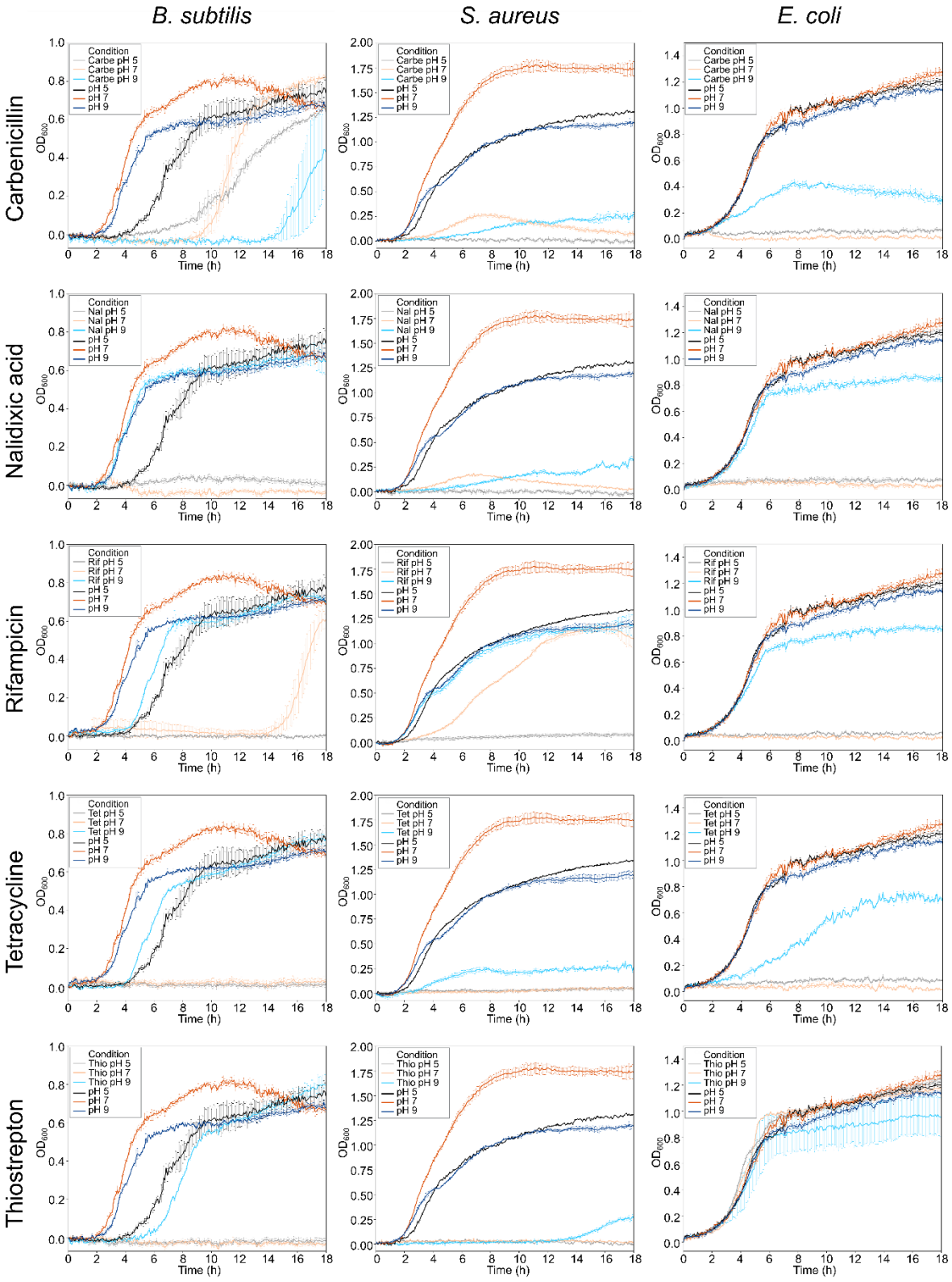

Figure 4.11. Growth curves of *B. subtilis*, *S. aureus* and *E. coli* in MH broth at pH 5 (black), pH 7 (orange), pH 9 (blue) and in MH broth containing carbenicillin, nalidixic acid, rifampicin, tetracycline and thiostrepton at pH 5 (grey), pH 7 (light orange), pH 9 (light blue). All the antibiotics were present in the culture media before inoculation. All the growth curves were carried out in triplicate.

Last, in order to investigate how extracellular pH affects antibiotic transport across the cell membrane, a LC-MS based assay was carried out based on previous work (Zhou *et al.*, 2015). Due to the COVID-19 lockdown, only a first preliminary test was carried out. First, calibrated standard curves were prepared for each commercial antibiotic (Figure S4.12-S4.16). However, concentrations would need to be adjusted, as some of the dilutions were too diluted for the antibiotic to be detected and thus some standard curves only have 2-4 points. Moreover, the serial dilutions of the standards needed to be injected three times each on different days to account for calibration errors. The antibiotic uptake assay was carried out just once in *E. coli* and the supernatants were analysed via LC-MS, however replicates could not be made and the results obtained indicate important handling errors since the controls should never have less antibiotic than the experimental samples (Figure 4.12B). Moreover, none of the samples nor controls should contain more antibiotic than the amount added of each (25 µg/ml nalidixic acid, 50 µg/ml carbenicillin 50 µg/ml thiostrepton, 50 µg/ml rifampicin and 10 µg/ml tetracycline).

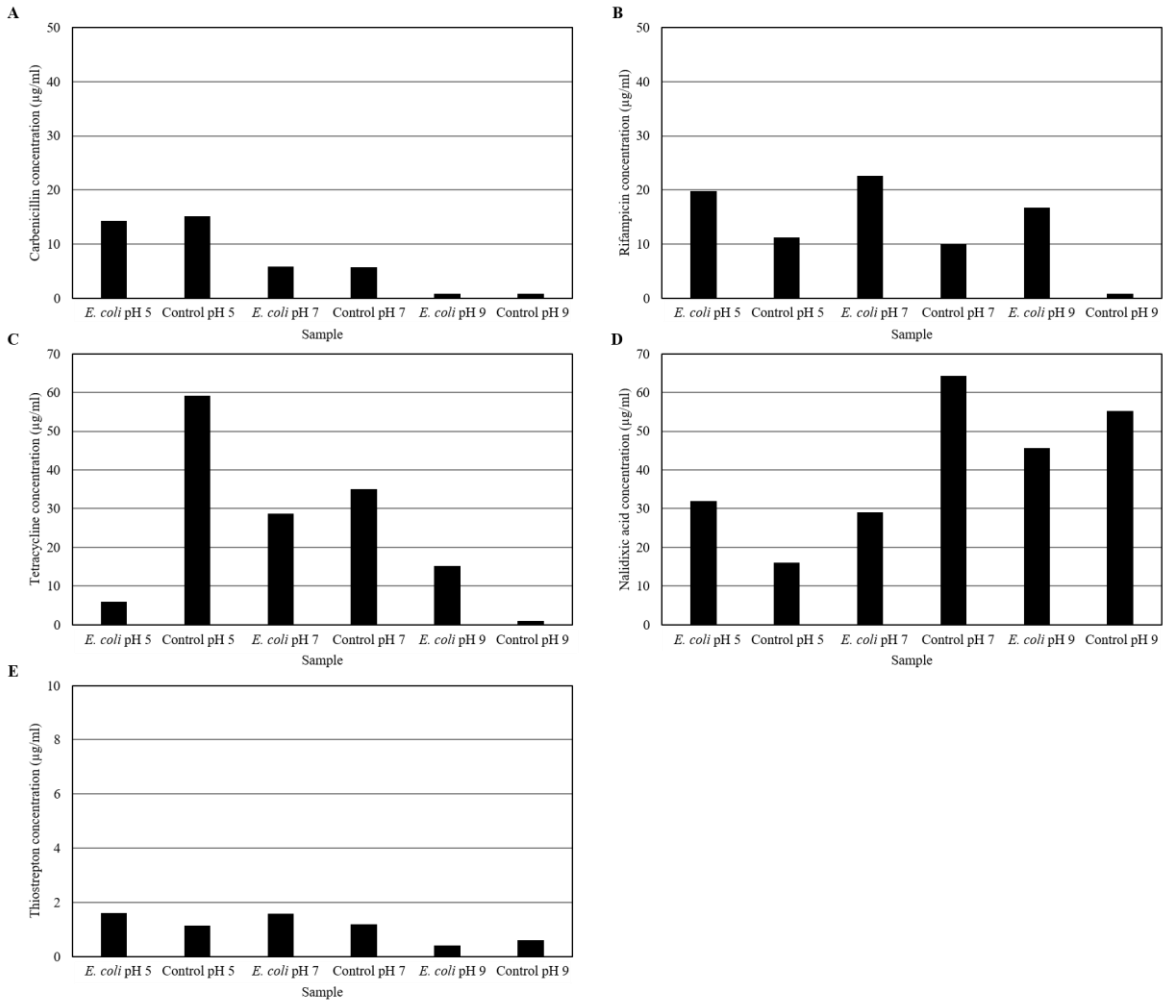

Figure 4.12. A) Carbenicillin, **B)** rifampicin, **C)** tetracycline, **D)** nalidixic acid and **E)** thiostrepton quantification by LC-MS after incubation with and without cells. The data shown is not based on replicates nor representative repeats.

Finally, in order to analyse if our observations on the sensitivity of antimicrobial diffusion assays could be practically applied to facilitate the discovery of non-abundant antimicrobial compounds, ten *Actinobacteria* grown in liquid (organic extracts) and in agar plates were screened. The plugs of strains 1, 3 and 4 generated the largest ZOI against *B. subtilis*, more specifically isolate 1 generated the largest ZOI when *B. subtilis* was grown in MH at pH 5, NA at pH 5 and CDM at pH 5, isolate 3 generated the largest ZOI when *B. subtilis* was grown in NA at pH 9, LB at pH 9 and CDM at pH 9 and isolate 4 generated the largest ZOI when *B. subtilis* was grown in NA at pH 9, CDM at pH 5 and CDM at pH 7 (Figure 4.13A). The fermentation extracts of strains 2, 4 and 9 generated the largest ZOI when *B. subtilis* was grown in MH at pH 5, NA at pH 5, LB at pH 5 and CDM at pH 5 (Figure 4.13B). The extracts from strains 8 and 10 showed the best results in CDM at pH 5 and MH at pH 5 respectively. When tested against *S. aureus*, the plugs of strains 1, 2, 3 and 4 showed the largest ZOI in LB at pH 9, followed by MH at pH 9 and NA at pH 9 for strain 3 (Figure 4.13C). The fermentation extract of strain 2 generated the largest ZOI against *S. aureus* when it was grown in MH at pH 5, NA at pH 5 and LB at pH 5 and 0.5 mm ZOI were generated by the extracts of strains 8 and 9 also in MH at pH 5, NA at pH 5 and LB at pH 5 (Figure 4.13D).

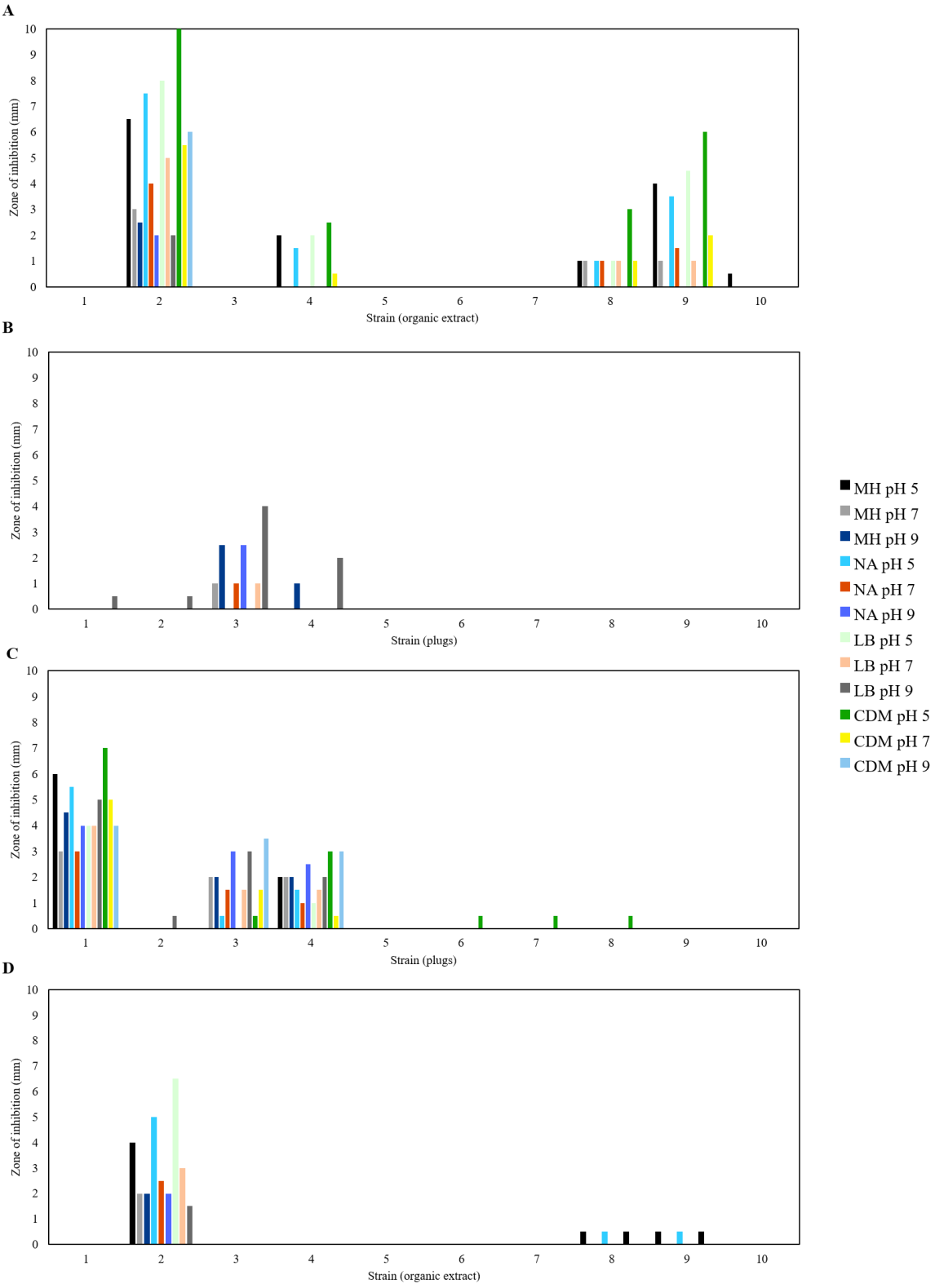

Figure 4.13. A) Zones of inhibition generated by plugs of different *Actinobacteria* isolates against *B. subtilis* grown in MH, LB, NA and CDM at pH 5, pH 7 and pH 9. **B)** Zones of inhibition generated by the organic extracts of the same isolates against *B. subtilis* grown in MH, LB, NA and CDM at pH 5, pH 7 and pH 9. **C)** Zones of inhibition generated by plugs of different *Actinobacteria* isolates against *S. aureus* grown in MH, LB and NA at pH 5, pH 7 and pH 9. **D)** Zones of inhibition generated by the organic extracts of the same isolates against *S. aureus* grown in MH, LB and NA at pH 5, pH 7 and pH 9. Due to time limitations this experiment was carried out with just one replicate for each sample.

4.4 Discussion.

4.4.1 Ribosome engineering to enhance abyssomicin production in M. maris *AB-18-032.*

Ribosome engineering is a methodology that has been used during the last 20 years in over 30 different bacterial and fungal species to activate or enhance the production of secondary metabolites (Ochi & Hosaka, 2013). It relies in the isolation of spontaneously developed drug-resistant mutants that often show a point mutation or a deletion mutation within a ribosomal component such as *rpsL* (ribosomal protein S12) or *rpoB* (RNA polymerase beta subunit) genes, indicating that bacterial gene expression can be altered dramatically by modulating ribosomal proteins and/or rRNA (Nishimura *et al.*, 2007; Okamoto *et al.*, 2007). In this work, the application of ribosome engineering to increase abyssomicin C production in *M. maris* AB-18-032 was presented, which is the first time this approach was applied to a *Micromonospora* species. Over 140 single and double drug-resistant mutants were isolated, and, from those, nine mutants generated ZOI larger than those produced by the wild-type strain (Figure 4.5). Ribosomal genes were sequenced depending on the antibiotic on which the mutants arose. Three streptomycin resistant mutants that displayed an amino acid substitution at the 88 position in RpsL were found, followed by two mutants that displayed three identical consecutive nucleotide substitutions that resulted in the amino acid substitution P91Y in RpsL, one mutant with the H419R substitution in RpoB and two double drug-resistant mutants displaying the amino acid substitutions H419Y and A424T in RpoB (Table 4.4). Interestingly, the point mutations in amino acid 88 of RpsL were previously described for different *Streptomyces* species and linked to transcriptional activation of secondary metabolite biosynthetic gene clusters and enhanced production (Wang *et al.*, 2014; Ochi, 2017). On the contrary, the P91Y substitution in RpsL has never been described before, suggesting it could confer species specific advantages in secondary metabolite production. Both substitutions in the amino acid 419 of RpoB (corresponding to the amino acid 437 in *S. coelicolor* A3(2)) were also found previously in *P. shermanii*, various *Streptomyces*, *A. orientalis* and *S. erythraea* (Ochi, 2017; Tanaka *et al.*, 2017), but there is not any report describing the amino acid substitution in the position 424 of RpoB. No mutations were found in *rpsL*, *rpoB* nor *gidB* genes for the chloramphenicol resistant mutants, suggesting spontaneous mutations could have taken place in a different component within the ribosomal machinery.

Following that, the maximum abyssomicin C production yield was calculated for each strain via analytical methods. This confirmed that four mutants were capable of producing up to 3.4-fold abyssomicin C in comparison to the wild-type strain (Figure 4.6). Interestingly, mm40 and

mm41 displayed a similar abyssomicin C yield, although they presented a different amino acid substitution in RpsL, suggesting both could be equally efficient, although mm42 and mm41 displayed the same point mutation but mm42 produced half the yield of mm41, suggesting there may be other mutations somewhere else in the genome also involved in the increased production or other mutations that decrease/supress the increased production (Table 4.4, Figure 4.6). Future work could focus on measuring strains growth so that abyssomicin C production could be normalised with biomass. Finally, the genomes of mm41 and its parent strain *M. maris* AB-18-032 were sequenced in order to identify potential mutations outside *rpsL*, *rpoB* and *gidB* that could also be responsible for the enhanced production. Among the mutations found there were five amino acid substitutions in a D-alanyl-D-alanine carboxypeptidase, one point mutation in the promoter sequence of the MurR/RpirR family transcriptional regulator of the *ngcEFG* operon involved in N-acetylglucosamine import and one point mutation in the intergenic region between AbyU and AbyI where potentially both terminators should be (Table 4.5). D-alanyl-D-alanine carboxypeptidases are proteins inhibited by ß-lactam antibiotics involved in peptidoglycan synthesis. Previous work demonstrated that these enzymes are involved in cell wall maturation, resistance and germination (Rioseras *et al.*, 2016), however, there are not studies relating them with ribosome engineering and the increased production of secondary metabolites. Similarly, there are not reports relating the *ngcEFG* operon to an increased antibiotic production, however, experiments with [^{14}C]-labelled N-acetylglucosamine could provide interesting information as how this substrate is used by *M. maris* (Xiao *et al.*, 2002; Iinuma *et al.*, 2018). It would be more logical to relate abyssomicin C yield to the mutation observed within the *aby* BGC. Although it is present in an intergenic region, the fact that it is located right in the terminator sequence of the main pathway activator draws someone's attention. In order to demonstrate if any of these mutations are responsible for the enhancement of abyssomicin C production, further work would be necessary and could focus on reintroducing each of these mutations in the wild-type *M. maris* AB-18-032. Moreover, other strategies to further enhance abyssomicin C production could be put into place in the future, including carrying out more rounds of ribosome engineering with other ribosome-targeting antibiotics and using the best ribosome engineering mutants for genome shuffling (Liu *et al.*, 2020).

4.4.2 Extracellular pH influences the antimicrobial detection limit in plate diffusion assays.

The abyssomicin C production yield of *M. maris* is relatively low and, initially, it was specially challenging detecting production via plug diffusion assays despite counting with a previously

developed diffusion assay specifically developed for the detection of antimicrobial compounds targeting the folate pathway (Riedlinger *et al.*, 2004). At the beginning of this project, when carrying out plate diffusion assays with PABA for the first time, PABA was added into the media right before pouring the media on plates, that is, after the media had been adjusted to pH 7.3 and autoclaved. Contrary to the reduced ZOI that were supposed to be observed in the presence of PABA, the ZOI were significatively larger, the presence of PABA was somehow making the antimicrobial diffusion assays more sensitive. Initially, it was thought that the concentration of PABA needed to be adjusted, as it could be being toxic for the test strain, however, PABA is a weak acid (pKa = 2.50 (carboxyl; H_2O); pKa = 4.87 (amino; H_2O)) (Cismesia *et al.*, 2017) and it likely modifies the pH of the media. With that in mind, a media screening in which four different media at pH 5, pH 7 and pH 9 with and without PABA were used to find the conditions that eased the detection of small amounts of abyssomicin C was carried out. The rescreen of *M. maris* AB-18-032 and its ribosome engineering mutants revealed that the plugs generated larger ZOI when *B. subtilis* was grown at pH 9 regardless of the media (Figure 4.7).

Then, in order to investigate whether this behaviour could also be observed when trying to detect other antibiotics, five commercial antibiotics with different structures and modes of action (Table S4.5) were used to carry out plate diffusion assays in MH, LB, NA and CDM at pH 5, pH 7 and pH 9, and the sensitivity of antimicrobial plate diffusion assays was assessed by measuring the detection limit for each antibiotic as the amount of antibiotic that generated a visible ZOI of 0.5 mm. The results showed that using media at pH 5 rather than at pH 7 to support the growth of the test strain could increase the sensitivity of the assay up to five times depending on the antibiotic and the test strain, whereas at pH 9, in most of the cases, concentrated stocks of the antibiotics were needed to be able to see that minimal ZOI (Figure 4.8 and Figure 4.9). This phenomenon had been observed before, however, it had never been applied as a potential method to facilitate antimicrobial drug discovery. Already in the 1940s, several studies showed that the pH of the medium affects significatively the activity of certain antibiotics against both Gram-positive and Gram-negative bacteria (Abraham & Duthie, 1946; Pratt & Dufrenoy, 1948; Chang & Stafseth, 1950). During decades, this was considered a hot research topic, as the enhanced antimicrobial activity could be useful in the treatment of urinary tract infections. In the 1970s, researchers reported that acidic antibiotics like penicillin G, ampicillin, cephalothin, cephaloridine, and novobiocin were considerably more active at pH 5.0, whereas the activity of basic antibiotics such as erythromycin and gentamicin was higher at pH 8.5 (Toala *et al.*, 1970; Miller & Perkins, 1973). Rifampin was found maximum stable at

pH 4 (Jindal *et al.*, 1995), whereas norfloxacin, a nalidixic acid analogue, was found to bind to DNA more efficiently at a pH between 4.5 and 6.5 (Shen & Pernet, 1985; Baudoux *et al.*, 2006).

At this point, three main hypotheses were considered for the research question "Why does the sensitivity of antimicrobial plate diffusion assays varies with extracellular pH?". The first hypothesis was that the extracellular pH could modify the pH in the cytosol affecting protein-ligand interaction, however, there are plenty of studies on the maintenance of pH homeostasis and for example, *E. coli* responds very efficiently to external pH variations, fully recovering its intracellular pH (7.2 to 7.8) within 5 minutes after acidification (Wilks & Slonczewski, 2007; Martinez *et al.*, 2012). Cytoplasmic pH homeostasis has also been comprehensively studied in *B. subtilis*, which maintains cytoplasmic pH within approximately pH 7.3 to pH 7.6 during vegetative growth over a range of environmental pH (Cotter & Hill, 2003; Padan *et al.*, 2005; Wilks *et al.*, 2009). Indeed, in the growth curves carried out to study how each media supports the growth of the antimicrobial test strains, it was possible to observe that there were no differences in the growth of *E. coli* regardless of the pH of the media (Figure 4.10C).

The second hypothesis was related to how pH modifies the density of the agar and thus the diffusion of the antibiotics through it. In order to study if the sensitivity of the antimicrobial disk assays could be related to differential diffusion rates through the agar due to the change in density of the agar at each different pH, growth curves of *B. subtilis*, *S. aureus* and *E. coli* were carried out in MH at pH 5, pH 7 and pH 9 containing the five commercial antibiotics used before. During this experiment, growth of all the strains was observed under the presence of all antibiotics when the media was adjusted to pH 9 whereas at pH 5 and pH 7 bacterial growth was inhibited in most cases (Figure 4.11). These results follow the trends observed when carrying out the antimicrobial plate diffusion assays (Figure 4.8 and Figure 4.9). Both experiments suggest that the media pH plays a key role in the detection of antimicrobial compounds and that cells are more sensitive at low pH, at least for the antibiotics tested. Moreover, the fact that these observations are similar in liquid and solid media discards the hypothesis that antimicrobial plate diffusion assays at pH 5 were more sensitive because the low pH decreased the agar density facilitating the diffusion of the compounds through the media.

Finally, the third hypothesis was that the extracellular pH affected antibiotic transport across the cell membrane, facilitating the entrance of small molecules depending on their charge. For example, previous work identified that pyrazinamide activity is linked to the acidity of the

medium, being most active at pH 5.5 and almost inactive at pH 7 (Zhang *et al.*, 2002a). In this case, it was demonstrated that acid pH enhanced the uptake and accumulation of the active derivative pyrazinoic acid, as weak acids spontaneously diffuse through the membrane in their protonated forms (Nikaido & Thanassi, 1993; Zhang *et al.*, 1999). However, transport across the cell membrane is complex. Passage of drugs across the outer membrane of Gram-negative bacteria can occur by passive diffusion through porin channels for most hydrophilic and amphiphilic antibiotics (e.g. beta-lactams and tetracyclines), although it is influenced by other factors such as drug hydrophobicity, size and net charge, and increased hydrophobicity, size or net negative charge tend to decrease the rate of permeation through porin channels (Chopra, 1988). Larger antibiotics such as erythromycin or novobiocin, are believed to cross the outer membrane by diffusion through the lipid bilayer, which is more slow (Zgurskaya *et al.*, 2016) and other transport options include facilitated diffusion using specific carriers (e.g. albomycin), or by self-promoted uptake (e.g. aminoglycosides and polymyxins) (Chopra, 1988). Then, the transfer of antibiotics across the bacterial cytoplasmic membrane is usually mediated by active, carrier-mediated, transport systems normally operating to transport essential solutes into the cell (Chopra, 1988). Interestingly, there is not a simple explanation for how small molecules enter bacterial cells. For example, research on tetracycline transport across the cell membrane suggests that tetracycline may cross the outer membrane as a cationic chelate of magnesium via OmpF channels (Chopra & Ball, 1982; Hancock, 1987; Chopra & Roberts, 2001). Then, passage of tetracycline across the cytoplasmic membrane occurs both by energy independent (passive diffusion) and energy-dependent (active) processes (Chopra, 1988) via two active carrier-mediated transport systems, one that is ATP dependent and a second one that is coupled to the proton motive force (McMurry & Levy, 1978; Argast & Beck, 1984; Chopra *et al.*, 1992).

Recent work in this field has focused on developing assays that allow to measure permeation and accumulation of small molecules inside the cells, taking into account variables such as the presence of multiple porins, efflux pumps and varying lipopolysaccharides (Richter *et al.*, 2017; Gayen *et al.*, 2019). In order to test our third hypothesis, a previously developed protocol to measure intracellular drug penetration based on LC-MS detection for the assessment of antibiotic uptake under different external pH was adapted (Zhou *et al.*, 2015). Due to time limitations, only a preliminary test was carried out. Future efforts will focus on repeating this work and optimising the experimental pipeline to reduce error. Once results are obtained for *E. coli*, in order to identify differences between different bacteria, the plan would be to repeat everything with *B. subtilis* and *S. aureus*. Through this approach it should be possible to obtain data to confirm the relevance of extracellular pH on the uptake of small molecules.

Finally, in order to assess whether it would be possible to take advantage of the knowledge acquired regarding how extracellular pH affects cells sensitivity to antimicrobial compounds to facilitate the discovery of non-abundant antimicrobial compounds, the ZOI generated by plugs and extracts from ten *Actinobacteria* against *B. subtilis* and *S. aureus* grown in various media at pH 5, pH 7 and pH 9 were analysed. Generally, it was possible to observe that the plugs produced largest ZOI at pH 9, whereas all the extracts generated largest ZOI when the test strains were grown at pH 5 (Figure 4.13). These results coincide with the plug assay results obtained when *M. maris* and its ribosome engineering mutants were screened (Figure 4.7), as well as the results obtained when using the purified stocks of commercial antibiotics (Figure 4.8, Figure 4.9 and Figure 4.11), suggesting that plate diffusion screenings at pH 5 could be better to detect secondary metabolites produced in small amounts from plugs, whilst plate diffusion screenings at pH 9 could be a better option when using fermentation extracts for the screening. Future work could focus on screening a representative number of commercial antibiotics as well as strains in order to further understand if these results are also subjected to the chemistry or the mode of action of the compound. It would also be interesting to understand why plugs and extracts display opposite trends and what kind of fluxes and gradients originate when the plugs enter in contact with the test media.

4.5 Conclusions and future perspectives.

The aim of this work was to enhance the yield of abyssomicin C produced by the slow growing marine strain *M. maris* AB-18-032. Through ribosome engineering various mutants with up to 3.4-fold increase in abyssomicin C production in comparison to the wild-type strain were obtained. All the mutants presented mutations in ribosomal genes. The genome of the best producer strain was sequenced and several mutations that could also be involved in the increase in abyssomicin C production were found. Future work could focus on using statistical design of experiments (DOE) to reintroduce the different mutations found in ribosome engineering mutants with enhanced abyssomicin C production into the wild-type *M. maris* AB-18-032, for example via CRISPR, individually or in different combinations, to determine which ones are responsible of the enhanced production. To further increase abyssomicin C production other approaches such as genome shuffling could be applied.

In addition to that, in this chapter, the role media pH plays in antimicrobial plate diffusion assays was also analysed. The initial intention was to find the best plate diffusion assay conditions for the detection of small amounts of abyssomicin C. This assay showed that the largest ZOI were produced using media at pH 9, thus, further investigations were carried out. The results obtained suggest that media at pH 5 could favour the observation of antimicrobial activity from plugs, whereas using media at pH 9 could aid when analysing fermentation extracts for antimicrobial activity. This could be due to the influence of extracellular pH on the transport of metabolites across the cell membrane. Future work could test this last hypothesis and investigate whether these assays could be broadly applied for the discovery of antimicrobial compounds produced in low amounts.

4.6 References.

Abraham, E. P. & Duthie, E. S. (1946). Effect of pH of the medium on activity of streptomycin and penicillin and other chemotherapeutic substances. *Lancet* **247**, 455–459.

Ahmetagic, A. & Pemberton, J. M. (2011). Antibiotic resistant mutants of *Escherichia coli* K12 show increases in heterologous gene expression. *Plasmid* **65**, 51–57.

Argast, M. & Beck, C. F. (1984). Tetracycline diffusion through phospholipid bilayers and binding to phospholipids. *Antimicrob Agents Chemother* **26**, 263–265.

Artsimovitch, I., Patlan, V., Sekine, S. I., Vassylyeva, M. N., Hosaka, T., Ochi, K., Yokoyama, S. & Vassylyev, D. G. (2004). Structural basis for transcription regulation by alarmone ppGpp. *Cell* **117**, 299–310.

Balouiri, M., Sadiki, M. & Ibnsouda, S. K. (2016). Methods for *in vitro* evaluating antimicrobial activity: A review. *J Pharm Anal* **6**, 71–79.

Baudoux, P., Bles, N., Lemaire, S., Mingeot-Leclercq, M.-P., Tulkens, P. M. & Van Bambeke, F. (2006). Combined effect of pH and concentration on the activities of gentamicin and oxacillin against *Staphylococcus aureus* in pharmacodynamic models of extracellular and intracellular infections. *J Antimicrob Chemother* **59**, 246–253.

Bentley, S. D., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Harper, D., Bateman, A., Brown, S., Collins, M. & other authors. **(2002).** Complete genome sequence of the model actinomycete *Streptomyces*. *Nature* **417**, 141–147.

Bihelovic, F. & Saicic, R. N. (2012). Total synthesis of (-)-atrop-abyssomicin C. *Angew Chemie - Int Ed* **51**, 5687–5691.

Bose, J. L. (2016). Chemical and UV mutagenesis. *Methods Mol Biol*, 111–115.

Chait, R., Shrestha, S., Shah, A. K., Michel, J.-B. & Kishony, R. (2010). A differential drug screen for compounds that select against antibiotic resistance. *PLoS One* **5**, e15179.

Chang, K.-H. & Stafseth, H. J. (1950). Influence of various factors on the bacteriostatic and bactericidal action of streptomycin on *Salmonella pullorum*. *Poult Sci* **29**, 130–138.

Chen, L., Shang, G., Yuan, W., Wu, Y. & Bai, F. (2012). Screening of *Clostridium* strains through ribosome engineering for improved butanol production. *Sheng Wu Gong Cheng Xue Bao* **28**, 1048–58.

Chopra, I. (1988). Molecular mechanisms involved in the transport of antibiotics into bacteria. *Parasitology* **96**, S25–S44.

Chopra, I., Hawkey, P. M. & Hinton, M. (1992). Tetracyclines, molecular and clinical aspects. *J Antimicrob Chemother* **29**, 245–277.

Chopra, I. & Ball, P. (1982). Transport of antibiotics into bacteria. *Adv Microb Physiol* **23**, 183–240.

Chopra, I. & Roberts, M. (2001). Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* **65**, 232–260.

Cismesia, A. P., Nicholls, G. R. & Polfer, N. C. (2017). Amine vs. carboxylic acid protonation in ortho-, meta-, and para-aminobenzoic acid: An IRMPD spectroscopy study. *J Mol Spectrosc* **332**, 79–85.

CLSI. (2018). *Performance Standards for Antimicrobial Disk Susceptibility Tests*, 13th Edition. (P. C. and L. S. I. Wayne, Ed.).

Cotter, P. D. & Hill, C. (2003). Surviving the acid test: Responses of Gram-positive bacteria to low pH. *Microbiol Mol Biol Rev* **67**, 429–453.

Freundlich, J. S., Lalgondar, M., Wei, J.-R., Swanson, S., Sorensen, E. J., Rubin, E. J. & Sacchettini, J. C. (2010). The abyssomicin C family as *in vitro* inhibitors of *Mycobacterium tuberculosis*. *Tuberculosis* **90**, 298–300.

Gao, X., Zhao, H., Zhang, G., He, K. & Jin, Y. (2012). Genome shuffling of *Clostridium acetobutylicum* CICC 8012 for improved production of acetone-butanol-ethanol (ABE). *Curr Microbiol* **65**, 128–132.

Gayen, A., Kumar, D., Matheshwaran, S. & Chandra, M. (2019). Unveiling the modulating role of extracellular pH in permeation and accumulation of small molecules in subcellular compartments of Gram-negative *Escherichia coli* using nonlinear spectroscopy. *Anal Chem* **91**, 7662–7671.

Gomez-Escribano, J. P. & Bibb, M. J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **4**, 207–215.

Gong, J., Zheng, H., Wu, Z., Chen, T. & Zhao, X. (2009). Genome shuffling: Progress and applications for phenotype improvement. *Biotechnol Adv* **6**, 996-1005.

Gottardi, E. M., Krawczyk, J. M., Von Suchodoletz, H., Schadt, S., Mühlenweg, A., Uguru, G. C., Pelzer, S., Fiedler, H. P., Bibb, M. J. & other authors. (2011). Abyssomicin biosynthesis: Formation of an unusual polyketide, antibiotic-feeding studies and genetic analysis. *ChemBioChem* **12**, 1401–1410.

Hancock, R. E. W. (1987). Role of porins in outer membrane permeability. *J Bacteriol* **3**, 929–933.

Hosaka, T., Xu, J. & Ochi, K. (2006). Increased expression of ribosome recycling factor is responsible for the enhanced protein synthesis during the late growth phase in an antibiotic-overproducing *Streptomyces coelicolor* ribosomal *rpsL* mutant. *Mol Microbiol* **61**, 883–897.

Hosaka, T., Ohnishi-Kameyama, M., Muramatsu, H., Murakami, K., Tsurumi, Y., Kodani, S., Yoshida, M., Fujie, A. & Ochi, K. (2009). Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat Biotechnol* **27**, 462–464.

Hutter, B., Fischer, C., Jacobi, A., Schaab, C. & Loferer, H. (2004). Panel of *Bacillus subtilis* reporter strains indicative of various modes of action. *Antimicrob Agents Chemother* **48**, 2588–2594.

Iinuma, C., Saito, A., Ohnuma, T., Tenconi, E., Rosu, A., Colson, S., Mizutani, Y., Liu, F., Świątek-Połatyńska, M. & other authors. **(2018).** NgcE^{Sco} acts as a lower-affinity binding protein of an ABC transporter for the uptake of N,N′-diacetylchitobiose in *Streptomyces coelicolor* A3(2). *Microbes Environ* **33**, 272–281.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**, 526–531.

Jeschek, M., Gerngross, D. & Panke, S. (2017). Combinatorial pathway optimization for streamlined metabolic engineering. *Curr Opin Biotechnol* **47**, 142-151.

Jindal, K. C., Chaudhary, R. S., Singla, A. K., Gangwal, S. S. & Khanna, S. (1995). Effects of buffers and pH on rifampicin stability. *Pharm Ind* **57**, 420–422.

Jorgensen, J. H. & Ferraro, M. J. (2009). Antimicrobial susceptibility testing: A review of general principles and contemporary practices. *Clin Infect Dis* **49**, 1749–1755.

Keller, S., Nicholson, G., Drahl, C., Sorensen, E., Fiedler, H.-P. & Süssmuth, R. D. (2007a). Abyssomicins G and H and atrop-abyssomicin C from the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **60**, 391–394.

Keller, S., Schadt, H. S., Ortel, I. & Süssmuth, R. D. (2007b). Action of atrop-abyssomicin C as an inhibitor of 4-amino-4-deoxychorismate synthase PabB. *Angew Chemie - Int Ed* **46**, 8284–8286.

Khan, Z. A., Siddiqui, M. F. & Park, S. (2019). Current and emerging methods of antibiotic susceptibility testing. *Diagnostics* 2, 49.

Kubo, Y., Inaoka, T., Hachiya, T., Miyake, M., Hase, S., Nakagawa, R., Hasegawa, H., Funane, K., Sakakibara, Y. & Kimura, K. (2013). Development of a rifampicin-resistant *Bacillus subtilis* strain for natto-fermentation showing enhanced exoenzyme production. *J Biosci Bioeng* **115**, 654–657.

Lai, C., Xu, J., Tozawa, Y., Okamoto-Hosoya, Y., Yao, X. & Ochi, K. (2002). Genetic and physiological characterization of *rpoB* mutations that activate antibiotic production in *Streptomyces lividans*. *Microbiology* **148**, 3365–3373.

Leeds, J. A., Schmitt, E. K. & Krastel, P. (2006). Recent developments in antibacterial drug discovery: Microbe-derived natural products - From collection to the clinic. *Expert Opin Investig Drugs* **3**, 211-226.

Liu, H., Jiang, C., Lin, J., Zhuang, Z., Kong, W., Liu, L., Huang, Y., Duan, Y. & Zhu, X. (2020). Genome shuffling based on different types of ribosome engineering mutants for enhanced production of 10-membered enediyne tiancimycin-A. *Appl Microbiol Biotechnol* **10**, 4359-4369.

Liu, Z., Zhao, X. & Bai, F. (2013). Production of xylanase by an alkaline-tolerant marine-derived *Streptomyces viridochromogenes* strain and improvement by ribosome engineering. *Appl Microbiol Biotechnol* **97**, 4361–4368.

Magocha, T. A., Zabed, H., Yang, M., Yun, J., Zhang, H. & Qi, X. (2018). Improvement of industrially important microbial strains by genome shuffling: Current status and future prospects. *Bioresour Technol* **257**, 281-289.

Marchiaro, P., Ballerini, V., Spalding, T., Cera, G., Mussi, M. A., Morán-Barrio, J., Vila, A. J., Viale, A. M. & Limansky, A. S. (2008). A convenient microbiological assay employing cell-free extracts for the rapid characterization of Gram-negative carbapenemase producers. *J Antimicrob Chemother* **62**, 336–344.

Martinez, K. A., Kitko, R. D., Mershon, J. P., Adcox, H. E., Malek, K. A., Berkmen, M. B. & Slonczewski, J. L. (2012). Cytoplasmic pH response to acid stress in individual cells of *Escherichia coli* and *Bacillus subtilis* observed by fluorescence ratio imaging microscopy. *Appl Environ Microbiol* **78**, 3706–3714.

Mazák, K. & Noszál, B. (2016). Advances in microspeciation of drugs and biomolecules: Species-specific concentrations, acid-base properties and related parameters. *J Pharm Biomed Anal* **130**, 390–403.

McMurry, L. & Levy, S. B. (1978). Two transport systems for tetracycline in sensitive *Escherichia coli*: critical role for an initial rapid uptake system insensitive to energy inhibitors. *Antimicrob Agents Chemother* **14**, 201–209.

Miller, M. A. & Perkins, R. L. (1973). Effect of pH on *in vitro* activity of carbenicillin against *Proteus mirabilis*. *J Infect Dis* **127**, 689–693.

Nicolaou, K. C., Harrison, S. T. & Chen, J. S. (2009). Discoveries from the abyss: The abyssomicins and their total synthesis. *Synthesis (Stuttg)* **1**, 33–42.

Nikaido, H. & Thanassi, D. G. (1993). Penetration of lipophilic agents with multiple protonation sites into bacterial cells: Tetracyclines and fluoroquinolones as examples. *Antimicrob Agents Chemother* **7**, 1393–1399.

Nishimura, K., Hosaka, T., Tokuyama, S., Okamoto, S. & Ochi, K. (2007). Mutations in *rsmG*, encoding a 16S rRNA methyltransferase, result in low-level streptomycin resistance and antibiotic overproduction in *Streptomyces coelicolor* A3(2). *J Bacteriol* **189**, 3876–3883.

Ochi, K. (2007). From microbial differentiation to ribosome engineering. *Biosci Biotechnol Biochem* **71**, 1373–1386.

Ochi, K. (2017). Insights into microbial cryptic gene activation and strain improvement: Principle, application and technical aspects. *J Antibiot (Tokyo)* **1**, 25-40.

Ochi, K. & Hosaka, T. (2013). New strategies for drug discovery: Activation of silent or weakly expressed microbial gene clusters. *Appl Microbiol Biotechnol* **97**, 87–98.

Ochi, K., Okamoto, S., Tozawa, Y., Inaoka, T., Hosaka, T., Xu, J. & Kurosawa, K. (2004). Ribosome engineering and secondary metabolite production. *Adv Appl Microbiol* **56**, 155–184.

Okamoto-Hosoya, Y., Hosaka, T. & Ochi, K. (2003). An aberrant protein synthesis activity is linked with antibiotic overproduction in *rpsL* mutants of *Streptomyces coelicolor* A3(2). *Microbiology* **149**, 3299–3309.

Okamoto, S., Tamaru, A., Nakajima, C., Nishimura, K., Tanaka, Y., Tokuyama, S., Suzuki, Y. & Ochi, K. (2007). Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. *Mol Microbiol* **63**, 1096–1106.

Padan, E., Bibi, E., Ito, M. & Krulwich, T. A. (2005). Alkaline pH homeostasis in bacteria: New insights. *Biochim Biophys Acta - Biomembr* **2**, 67-88.

Pratt, R. & Dufrenoy, J. (1948). Cytochemical interpretation of the mechanism of penicillin action. *Bacteriol Rev* **12**, 79–103.

Reen, F. J., Romano, S., Dobson, A. D. W. & O'Gara, F. (2015). The sound of silence: Activating silent biosynthetic gene clusters in marine microorganisms. *Mar Drugs* **8**, 4754–4783.

Richter, M. F., Drown, B. S., Riley, A. P., Garcia, A., Shirai, T., Svec, R. L. & Hergenrother, P. J. (2017). Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature* **545**, 299–304.

Riedlinger, J., Reicke, A., Zähner, H., Krismer, B., Bull, A. T., Maldonado, L. A., Ward, A. C., Goodfellow, M., Bister, B. & other authors. **(2004).** Abyssomicins, inhibitors of the para-aminobenzoic acid pathway produced by the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **57**, 271–279.

Rioseras, B., Yaguë, P., López-Garciá, M. T., Gonzalez-Quinõnez, N., Binda, E., Marinelli, F. & Manteca, A. (2016). Characterization of SCO4439, a D-alanyl-D-alanine carboxypeptidase involved in spore cell wall maturation, resistance, and germination in *Streptomyces coelicolor*. *Sci Rep* **6**, 1–15.

Sadaka, C., Ellsworth, E., Robert Hansen, P., Ewin, R., Damborg, P. & Watts, J. L. (2018). Review on abyssomicins: Inhibitors of the chorismate pathway and folate biosynthesis. *Molecules* **23**.

Saicic, R. N., Matovic, R., Bihelovic, F. & Gruden-Pavlovic, M. (2014). Total synthesis and biological evaluation of atrop-O-benzyl-desmethylabyssomicin C. *Org Biomol Chem* **11**.

Saito, N., Xu, J., Hosaka, T., Okamoto, S., Aoki, H., Bibb, M. J. & Ochi, K. (2006). EshA accentuates ppGpp accumulation and is conditionally required for antibiotic production in *Streptomyces coelicolor* A3(2). *J Bacteriol* **188**, 4952–4961.

Shen, L. L. & Pernet, A. G. (1985). Mechanism of inhibition of DNA gyrase by analogues of nalidixic acid: The target of the drugs is DNA. *Proc Nati Acad Sci USA* **82**, 307–311.

Shima, J., Hesketh, A., Okamoto, S., Kawamoto, S. & Ochi, K. (1996). Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* **178**, 7276–7284.

Singh, S. B., Young, K. & Miesel, L. (2011). Screening strategies for discovery of antibacterial natural products. *Expert Rev Anti Infect Ther* **8**, 589-613.

Suzuki, T., Seta, K., Nishikawa, C., Hara, E., Shigeno, T. & Nakajima-Kambe, T. (2015). Improved ethanol tolerance and ethanol production from glycerol in a streptomycin-resistant *Klebsiella variicola* mutant obtained by ribosome engineering. *Bioresour Technol* **176**, 156–162.

Tanaka, Y., Kasahara, K., Izawa, M. & Ochi, K. (2017). Applicability of ribosome engineering to vitamin B12 production by *Propionibacterium shermanii*. *Biosci Biotechnol Biochem* **81**, 1636–1641.

Toala, P., Wilcox, C. & Finland, M. (1970). Effect of pH of medium and size of inoculum on activity of antibiotics against group D *Streptococcus* (*Enterococcus*). *Appl Microbiol* **4**, 629–637.

Wang, G., Hosaka, T. & Ochi, K. (2008). Dramatic activation of antibiotic production in *Streptomyces coelicolor* by cumulative drug resistance mutations. *Appl Environ Microbiol* **74**, 2834–2840.

Wang, Q., Zhang, D., Li, Y., Zhang, F., Wang, C. & Liang, X. (2014). Genome shuffling and ribosome engineering of *Streptomyces actuosus* for high-yield nosiheptide production. *Appl Biochem Biotechnol* **173**, 1553–1563.

Wilks, J. C. & Slonczewski, J. L. (2007). pH of the cytoplasm and periplasm of *Escherichia coli*: Rapid measurement by green fluorescent protein fluorimetry. *J Bacteriol* **189**, 5601–5607.

Wilks, J. C., Kitko, R. D., Cleeton, S. H., Lee, G. E., Ugwu, C. S., Jones, B. D., Bondurant, S. S. & Slonczewski, J. L. (2009). Acid and base stress and transcriptomic responses in *Bacillus subtilis*. *Appl Environ Microbiol* **75**, 981–990.

Xiao, X., Wang, F., Saito, A., Majka, J., Schlosser, A. & Schrempf, H. (2002). The novel *Streptomyces olivaceoviridis* ABC transporter Ngc mediates uptake of N-acetylglucosamine and N, N′-diacetylchitobiose. *Mol Genet Genomics* **267**, 429–439.

Xu, J., Tozawa, Y., Lai, C., Hayashi, H. & Ochi, K. (2002). A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Genet Genomics* **268**, 179–189.

Xu, M. & Wright, G. D. (2019). Heterologous expression-facilitated natural products' discovery in actinomycetes. *J Ind Microbiol Biotechnol* **46**, 415–431.

Zgurskaya, H. I., López, C. A. & Gnanakaran, S. (2016). Permeability barrier of Gram-negative cell envelopes and approaches to bypass it. *ACS Infect Dis* **1**, 512–522.

Zhang, Y., Scorpio, A., Nikaido, H. & Sun, Z. (1999). Role of acid pH and deficient efflux of pyrazinoic acid in unique susceptibility of *Mycobacterium tuberculosis* to pyrazinamide. *J Bacteriol* **181**, 2044–2049.

Zhang, Y., Permar, S. & Sun, Z. (2002a). Conditions that may affect the results of susceptibility testing of *Mycobacterium tuberculosis* to pyrazinamide. *J Med Microbiol* **51**, 42–49.

Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P. C. & Del Cardayré, S. B. (2002b). Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* **415**, 644–646.

Zhao, X. Q., Gust, B. & Heide, L. (2010). S-Adenosylmethionine (SAM) and antibiotic biosynthesis: Effect of external addition of SAM and of overexpression of SAM biosynthesis genes on novobiocin production in *Streptomyces*. *Arch Microbiol* **192**, 289–297.

Zhou, Y., Joubran, C., Miller-Vedam, L., Isabella, V., Nayar, A., Tentarelli, S. & Miller, A. (2015). Thinking outside the 'bug': A unique assay to measure intracellular drug penetration in Gram-negative bacteria. *Anal Chem* **87**, 3579–3584.

Zhu, S., Duan, Y. & Huang, Y. (2019). The application of ribosome engineering to natural product discovery and yield improvement in *Streptomyces*. *Antibiotics* **3**,133.

Zhuang, Z., Jiang, C., Zhang, F., Huang, R., Yi, L., Huang, Y., Yan, X., Duan, Y. & Zhu, X. (2019). Streptomycin-induced ribosome engineering complemented with fermentation optimization for enhanced production of 10-membered enediynes tiancimycin-A and tiancimycin-D. *Biotechnol Bioeng* **116**, 1304–1314.

Chapter 5. Genetic manipulation of *M. maris*: a DOE-guided electroporation protocol and CRISPR-mediated genome editing.

Abstract

Members of the genus *Micromonospora* are well-known for their ability to produce secondary metabolites with biomedical applications and for playing a very important role in association with higher organisms such as plants. Although their genetics has been studied for decades, most *Micromonospora* species remain recalcitrant to bioengineering techniques due to the lack of efficiency of basic molecular microbiology methods. In Chapter 4, ribosome engineering was used to generate *M. maris* mutants able to produce more abyssomicin C than the wild-type strain. Here, a statistical design of experiments approach was followed to establish an electroporation method for this strain and understand how the different factors involved in the process of DNA electroporation affect the obtention of *M. maris* electrotransformants. For that, two sets of 18 different electroporation experiments were carried out varying the values of the factors determined to have a significant impact on performance and a specific electroporation condition that yielded over 100 transformants per µg of DNA was found. This is a promising starting point and future work could look at applying a similar approach to electroporate other *Micromonospora* species. Finally, in an attempt to optimise a genome editing tool that allows quick and reliable engineering in *M. maris*, some preliminary work on the application of CRISPR-Cas9 systems for genome editing *M. maris* was also carried out by building a library of CRISPR vectors based on the pCRISPomyces-2 system.

5.1 Introduction.

5.1.1 Biotechnological potential of the genus Micromonospora.

Micromonospora is a genus of bacteria of the phylum *Actinobacteria* that is comprised of 103 species validly described as of the time of writing this manuscript (http://www.bacterio.net/micromonospora.html) (Parte, 2014) . They are Gram-positive, spore-forming, aerobic to microaerophilic, chemoorganotrophic and form branched mycelium (Barka *et al.*, 2016). Despite not being a dominant genus within environmental microbiomes, they are ubiquitously distributed in nature, inhabiting mainly terrestrial ecosystems but also marine and fresh-water environments (Supong *et al.*, 2013). More recently, they were also found associated to higher organisms in actinorhizal nodules (Carro *et al.*, 2013); living in deep association with plants having evolved from a purely terrestrial to a facultative endophytic lifestyle (Barka *et*

al., 2016), and associated to marine sponges (Contreras-Castro *et al.*, 2019) and molluscs (Manivasagan *et al.*, 2014). It has been demonstrated on several occasions that *Micromonospora* species can also inhabit extreme environments such as the Antarctic (Hirsch *et al.*, 2004), Mongolian deserts (Norovsuren *et al.*, 2007), saline soils from arid territories (Lubsanova *et al.*, 2014) and hyper-arid Atacama Desert soils (Carro *et al.*, 2019).

The genus *Micromonospora* is known for being the second most prolific source of antimicrobial compounds after the genus *Streptomyces* (Barka *et al.*, 2016; Carro *et al.*, 2019). However, it was not until the discovery of gentamicin, produced by *M. echinospora*, that the genus started attracting the interest of natural products researchers (Weinstein *et al.*, 1963; Ganguly, 1982). *Micromonospora* species are producers of the aminoglycosides gentamicin (Weinstein *et al.*, 1963), sagamicin (Okachi *et al.*, 1974), sisomicin (Weinstein *et al.*, 1970) and verdamicin (Weinstein *et al.*, 1975); the oligosaccharide antibiotic everninomicin (Weinstein *et al.*, 1964); the ansamycin antibiotic halomicin (Weinstein *et al.*, 1967) and the macrolide antibiotic and antiviral megalomicin (Weinstein *et al.*, 1969) and mycinamicin (Satoi *et al.*, 1980). With hundreds of antimicrobial compounds isolated from diverse *Micromonospora* species, bioinformatic methods have demonstrated that the hidden potential is still greater (Tiwari & Gupta, 2012; Carro *et al.*, 2018). Indeed, several clinically relevant *Micromonospora* natural products, such as ECO-04601, were uncovered through genome mining (Wilkinson & Micklefield, 2007). In addition to antimicrobial agents, *Micromonospora* species synthesise other secondary metabolites with antitumoral bioactivity such as calicheamicin (Love *et al.*, 1992), quinolidomicin (Yoichi *et al.*, 1993), macquarimicin (Jackson *et al.*, 1995), pyrrolosporin A (Lam *et al.*, 1996) and lupinacidin C (Igarashi *et al.*, 2011) amongst others, the antifungal spartanamicin (Pandey *et al.*, 1992) and the anti-inflammatory compounds diazepinomicin (Charan *et al.*, 2004) and BU-4664L (Igarashi *et al.*, 2005; Boumehira *et al.*, 2016).

The genus *Micromonospora* is not just of interest due to its potential as source of bioactive compounds. During the last two decades, many studies have focused on understanding plant-*Micromonospora* relationships, as strains from this genus have been isolated from roots of rice (Tian *et al.*, 2007) and wheat (Coombs & Franco, 2003) and from root nodules of the Australian pine tree (Valdés *et al.*, 2005), *Coriaria myrtifolia* (Trujillo *et al.*, 2006), *Lupinus angustifolius* (Trujillo *et al.*, 2007) and *Pisum sativum* (Garcia *et al.*, 2010; Hirsch & Valdés, 2010; Carro *et al.*, 2016). More specifically, *Micromonospora* species have been studied for their ability to influence plant growth and development. It has been demonstrated that some *Micromonospora*

strains act as Rhizobia Helper Bacteria (RHB) promoting symbiosis between *Discaria trinervis-Frankia* (Solans, 2007) and *Medicago sativa-Sinorhizobium meliloti* (Solans *et al.*, 2009) and significantly they were shown to increase the nodulation of alfalfa by *Rhizobium* species improving nutrition efficiency (Martínez-Hidalgo *et al*., 2014). Moreover, it has been demonstrated that the rhizosphere isolate *Micromonospora endolithica* promotes the growth of bean (*Phaseolus vulgaris L.*) (El-Tarabily *et al.*, 2008) and the endophytic *Micromonospora chalcea* can directly enhance plant growth, photosynthetic pigment production and seed yield in the halophilic crop *Salicornia bigelovii* (El-Tarabily *et al*., 2019).

Finally, another relevant application of *Micromonospora* strains includes their use as biocontrol agents. It has been demonstrated that *Micromonospora* can counteract different plant pathogens both *in vitro* and *in planta* through essential genes in the systemic acquired resistance or jasmonate/ethylene pathways (Hirsch & Valdés, 2010). *Micromonospora* strains have been reported to reduce root rot disease in *Banksia grandis L.* (El-Tarabily *et al.*, 1996), protect carrots from *Pythium coloratum* (El-Tarabily *et al.*, 1997), *Arabidopsis thaliana* from the bacterial pathogen *Erwinia carotovora* (Conn *et al.*, 2008), cucumber from *Pythium aphanidermatum* (El-Tarabily *et al.*, 2009), mango from *Lasiodiplodia theobromae* (Kamil *et al.*, 2018) and tomato plants from *Botrytis cinerea* (Martínez-Hidalgo *et al.*, 2015).

Other potential biotechnological applications of *Micromonospora* species involve their ability to produce plant biomass degrading enzymes for biofuel production (Hirsch & Valdés, 2010; Trujillo *et al.*, 2014) and their potential use as probiotics in aquacultures (Das *et al.*, 2008).

5.1.2 Engineering Micromonospora *species: state of the art.*

Micromonospora species are well-known for their natural ability synthesising a wide range of natural products, influencing plant growth and plant-microorganism interactions and acting as biocontrol agents among others. In order to fully exploit this potential, the domestication of strains and the standardisation of genetic engineering techniques play a key role. To date, there have been a considerable number of genetic studies on *Micromonospora* species, however, there is a lack of efficient tools and techniques for their genetic manipulation (Barka *et al.*, 2016).

During the 1980s and beginning of the 1990s, several studies focused on isolating plasmids from *Micromonospora* species in order to develop vectors to exploit the potential of the strains

within this genus. Plasmids isolated from *Micromonospora* species include pIM1 isolated from *M. inyoensis* (Parag & Goedeke, 1984), pMZ1 from *M. zionensis* (Oshida *et al.*, 1986), pMR1 and pMR2 from *M. rosaria* (Oshida *et al.*, 1986; Vukov & Vasiljevic, 1998), pMH101 isolated from *M. halophytica*, pMZ101 and pMZ201 from *M. zionensis* NRRL 5466, pMO101, pMO201 and pMO301 from *M. olivasterospora* (Hasegawa *et al.*, 1991), p11725a from *M. griseorubida* (Takada *et al.*, 1994) and pJTU112 from *Micromonospora* sp. strain 40027 (Li *et al.*, 2003). The first vectors specifically designed for cloning in *Micromonospora* were constructed from pMO101 and pMO201 and yielded pMO217, pMO133, pMO116, pMO126 and pMO136 (Hasegawa *et al.*, 1991). Also the vectors pTM1009, pTM1008 and pTM2001 and the *M. griseorubida-E. coli* shuttle cosmid pTYS507 were developed from p11725a (Inouye *et al.*, 1994; Takada *et al.*, 1994). The low-copy-number conjugative plasmid pJTU112 was also used to develop the cloning vectors pHZ199 and pOJ260 (Li *et al.*, 2003). Unfortunately, the characterisation of these vectors was limited to approximate estimations of the copy numbers and basic restriction enzyme maps. The first *Micromonospora* vector ever sequenced was pSPRX740 derived from the *M. rosaria* pMR2 plasmid and contained a codon optimised integrase that allowed direct site-specific integration into the chromosome of *M. halophytica* (Hosted *et al.*, 2005).

Meanwhile, members of the genus *Micromonospora* were successfully transformed, although with low efficiency, by protoplast transformation and conjugation using different vectors originally designed for other bacteria. The broad-host range *Streptomyces* plasmid pIJ702 was transformed into *M. rosaria*, *M. echinospora*, *M. melanospora* and *M. purpurea* (Matsushima & Baltz, 1988; Kelemen *et al.*, 1989; Kojic *et al.*, 1991; Love *et al.*, 1992), the *Streptomyces* cosmid vector pGM446 and the *E. coli* plasmid pTO1, possessing the *int* gene and *attP* site from *Streptomyces* phage phiC31 were introduced in *M. aurantica* (Voeykova *et al.*, 1998; Rose & Steinbuchel, 2002) and the *E. coli* plasmid pSET152 was for the first time used for heterologous gene cloning and expression in *Micromonospora* sp. strain 40027 (Li *et al.*, 2003). Only the *Micromonospora*-based plasmid pMO116 was used to study fortimicin A biosynthesis in *M. olivasterospora* (Dairi *et al.*, 1992). To date, just one study has reported the electroporation of a *Micromonospora* species (Li *et al.*, 2007).

During the last decade, many genetic studies have been carried out in *Micromonospora* species, mainly to characterise and enhance the production of antibacterial secondary metabolites. These studies include the description of the tetrocarcin A BGC from *M. chalcea* (Fang *et al.*, 2008), engineering rosamicin biosynthesis in *M. rosaria* IFO13697 introducing D-mycinose

biosynthetic genes from *M. griseorubida* A11725 to produce mycinosyl rosamicin derivatives (Anzai *et al.*, 2009; Iizaka *et al.*, 2014), increasing the production of the aminoglycoside antibiotic G418 in *M. echinospora* engineering the gentamycin BGC (Ni *et al.*, 2014), the assembly of a novel biosynthetic pathway for gentamicin B production in *Micromonospora echinospora* by combining twelve steps of gentamicin biosynthesis and two steps of kanamycin biosynthesis (Ni *et al.*, 2016) and overproduction of gentamicin B in the industrial strain *Micromonospora echinospora* CCTCC M 2018898 by cloning of missing genes (Chang *et al*., 2019). Despite the efforts put before into developing vectors and genetic modification methods specific for *Micromonospora* species, all these studies were carried out using *E. coli-Streptomyces* shuttle vectors and general *Streptomyces* conjugation protocols.

5.1.3 Design of Experiments.

Design of Experiments (DOE), is a systematic approach to understand how processes and the parameters within those processes affect an output or response variable (Wagner, J. R., Mount, E. M., & Giles, 2014). DOE is a mathematical tool that uses a statistical methodology to understand how different variables interact between each other in order to optimise the system's performance with the minimum amount of experimentation. Thus, it is used to differentiate significant factors from less important factors, to estimate the magnitude of the significant factors and to identify and estimate the magnitude of significant factor interactions. The foundations of statistical DOE were laid by Sir Ronald Fisher and Frank Yates during the first half of the 20th century (Stanley, 1966). Since then, DOE has been popularly applied in protein crystallography (Carter *et al.*, 1988), proteomics (Riter *et al.*, 2005), drug discovery (Tye, 2004), bioremediation (Landaburu-Aguirre *et al.*, 2009) and the development of functional materials (Bader *et al.*, 2019; Ryan *et al.*, 2019) among others.

DOE works as follows. First, the researcher studies the experimental system to identify all the factors involved in the experiment. For that, a preliminary model is established with free fitting parameters that are selected to be sufficiently general and flexible and one or several objectives are established. Then, there are several types of screening designs, such as Factorial, Fractional Factorial, Plackett-Burman and Definitive Screening Designs (DSDs), that can be applied depending on time and resources availability and the number of factors under consideration. During the first experimental stages, when the number of factors is large and time and resources scarce, DSDs are generally applied, as they offer several advantages over standard screening designs. DSDs have three levels, provide estimates of main effects that are unbiased by any

second-order effect, detect nonlinearity and identify the responsible factors of nonlinear effects while avoiding confounding any terms up to second order (Errore *et al.*, 2017). In addition to that, for designs with over five factors, the full quadratic model in any three factors can be efficiently estimated (Jones & Nachtsheim, 2011, 2017). Overall, it generates a greater understanding of the system under investigation in less time when compared to its counterparts. Furthermore, it characterises curvature and main effects independent of two factor interactions and requires only one more than twice as many runs as there are factors (Jones & Nachtsheim, 2016).

In order to fully exploit the biotechnological potential of a microorganism, the development of an efficient method to transfer DNA into the target strain is one of the first bottlenecks researchers generally face when developing a new genetic system (Zhang *et al.*, 2015). DOE has been used to adapt this kind of methods for specifically challenging bacteria. Previous efforts focused, for example, in optimising the electrotransformation of *Streptococcus thermophilus*, where eight quantitative factors including electrical, physiological and chemical parameters were studied by fractional factorial designs to identify the effects of individual parameters and the interactions between them (Marciset & Mollet, 1994). That same design was applied later to electroporate various strains of *Streptococcus salivarius* and the final method resulted to be applicable to other Gram-positive bacteria including *Streptococcus sanguis*, *Streptococcus vestibularis*, *Streptococcus mutans* and *Lactococcus lactis* (Buckley *et al.*, 1999). More recently, and whilst the one factor at a time approach was still being used to set up electroporation protocols for some genus such as *Arthtobacter* (Zhang *et al.*, 2011b), response surface methodology (RSM), a type of multifactorial experimental design, was used to study the effects of seven factors (growth media, growth phase, electric field, concentration of weakening agent, electroporation buffer, plasmid quantity and heat treatment) on the electroporation efficiency of the iturin A-producing *B. subtilis* ZK (Zhang *et al.*, 2015). A similar approach was followed to evaluate the effects of five variables (growth phase, voltage, resistance, plasmid DNA concentration, and the concentration of magnesium) on the transformation efficiency of *Acinetobacter baumannii* (Yildirim *et al.*, 2016). In this case, researchers used a three-level fractional factorial design to screen these factors in nine runs that allowed them to identify that growth phase, voltage, and resistance significantly contributed to the transformation efficiency (Yildirim *et al.*, 2016).

5.1.4 CRISPR systems for Actinobacteria.

Genetic engineering of *Actinobacteria*, and more specifically of genera recognised for their ability to produce natural products, has proved to be a highly valuable approach for the discovery of new bioactive molecules. During decades, genome editing these genera has relied in the use of RecA-mediated single- or double-crossover homologous recombination or λ-Red-mediated homologous recombination (Kieser *et al.*, 2000). However, the recent development of several different clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems for such organisms has provoked a shift in use towards these new methods that enable rapid, easy and often multiplexed genome editing (Tao *et al.*, 2018).

To date, four different CRISPR/Cas9-based toolkits for *in vivo* editing of *Actinobacteria* from the genus *Streptomyces* have been developed (Table 5.1). The first one, pCRISPomyces, consists of two plasmids pCRISPomyces-1 and pCRISPomyces-2 that differ only in the fact that the second one includes a sgRNA cassette in which the tracrRNA and the crRNA are fused to give a chimeric molecule (Cobb *et al.*, 2015). This second plasmid, pCRISPomyces-2, has proved to be the most useful one in terms of applicability for genome editing species of the genera *Streptomyces*, *Actinoplanes*, *Synechococcus* and *Paenibacillus* (Alberti & Corre, 2019). A second system, pKCcas9dO, characterised for the inducible expression of Cas9 and a *cas9* gene codon optimised for the codon usage of *S. coelicolor*, was designed to edit the model organism *S. coelicolor*, for which the pCRISPomyces system did not work (Huang *et al.*, 2015). The third system developed, consisted of the plasmids pCRISPR-Cas9 and pCRISPR-dCas9. Similarly to pKCcas9dO, pCRISPR-Cas9 and pCRISPR-dCas9 display an inducible promoter for Cas9 expression and the *cas9* gene codon optimised for *S. coelicolor*, but the promoter for the sgRNA expression is *ermE*p* instead (Tong *et al.*, 2015). This system was successfully applied in *Streptomyces*, *Corynebacterium* and *Micromonospora* strains (Alberti & Corre, 2019). The last system reported, pWHU2653, featured the *Streptomyces* constitutive *ermE*p* promoter upstream a codon optimised *cas9* and the constitutive *aac(3)IVp* promoter to control the expression of the sgRNA. In addition to that, this system included a cytosine deaminase gene *codA(sm)* as counter-selectable marker to accelerate the screening process (Zeng *et al.*, 2015). Of these four systems, only the first two can be used for multiplex genome editing (Table 5.1).

Despite the success of some of these systems, there are critical species-specific issues related to the use of CRISPR/Cas9 systems. Several studies have reported, among others, limitations related to the inherent toxicity associated with the expression of Cas9, the lack of alternative selectable markers and the risk of off-target effects (Cobb *et al.*, 2015; Alberti & Corre, 2019). In order to solve some of these constrains, an additional CRISPR/Cas system for *Streptomyces* was developed more recently, CRISPR-Cpf1 (Li *et al.*, 2018). In this new system, Cpf1, also known as Cas12a, is the endonuclease responsible for the targeted double-strand break (DSB). Unlike Cas9, it recognises T-rich PAM sequences and it is guided by a single CRISPR RNA (crRNA) without trans-activating crRNA (tracrRNA), making it easier to perform multiplex editing. Moreover, this system can also be used for multiplex gene repression through CRISPRi (Li *et al.*, 2018).

Table 5.1. Main features of the CRISPR-Cas9 editing systems for *Actinobacteria*. Modified from (Alberti & Corre, 2019).

	pCRISPomyces-1	pCRISPomyces-2	pKCcas9dO11	pCRISPR-Cas9	pCRISPR-dCas9	pWHU2653
Selection marker	Apramycin	Apramycin	Apramycin	Apramycin	Apramycin	Apramycin
Cas9 promoter	Constitutive: *rpsLp*	Constitutive: *rpsLp*	Inducible: *tipAp*	Inducible: *tipAp*	Inducible: *tipAp*	Constitutive: *aac(3)IVp*
Guide RNA promoter	Constitutive: *gapdhp* for crRNA and *rpsLp* for tracrRNA	Constitutive: *gapdhp* for sgRNA	Constitutive: *j23199p* for sgRNA	Constitutive: *ermE*p* for sgRNA	Constitutive: *ermE*p* for sgRNA	Constitutive: *ermE*p* for sgRNA
Plasmid clearance	Temperature-sensitive rep origin from pSG5	Temperature-sensitive rep origin from pSG5	Temperature-sensitive rep origin from pSG5	Temperature-sensitive rep origin from pSG5	Temperature-sensitive rep origin from pSG5	CodA counter-selectable marker
Species on which the plasmid was tested (% efficiency)	*S. lividans* (20–25%)	*S. lividans* (100%), *S. viridochromogenes* (67–100%), *S. albus* (67–100%), *S. formicae*, *S. rimosus*, *S. showdoensis*, *S. roseosporus*, *S. venezuelae*, *Actinoplanes* sp. SE50/110, *S. elongatus*, *P. polymyxaa*, *S. coelicolor* M1152 (tested but not successful), *Streptomyces* sp. KY 40-1 (tested but not successful)	*S. coelicolor* M145 (29–100%), *S.pristinaespiralis* (94%), *S. cinnamonensis*	*S. coelicolor* A3(2) (∼100%), *Corynebacterium glutamicum*, *Streptomyces* SD-85, *Micromonospora chersina*	*S. coelicolor* A3(2)	*S. coelicolor* M145 (93–99%)
Addgene ref. #	61 736	61 737	62 552	—	—	—
Study	(Cobb *et al.*, 2015)	(Cobb *et al.*, 2015)	(Huang *et al.*, 2015)	(Tong *et al.*, 2015)	(Tong *et al.*, 2015)	(Zeng *et al.*, 2015)

Although all these systems were developed specifically for *Streptomyces*, some of them have been used to edit *Actinobacteria* from other genera, including *Micromonospora*. More specifically, pCRISPR-Cas9 system was used to delete *genR* and *genS* genes involved in the production of gentamicin B by the industrial strain *Micromonospora echinospora* CCTCC M 2018898 (Chang *et al*., 2019). Additionally, a chromogenic screening-based CRISPR/Cas9 system was developed and applied in the abyssomicin producer *Verrucosispora* sp. MS100137, a strain of the family *Micromonosporaceae* yet to be reclassified as a *Micromonospora* species, since previous work proposed that the whole genus demands to be reclassified (Nouioui *et al.*, 2018; Wang *et al.*, 2020). This system, based on pWHU2653 (Zeng *et al.*, 2015), was specifically designed to address the time- and labour- intensive process of clone screening and plasmid curing as it features two chromogenic reporter systems (GusA and IdgS). The new vectors, pQS-gusA and pQS-idgS, include the *Streptomyces* codon-optimised *scas9* under the control of the thiostrepton inducible *tipA* promoter, the sgRNA under the control of the *ermE*p* promoter, the pIJ101 origin of replication and different chromogenic genes in each plasmid. Using this system, a 100% efficiency was achieved when editing the *S. coelicolor* M145 actinorhodin BGC, and *V*. sp. MS100137 carotene and abyssomicin BGCs (Wang *et al.*, 2020).

In this chapter, the application of statistical DOE for the development of an electroporation method that would facilitate the introduction of foreign DNA into *M. maris* is presented. In addition, a library of pCRISPomyces-2 vectors featuring promoters with different strengths for the expression of Cas9 was built in attempt to start optimising a genome editing tool for targeted genome editing in *M. maris*.

5.2 Materials and methods.

5.2.1 DOE for electrotransformation of Micromonospora.

DOE was used to tackle the transfer of DNA via electroporation into *Micromonospora maris* AB-18-032. The target goals were set as to, first, transfer any external DNA into *M. maris*, and then, maximise transformation efficiency. The main factors involved in the process were identified, although finally only those with *a priori* more impact on the system's performance were selected for the analysis and were accordingly categorised as numerical or categorical. Two or three factor levels were considered for each factor. Recipient cells - days old were initially categorised as categorical, as this would depend on the culture media (liquid or solid). This factor was numerically transformed as follows: when recipient cells were cultured in liquid media, young equalled 24 hours and old, 120 hours after inoculation. Due to the natural complexity of the electroporation method, some parameters were initially disregarded (Table 5.2). Recipient cells were only cultured in liquid media as growth is faster, the culture media used was tryptic soy broth (TSB) because again growth is faster and the media does not contain flours that could difficult the washing process, the number of recipient cells was maintained constant at ~ 10^7 CFU/ml, a unique combination of wash and electroporation buffers was used (Pigac & Schrempf, 1995), TSB was used as resuspension media, incubation time was set as 3 hours at 30 °C and 200 rpm and, finally, plating was done in TSA containing 20 µg/ml of apramycin (Table 5.2).

Table 5.2. Factors left outside the initial DOE design. Categorisation (N = numerical, C = categorical) and constant value established for each.

Factor	N/C	Value
Recipient cells - liquid culture or plate	C	Liquid
Recipient cells - origin media	C	TSB
Recipient cells – amount (CFU/ml)	N	10^7
Wash/electroporation buffers	C	(Pigac & Schrempf, 1995)
Resuspension media	C	TSB
Incubation – time (h)	N	3
Incubation – speed (rpm)	N	200
Incubation – temperature (°C)	N	30
Plating - media	C	TSA
Antibiotics - time after plating (h)	N	0

The factors considered to play a key role in the electroporation process were used to design a DSD using JMP Pro 13.0 (JMP®, Version Pro 13.0. SAS Institute Inc., Cary, NC, 1989-2019). These factors comprised: the maturity of the cells (days old), a pre-treatment with lysozyme, the concentration of DNA and the field strength, capacitance and resistance used (Table 5.3).

Table 5.3. Factors for experimentation, categorisation (N = numerical, C = categorical) and levels.

Factor	N/C	Low Level	Centre point	High Level
Recipient cells – maturity (days old)	N	1	-	5
Pre-treatment	C	Lysozyme	-	No pre-treatment
Pulse conditions - DNA concentration (ng)	N	0.05	250	500
Pulse conditions – field strength (kV/cm)	N	2.00	8.00	15.00
Pulse conditions - capacitance (μF)	N	10	25	50
Pulse conditions – resistance (Ω)	N	100	300	600

For phase I of the DOE applied to *M. maris* electroporation, the DSD yielded 18 different experiments (Table 5.4), for which the fraction of design space plot showed that the minimum relative prediction variance was slightly over 0.15 and the maximum was around 0.55 (Figure S5.1). The colour map on correlations showed small correlations between distinct terms with very few exceptions. The mean absolute value of the correlation between two effects was near 0.2 (Figure S5.2).

Table 5.4. Experiments comprising the phase I of *M. maris* electroporation.

Experiment id number	Age of cells (days)	Pre-treatment	DNA Concentration (ng)	Field Strength (kV)	Capacitance (μF)	Resistance (Ω)
1	5	No	500	2	10	300
2	5	No	0.05	2	50	100
3	1	Yes	0.05	2	50	600
4	5	No	500	15	10	100
5	5	Yes	0.05	8	10	100
6	1	No	0.05	15	25	100
7	1	Yes	0.05	15	50	300
8	5	Yes	500	2	25	600
9	5	Yes	0.05	15	10	600
10	5	Yes	500	15	50	100
11	1	No	0.05	2	10	600
12	1	Yes	250	2	10	100
13	1	Yes	500	15	10	600
14	5	No	250	15	50	600
15	5	No	250	8	25	300
16	1	No	500	8	50	600
17	1	Yes	250	8	25	300
18	1	No	500	2	50	100

Prior to starting all the electroporations, in order to determine the MIC of apramycin and thiostrepton for *M. maris* AB-18-032, ~10^8 spores of *M. maris* AB-18-032 were plated in TSA plates containing increasing concentrations of apramycin (0.125-1 mg per plate) or thiostrepton (0.5-1.5 mg per plate). Plates were incubated for 15 days at 30 °C.

Electroporation was carried out as follows: *M. maris* was grown in 50 ml TSB in a 250 ml baffled flask at 30 °C and 200 rpm for 24 or 120 hours. After that, cells were harvested by centrifugation at 4 °C and 4,500 rpm for 5 minutes. Cells were resuspended in 50 ml ice-cold 10% sucrose and centrifuged at 4,500 rpm for 5 minutes. Cells were then resuspended in 50 ml of 15% ice-cold glycerol and centrifuged at 4,500 rpm for 5 minutes. The pellet was resuspended in 10 ml ice-cold 15% glycerol, from which 5 ml were transferred into a tube containing 100 μg/ml lysozyme and incubated at 37 °C and 60 rpm for 30 minutes. The remaining 5 ml of cells were centrifuged at 4,500 rpm for 5 minutes and washed twice with 5 ml ice-cold 15% glycerol. The pellet was resuspended in 1 ml of 30% PEG 1500, 10% glycerol and 6.5% sucrose. For each electroporation, 50 μl of cells were mixed with 1 μl of the plasmid pSET152 at the corresponding concentration for each experiment (Table 5.4). The suspension was transferred into an ice-cold 1 mm-gapped electrocuvette and subjected to the corresponding

pulse (Table 5.4). Cells were immediately diluted 10-fold with TSB and incubated at 30 °C for 3 hours. After that, cells were centrifuged for 5 minutes at 10,000 rpm and the pellet was plated onto pre-warmed TSA plates containing 20 µg/ml apramycin. Plates were incubated at 30 °C for 10 days. A negative control was carried out with 1 µl of sterile water instead of DNA and serial dilutions of the competent cells were plated to calculate the CFU/ml used for the electroporation. After 10 days, individual colonies were resuspended into 20 µl of sterile water and plated into a new TSA plate with apramycin. After a week, a genomic DNA extraction was carried for each electrotransformant grown and pSET152 presence was verified as described later. All phase I electroporation experiments were carried out in triplicate. Moreover, all phase I electroporation was carried out twice; for the first round of electroporation, pSET152 had been purified from *E. coli* DH5α. The second time, namely phase I bis, the experiments were performed using pSET152 purified from the Dam and Dcm deficient *E. coli* JM110.

A second phase of DOE was carried out in order to maximise transformation efficiency. Accordingly, a second DSD was designed to specifically account for the effect of the field strength, capacitance and resistance (Table 5.5). Electroporation was carried out as described above after *M. maris* was cultured for 5 days, without lysozyme treatment prior to the electroporation and using 250 ng of pSET152 purified from the Dam and Dcm deficient *E. coli* JM110.

Table 5.5. Experiments comprising the phase II of *M. maris* electroporation.

Experiment id number	Field Strength (kV)	Resistance (Ω)	Capacitance (µF)
1	2.5	1000	25
2	1.5	800	25
3	1.5	1000	25
4	2.5	600	50
5	2.5	800	50
6	2.5	600	50
7	2.0	600	25
8	2.5	1000	50
9	1.5	600	50
10	2.5	1000	25
11	1.5	1000	25
12	2.0	800	50
13	1.5	600	50
14	1.5	600	25
15	2.0	1000	50
16	2.0	800	25
17	1.5	1000	50
18	2.5	600	25

5.2.2 Verification of electrotransformants.

The presence of pSET152 in *M. maris* electrotransformants was checked by PCR using the primers apraF and apraR (Table S5.1). Insertion of pSET152 into *M. maris* genome was checked by inverse PCR as follows. First, the genome was *in silico* digested to identify the enzymes with higher cutting frequency using http://insilico.ehu.es/ with the following configuration: endonucleases yielding fragments within a range 0 – 15,000, minimum recognition size for restriction enzymes was set as 4 and for type of restriction enzymes, the option only restriction enzymes with known bases (no N,R,Y...) was selected (Table S5.2) (accessed January 2020) (Bikandi *et al.*, 2004). NruI restriction enzyme, yielding 1,803 fragments of genomic DNA of an average size of 3,701 bp was selected. Genomic DNA of *M. maris* wild-type and electrotranformants was extracted (GenElute Bacterial Genomic DNA Kit; Sigma-Aldrich) and 1 µg was further digested with NruI for 2 hours at 37 °C and subjected to PCR clean-up. All the resulting DNA was ligated at 16 °C for 2 hours and finally, inverse PCR was performed using 50 ng of ligation product as template. Primers intpsetF and intpsetR (Table S5.1), and Q5 polymerase with 8 minutes of elongation time for 30 cycles were used. The PCR

product was visualised by gel electrophoresis to check the number of insertions in the genome and Sanger sequenced to identify the insertion location.

5.2.3 CRISPR-based genome editing in Micromonospora.

pCRISPomyces-2 was used as backbone to construct a library of CRISPR vectors. In order to investigate the influence and possible toxic effects of Cas9 expression in *Micromonospora maris*, eleven constitutive and inducible promoters with different strengths (Figure 5.1A and Table S5.3) were selected and cloned into pCRISPomyces-2 by PCR with Q5 Hot Start High Fidelity polymerase (NEB), followed by KLD treatment (NEB) (Figure 5.1B). All primers used are listed in Table S5.1. In all vectors, a sgRNA designed to target the promoter region of the main activator of the *M. maris* abyssomicin BGC, AbyI, was cloned via Golden Gate as described before (Cobb *et al.*, 2015). A 4 kb repair template containing the strong constitutive *Streptomyces* promoter *kasOp** with 2 kb homologous arms upstream and downstream *abyIp* was also cloned in all vectors to introduce *kasOp** promoter upstream the natural promoter of *abyI* (Figure 5.1B). All the vectors were fully sequenced before proceeding. A pCRISPomyces-2 based reporter vector was also built by replacing the saCas9 by superfolder GFP (sGFP). For that, pCRISPomyces-2 backbone and sGFP were amplified by PCR with primers containing HindIII restriction sites to be ligated by restriction-ligation.

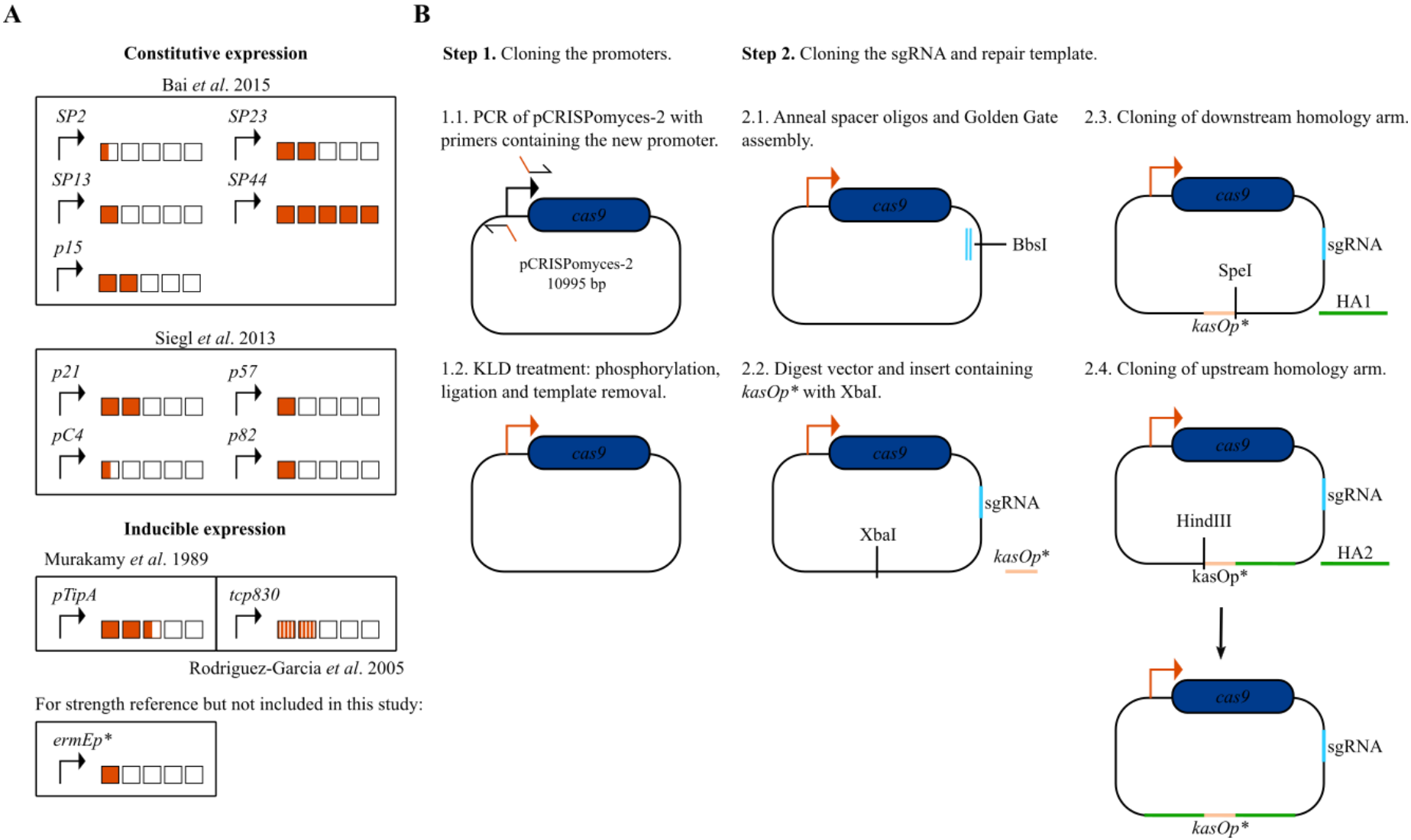

Figure 5.1. **A)** Promoters used for the expression of Cas9 and orientative strength in comparison to the strength of the *ermEp** promoter. **B)** Schematic representation of the construction of the promoter library to introduce *kasOp** upstream *abyI* in *M. maris*. HA stands for homology arm.

Vectors were transformed into chemically competent *E. coli* DH5α cells and miniprepped to send for sequencing and then transform chemically competent *E. coli* ET12567/pUZ8002 and *E. coli* JM110 cells. Both chemical competent cells were prepared as follows: a colony was inoculated into LB broth (with 50 µg/ml kanamycin and 25 µg/ml chloramphenicol for *E. coli* ET12567) and cultured overnight at 37 °C, the following day, 100 ml of LB broth (with 50 µg/ml kanamycin and 25 µg/ml chloramphenicol for *E. coli* ET12567) were inoculated with 1 ml of the overnight culture, cultured until OD_{600} reached 0.2-0.3, cooled in ice for 15 minutes, centrifuged at 1,000 x *g* for 5 minutes, washed once with 20 ml of ice cold 75 mM $MgCl_2$, washed once with 20 ml of ice cold 75 mM $CaCl_2$, resuspended in 1 ml of ice cold 75 mM $CaCl_2$ with 15% glycerol, aliquoted, frozen with liquid nitrogen and finally stored at -80 °C. Cells were transformed as follows: competent cells were thawed in ice for a minute, 100 µg of DNA was added in the tube, incubated in ice for 30 minutes, heat shocked for 30 seconds at 42 °C, incubated in ice for 5 minutes, resuspended in 1 ml LB broth, incubated at 37 °C for an hour, concentrated by centrifugation at 9,000 rpm for 2 minutes and spread in LB agar plates containing the appropriate antibiotics. *M. maris* transformations with all the CRISPR plasmids by electroporation and by conjugation could not be carried out.

In parallel, an attempt to test a second CRISPR editing system was made, this time using the chromogenic reporter based pQ-idgS vector (Wang *et al.*, 2020). The *kasOp** knock-in plasmid pQ-idgS-KasOp* was built as follows. The sgRNA scaffold was introduced using two primers containing the NcoI-XbaI terminal cohesive ends 5'-CATGG-N(20)-GTTTTAGAG-3' and 5'-CTAGCTCTAAAAC-N(20)-C-3' that were annealed and ligated into a NcoI-XbaI digested pQ-idgS backbone. The repair template containing *kasOp** was PCR amplified using the primers pQrtkasF and pQrtkasR from pCRISPOmyces-2 and introduced into pQ-idgS by Gibson assembly in the StuI site. The plasmid was sequence verified before transforming *E. coli* ET12567/pUZ8002 and *E. coli* JM110 cells, however, *M. maris* transformations were not carried out.

5.3 Results.

5.3.1 DOE for electrotransformation of Micromonospora.

In order to set up an electroporation protocol for *M. maris*, DOE was used to analyse how the different factors involved in the electroporation of DNA had an impact on the overall efficiency of the method. Those factors considered to play a key role in the electroporation process (cell maturity, lysozyme pre-treatment, DNA concentration and field strength, capacitance and resistance) were used to design a DSD. Prior to electroporation experiments, MIC curves of apramycin and thiostrepton for *M. maris* were also carried out to determine the optimal concentration of each antibiotic to be used during electroporation (Figure S5.3 and Figure S5.4).

During phase I of DOE applied to *M. maris* electroporation, the DSD yielded 18 different experiments (Table 5.4). From these electroporation experiments only four transformants per µg of DNA were obtained in experiment 16 (Figure 5.2). The time constant for each experiment was recorded (Figure S5.5) After that, all the experiments were repeated a second time using pSET152 purified from the Dam and Dcm deficient *E. coli* JM110. In this case, during phase I bis, four, eight and fourteen transformants per µg of DNA were obtained for experiments 4, 15 and 16, respectively and 104 transformants per µg of DNA were obtained from experiment 14 (Figure 5.2). The presence of pSET152 was confirmed in 15/15 randomly selected electrotransformants (Figure S5.7) and then insertion in the *attB* site present in *M. maris* chromosome was verified (Figure S5.8).

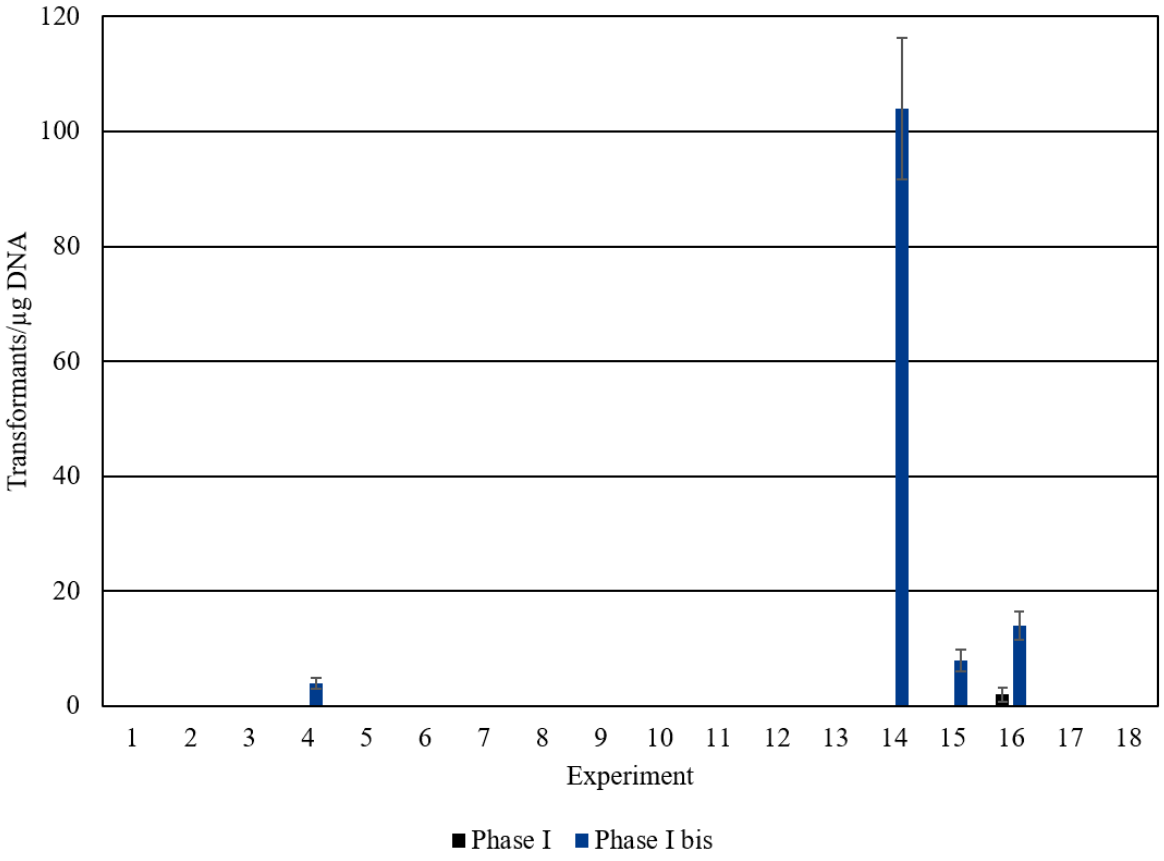

Figure 5.2. Transformants obtained per µg of DNA for each of the experiments carried out as part of the phase I and phase I bis.

Interestingly, the results obtained from phase I bis suggested that older cells, the absence of lysozyme pre-treatment and increasing field strength, capacitance and resistance had a positive effect on the recovery of *M. maris* electrotransformants (Figure 5.3). Increasing DNA concentration, although less notably, also had a positive impact on electrotransformant numbers.

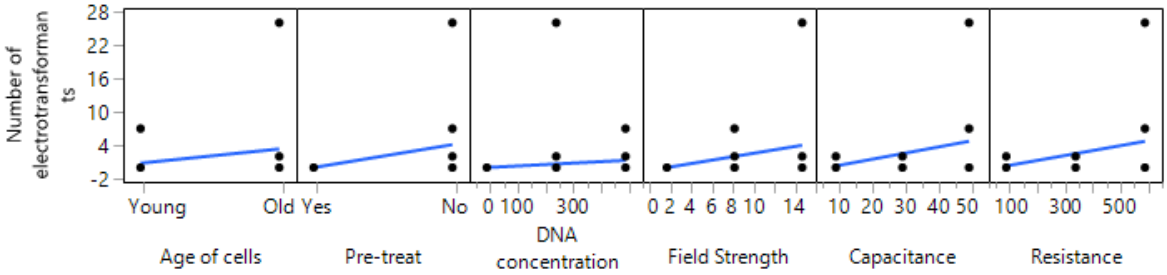

Figure 5.3. Main effects plot after phase I of *M. maris* electroporations.

Although more data would be needed in order to infer possible interactions between factors, *a priori*, using the data gathered from phase I bis electroporation, it seems likely that there were no interactions between the six factors analysed (Figure 5.4).

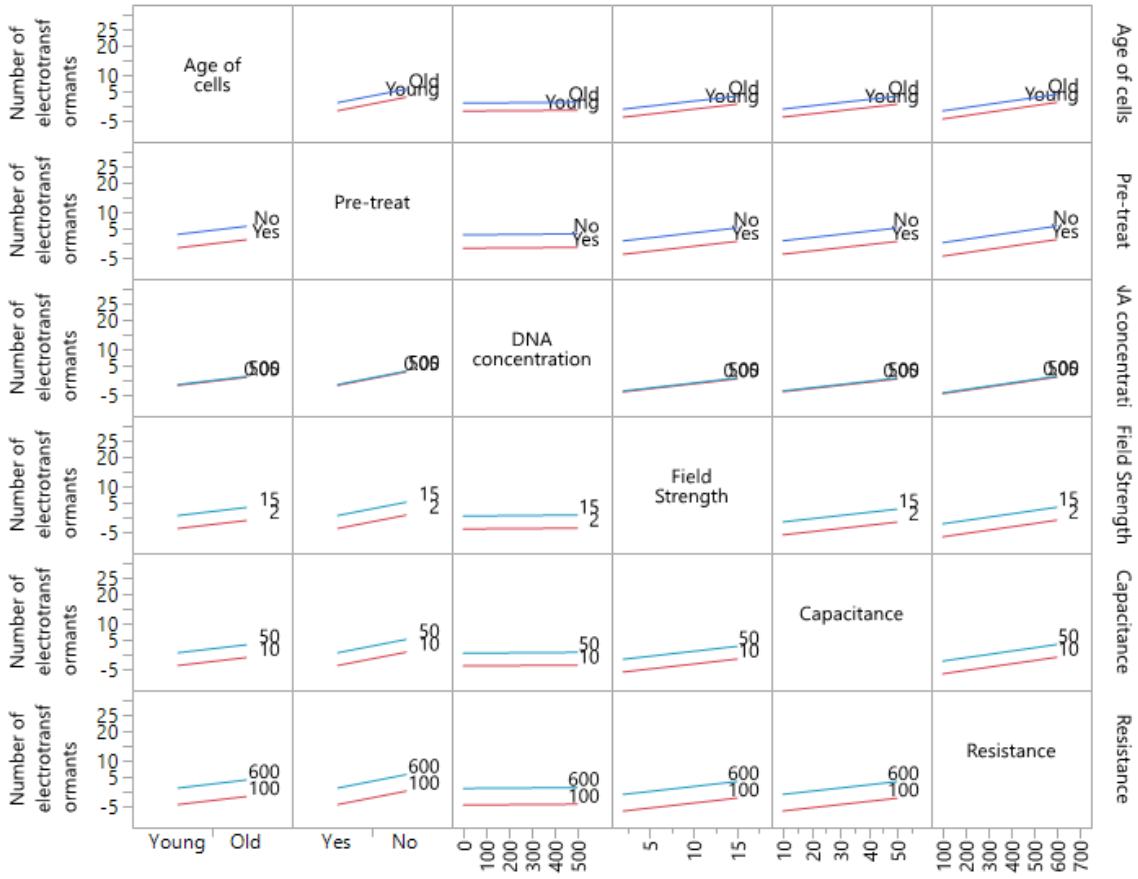

Figure 5.4. Interaction profile of the factors analysed during phase I bis of *M. maris* electroporation.

Since higher values of field strength, capacitance and resistance seemed to positively affect the recovery of electrotransformants, a second DSD was designed to specifically investigate the effect of the field strength, capacitance and resistance using the highest values possible for each factor (Table 5.5). When these experiments were carried out, 77.7% of the pulses arced, causing an important variability in the time constant values (Figure S5.9). The number of electrotransformants obtained did not improve on the results of phase I bis as only four transformants per µg of DNA were obtained from experiments 1, 13, 14 and 17 (Figure S5.10). These results were not sufficient to estimate the effects and interactions of the factors.

5.3.2 CRISPR-based genome editing in Micromonospora.

In order to investigate whether the use of CRISPR-Cas9 technologies could be a suitable approach to genome engineer *M. maris*, a library of CRISPR vectors was built. These were based on the pCRISPomyces-2 system developed previously for *Streptomyces* (Cobb *et al.*, 2015). First, conjugations in *M. maris* were carried out using the empty pCRISPomyces-2 (no sgRNA, no repair template), however, exconjugants were never obtained. Thus, with the aim of investigating the influence and possible toxic effects of Cas9 expression, eleven constitutive and inducible promoters with different strengths (Figure 5.1A and Table S5.3) were cloned into pCRISPomyces-2 replacing the native strong constitutive promoter *rpsLp*(XC). In all these vectors, a sgRNA targeting the promoter region of the main activator of *M. maris* abyssomicin BGC, AbyI, was cloned, as well as a 4 kb repair template containing the strong constitutive *Streptomyces* promoter *kasOp**. As a control to test if the pCRISPomyces-2 backbone was suitable for *M. maris*, a reporter vector was also built by replacing the saCas9 by sGFP. Vectors were transformed into *E. coli* DH5α, *E. coli* ET12567/pUZ8002 and *E. coli* JM110 cells, however, *M. maris* electroporations and conjugations were not carried out. In parallel, experiments to evaluate the chromogenic reporter based CRISPR editing system pQ-idgS were also initiated (Wang *et al.*, 2020). For that, a *kasOp** knock-in plasmid pQ-idgS-KasOp* was built to carry out the same edit as with pCRISPomyces-2, however, similarly, *M. maris* transformations were not carried out.

5.4 Discussion.

5.4.1 Genetic manipulation of M. maris.

DOE is a statistical approach to improve experimental efficiency and reliability. It is an iterative, empirical approach that systematically identifies and investigates the interactions between components within a system. The sets of experiments proposed through DOE are designed to uncover unexpected relationships between input variables (factors) and output variables (responses). In the present work, statistical DOE was used to study the effect of various factors involved in the electroporation of DNA into *Micromonospora maris* AB-18-032. This strain has previously shown its potential as producer of abyssomicins and proximicins (Roh *et al.*, 2011), however, its slow growth under laboratory conditions and the lack of tools to engineer it have prevented scientists from exploiting its full potential. Currently, the only reported method for introducing foreign DNA into *M. maris* is through conjugation via *E. coli* ET12567 as donor strain (Gottardi *et al.*, 2011; Wang *et al.*, 2020), however, this process is time consuming and the efficiency is low.

The aim of this work was to investigate whether it would be possible to establish an electroporation protocol for *M. maris*, using DOE to analyse the impact of each of the different factors involved in the process. Based on the results reported by previous researchers who studied the electroporation process at the detail, some of the factors that are critical include field strength, capacitance and resistance, as well as cell maturity, the use of a pre-treatment and the concentration of DNA (Rittich & Španová, 1996; Prasanna & Panda, 1997; Zhang *et al.*, 2011b). More recent work provided evidence on how the osmolytes used to stabilise weakened cells and the recovery time post-electroporation also contribute to higher transformation efficiencies (Bhattacharjee & Sorg, 2020; Sherba *et al.*, 2020). Here, the following six factors were selected to design the first DSD experiments: cell maturity, lysozyme pre-treatment, DNA concentration, field strength, capacitance and resistance. This yielded 18 different experiments (Table 5.4) that were carried out twice, the first time using DNA purified from *E. coli* DH5α and the second time using DNA purified from the Dam and Dcm deficient *E. coli* JM110 (Figure 5.2). The differences in the transformants/µg of DNA observed between these two rounds of experiments suggest that *M. maris* could, as many *Streptomyces*, possess a methyl-specific restriction system (Flett *et al.*, 2006). Moreover, the results obtained from phase I bis also suggested that older cells, the absence of lysozyme pre-treatment and increasing field strength, capacitance and resistance could have a positive effect on the obtention of *M. maris*

electrotransformants (Figure 5.3). It has been reported before that high-voltages facilitate DNA transfer as well as field strength and pulse duration (Prasanna & Panda, 1997). For that reason, a second DSD in which high values of field strength, capacitance and resistance were the only factors to study was designed (Table 5.5). Nonetheless, none of the experiments derived from this DSD improved the results obtained during phase I bis (Figure S5.10). On a positive note, from the electrotransformants checked from phases I, I bis and II, a 100% had been transformed and in all of them pSET152 had been inserted in the *attB* site present in *M. maris* chromosome (Figure S5.7 and Figure S5.8), indicating the absence of false positives.

Overall, it is clear from the literature that the electroporation conditions are species specific, at least for bacteria, and that these vary depending on the composition and thus rigidity of the cell wall (Kieser *et al.*, 2000; Zeaiter *et al.*, 2018). Statistical DOE proved to be useful in the past to establish electroporation methods for species such as *S. thermophilus* (Marciset & Mollet, 1994), *Staphylococcus carnosus* (Löfblom *et al.*, 2007), *Arthrobacter* species (Zhang *et al.*, 2011b), *Bacillus amyloliquefaciens* (Zhang *et al.*, 2011a) and *A. baumannii* (Yildirim *et al.*, 2016). In this chapter, promising preliminary results suggest that the optimal electroporation conditions for *M. maris* AB-18-032 would be using 5 days old liquid cultures, without any pre-treatment, using 250 ng of DNA and a pulse of 15 kV, 50 μF and 600 Ω. Future work could look at trying to increase the electroporation efficiency for example by designing a new DSD with the factors initially disregarded. Moreover, the new protocol would need to be tested with other vectors larger in size. For the matter of this dissertation, *M. maris* AB-18-032 could have been electrotransformed with the abyssomicin BGC containing phage artificial chromosome pESAC13_*aby* (139 kb), for example, as it contains the same *attP* site of pSET152 and would have inserted a second copy of the *aby* BGC in *M. maris* genome. Moreover, it would be interesting to test this electroporation protocol in other *Micromonospora* such as *M. chalcea*, *M. echinospora* or *M. rosaria* and, in case of it not been applicable for these species, the first DSD experiments presented here could be tested, as it was initially designed to be sufficiently general.

Finally, in order to expand *M. maris* toolbox, further work on the use of CRISPR-Cas9 would need to be carried out. The library of CRISPR-Cas9 vectors that was built based on the pCRISPomyces-2 system could be transferred into *M. maris* both via conjugation and electroporation. This way, both methods could be compared in terms of experimental time taken, cost, number of correct transformants recovered and stability of the transformants once

subcultured. Specially because, according to my own experience, false positives are quite likely to appear from conjugations as well as unstable transformants that stop growing after subculturing, whereas none of this happened when carrying out the electroporations. Additionally, conjugations can result in rearrangements within the vector while it is in the donor strain. Finally, by using a second CRISPR vector such as pQ-idgS (Wang *et al.*, 2020) to carry out the same edits, it would be possible to compare the genome editing efficiency of both systems.

5.5 Conclusions and future perspectives.

The aim of this part of my dissertation was to start developing general molecular microbiology methods for *M. maris* that would facilitate engineering it in the future. Statistical DOE proved to be an effective tool to identify the most optimal electroporation conditions for this organism. Older cells, the absence of lysozyme pre-treatment and increasing field strength, capacitance and resistance seemed to have a positive effect on the obtention of *M. maris* electrotransformants. The highest electroporation efficiency was obtained when 5 days old cells were electroporated, without being subjected to lysozyme treatment, with 250 ng of DNA at a field strength of 15 kV, a capacitance of 50 μF and a resistance of 600 Ω. To validate the protocol, other plasmids would also need to be introduced into *M. maris*. Moreover, the results indicate that *M. maris* has a methyl-specific restriction system. Future work could focus on the design of a new set of experiments with the factors initially disregarded to improve the results presented here. Finally, the final electroporation method, or in case of negative results the sets of experiments presented in Table 5.4, could be used to electroporate other *Micromonospora* species.

Lastly, more work is still needed to find and optimise a genome editing tool that allows quick and reliable engineering in *M. maris*. For that, different approaches could be followed, including the development from scratch of a completely new system, however, generally, it has been more fruitful to redesign or adapt systems used in closely related species. Currently, the most promising system seems to be the chromogenic screening-based CRISPR/Cas9 system developed for the abyssomicin producer *Verrucosispora* sp. MS100137, although it could also be interesting to study how efficient different CRISPR-associated endonucleases under the expression of various promoters would be for genome editing *M. maris*.

5.6 References.

Alberti, F. & Corre, C. (2019). Editing *Streptomycete* genomes in the CRISPR/Cas9 age. *Nat Prod Rep* **36**, 1237–1248.

Anzai, Y., Iizaka, Y., Li, W., Naoki, ·, Shu-Ichi Tsukada, I. ·, Koike, K., Kinoshita, K. & Kato, F. (2009). Production of rosamicin derivatives in *Micromonospora rosaria* by introduction of D-mycinose biosynthetic gene with C31-derived integration vector pSET152. *J Ind Microbiol Biotechnol* **36**, 1013–1021.

Bader, A., Toenjes, A., Wielki, N., Mändle, A., Onken, A.-K., Hehl, A. von, Meyer, D., Brannath, W. & Tracht, K. (2019). Parameter optimization in high-throughput testing for structural materials. *Materials (Basel)* **12**, 3439.

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y. & van Wezel, G. P. (2016). Taxonomy, physiology, and natural products of *Actinobacteria*. *Microbiol Mol Biol Rev* **80**, 1–43.

Bhattacharjee, D. & Sorg, J. A. (2020). Factors and conditions that impact electroporation of *Clostridioides difficile* strains. *mSphere* **5**.

Bikandi, J., Millan, R. S., Rementeria, A. & Garaizar, J. (2004). *In silico* analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics* **20**, 798–799.

Boumehira, A. Z., El-Enshasy, H. A., Hacène, H., Elsayed, E. A., Aziz, R. & Park, E. Y. (2016). Recent progress on the development of antibiotics from the genus *Micromonospora*. *Biotechnol Bioprocess Eng* **21**, 199–223.

Buckley, N. D., Vadeboncoeur, C., LeBlanc, D. J., Lee, L. N. & Frenette, M. (1999). An effective strategy, applicable to *Streptococcus salivarius* and related bacteria, to enhance or confer electroporation competence. *Appl Environ Microbiol* **65**, 3800–4.

Carro, L., Pujic, P., Trujillo, M. E. & Normand, P. (2013). *Micromonospora* is a normal occupant of actinorhizal nodules. *J Biosci* **38**, 685–693.

Carro, L., Riesco, R., Spröer, C. & Trujillo, M. E. (2016). *Micromonospora luteifusca* sp. nov. isolated from cultivated *Pisum sativum*. *Syst Appl Microbiol* **39**, 237–242.

Carro, L., Nouioui, I., Sangal, V., Meier-Kolthoff, J. P., Trujillo, M. E., Montero-Calasanz, M. del C., Sahin, N., Smith, D. L., Kim, K. E. & other authors. (2018). Genome-based classification of *Micromonosporae* with a focus on their biotechnological and ecological potential. *Sci Rep* **8**, 525.

Carro, L., Castro, J. F., Razmilic, V., Nouioui, I., Pan, C., Igual, J. M., Jaspars, M., Goodfellow, M., Bull, A. T. & other authors. **(2019).** Uncovering the potential of novel *Micromonosporae* isolated from an extreme hyper-arid Atacama Desert soil. *Sci Rep* **9**.

Carter, C. W., Baldwin, E. T. & Frick, L. (1988). Statistical design of experiments for protein crystal growth and the use of a precrystallization assay. *J Cryst Growth* **90**, 60–73.

Chang, Y., Chai, B., Ding, Y., He, M., Zheng, L., Teng, Y., Deng, Z., Yu, Y. & Liu, T. (2019). Overproduction of gentamicin B in industrial strain *Micromonospora echinospora* CCTCC M 2018898 by cloning of the missing genes genR and genS. *Metab Eng Commun* e00096.

Charan, R. D., Schlingmann, G., Janso, J., Bernan, V., Feng, X. & Carter, G. T. (2004). Diazepinomicin, a new antimicrobial alkaloid from a marine *Micromonospora* sp. *J Nat Prod* **67**, 1431–1433.

Cobb, R. E., Wang, Y. & Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth Biol* 2–7.

Conn, V. M., Walker, A. R. & Franco, C. M. M. (2008). Endophytic *Actinobacteria* induce defense pathways in *Arabidopsis thaliana*. *Mol Plant-Microbe Interact* **208**, 208–218.

Contreras-Castro, L., Maldonado, L. A., Quintana, E. T., Carro, L. & Klenk, H.-P. (2019). Genomic insight into three marine *Micromonospora* sp. Strains from the Gulf of California. *Microbiol Resour Announc* **8**.

Coombs, J. T. & Franco, C. M. M. (2003). Isolation and identification of *Actinobacteria* from surface-sterilized wheat roots. *Appl Environ Microbiol* **69**, 5603–5608.

Dairi, T., Ohta, T., Hashimoto, E. & Hasegawa, M. (1992). Self cloning in *Micromonospora olivasterospora* of *fms* genes for fortimicin A (astromicin) biosynthesis. *MGG Mol Gen Genet* **232**, 262–270.

Das, S., Ward, L. R. & Burke, C. (2008). Prospects of using marine *Actinobacteria* as probiotics in aquaculture. *Appl Microbiol Biotechnol* **81**, 419–429.

El-Tarabily, K. A., Nassar, A. H., Hardy, G. E. S. J. & Sivasithamparam, K. (2009). Plant growth promotion and biological control of *Pythium aphanidermatum*, a pathogen of cucumber, by endophytic actinomycetes. *J Appl Microbiol* **106**, 13–26.

El-Tarabily, K. A., Sykes, M. L., Kurtböke, I. D., Hardy, G. E. S. J., Barbosa, A. M. & Dekker, R. F. H. (1996). Synergistic effects of a cellulase-producing *Micromonospora carbonacea* and an antibiotic-producing *Streptomyces violascens* on the suppression of *Phytophthora cinnamomi* root rot of *Banksia grandis*. *Can J Bot* **74**, 618–624.

El-Tarabily, K. A., Hardy, G. E. S. J., Sivasithamparam, K., Hussein, A. M. & Kurtböke, D. I. (1997). The potential for the biological control of cavity-spot disease of carrots, caused by *Pythium coloratum*, by streptomycete and non-streptomycete actinomycetes. *New Phytol* **137**, 495–507.

El-Tarabily, K. A., Nassar, A. H. & Sivasithamparam, K. (2008). Promotion of growth of bean (*Phaseolus vulgaris* L.) in a calcareous soil by a phosphate-solubilizing, rhizosphere-competent isolate of *Micromonospora endolithica*. *Appl Soil Ecol* **39**, 161–171.

El-Tarabily, K. A., AlKhajeh, A. S., Ayyash, M. M., Alnuaimi, L. H., Sham, A., ElBaghdady, K. Z., Tariq, S. & AbuQamar, S. F. (2019). Growth promotion of *Salicornia bigelovii* by *Micromonospora chalcea* UAE1, an endophytic 1-aminocyclopropane-1-carboxylic acid deaminase-producing actinobacterial isolate. *Front Microbiol* **10**.

Errore, A., Jones, B., Li, W. & Nachtsheim, C. J. (2017). Using definitive screening designs to identify active first-and second-order factor effects. *J Qual Technol* **49**, 244–264.

Fang, J., Zhang, Y., Huang, L., Jia, X., Zhang, Q., Zhang, X., Tang, G. & Liu, W. (2008). Cloning and characterization of the tetrocarcin A gene cluster from *Micromonospora chalcea* NRRL 11289 reveals a highly conserved strategy for tetronate biosynthesis in spirotetronate antibiotics †. *J Bacteriol* **190**, 6014–6025.

Flett, F., Mersinias, V. & Smith, C. P. (2006). High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting *Streptomycetes*. *FEMS Microbiol Lett* **155**, 223–229.

Ganguly, A. K. (1982). Antibiotics from *Micromonospora*. *Ann Rev Microbiol*, 38–43.

Garcia, L. C., Martinez-Molina, E. & Trujillo, M. E. (2010). *Micromonospora pisi* sp. nov., isolated from root nodules of *Pisum sativum*. *Int J Syst Evol Microbiol* **60**, 331–337.

Gottardi, E. M., Krawczyk, J. M., Von Suchodoletz, H., Schadt, S., Mühlenweg, A., Uguru, G. C., Pelzer, S., Fiedler, H. P., Bibb, M. J. & other authors. **(2011).** Abyssomicin biosynthesis: Formation of an unusual polyketide, antibiotic-feeding studies and genetic analysis. *ChemBioChem* **12**, 1401–1410.

Hasegawa, M., Dairi, T., Ohta, T. & Hashimoto, E. (1991). A novel, highly efficient gene-cloning system for *Micromonospora* strains. *J Bacteriol* **173**, 7004–7011.

Hirsch, A. M. & Valdés, M. (2010). *Micromonospora*: An important microbe for biomedicine and potentially for biocontrol and biofuels. *Soil Biol Biochem* **42**, 536–542.

Hirsch, P., Mevs, U., Kroppenstedt, R. M., Schumann, P. & Stackebrandt, E. (2004). Cryptoendolithic *Actinomycetes* from antarctic sandstone rock samples: *Micromonospora*

endolithica sp. nov. and two isolates related to *Micromonospora coerulea* Jensen 1932. *Syst Appl Microbiol* **27**, 166–174.

Hosted, T. J., Wang, T. & Horan, A. C. (2005). Characterization of the *Micromonospora rosaria* pMR2 plasmid and development of a high G + C codon optimized integrase for site-specific integration. *Plasmid* **54**, 249–258.

Huang, H., Zheng, G., Jiang, W., Hu, H. & Lu, Y. (2015). One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim Biophys Sin (Shanghai)* **47**, 231–43.

Igarashi, Y., Miyanaga, S., Onaka, H., Takeshita, M. & Furumai, T. (2005). Revision of the structure assigned to the antibiotic BU-4664L from *Micromonopora*. *J Antibiot (Tokyo)* **58**, 350–352.

Igarashi, Y., Yanase, S., Sugimoto, K., Enomoto, M., Miyanaga, S., Trujillo, M. E., Saiki, I. & Kuwahara, S. (2011). Lupinacidin C, an inhibitor of tumor cell invasion from *Micromonospora lupini*. *J Nat Prod* **74**, 862–865.

Iizaka, Y., Higashi, N., Li, W., Fukumoto, · Atsushi, Anzai, Y. & Kato, F. (2014). A new mycinosyl rosamicin derivative produced by an engineered *Micromonospora rosaria* mutant with a cytochrome P450 gene disruption introducing the d-mycinose biosynthetic gene. *J Ind Microbiol Biotechnol* **41**, 1451–1456.

Inouye, M., Takada, Y., Muto, N., Beppu, T. & Horinouchi, S. (1994). Characterization and expression of a P-450-like mycinamicin biosynthesis gene using a novel *Micromonospora-Escherichia coli* shuttle cosmid vector. *MGG Mol Gen Genet* **245**, 456–464.

Jackson, M., Karwowski, J. P., Theriault, R. J., Rasmussen, R. R., Hensey, D. M., Humphrey, P. E., Swanson, S. J., Barlow, G. J., Premachandran, U. & McAlpine, J. B. (1995). Macquarimicins, microbial metabolites from *Micromonospora*: I. Discovery, taxonomy, fermentation and biological properties. *J Antibiot (Tokyo)* **48**, 462–466.

Jones, B. & Nachtsheim, C. J. (2011). A class of three-level designs for definitive screening in the presence of second-order effects. *J Qual Technol* **43**, 1–15.

Jones, B. & Nachtsheim, C. J. (2016). Blocking schemes for definitive screening designs. *Technometrics* **58**, 74–83.

Jones, B. & Nachtsheim, C. J. (2017). Effective design-based model selection for definitive screening designs. *Technometrics* **59**, 319–329.

Kamil, F. H., Saeed, E. E., El-Tarabily, K. A. & AbuQamar, S. F. (2018). Biological control of mango dieback disease caused by *Lasiodiplodia theobromae* using streptomycete and non-streptomycete *Actinobacteria* in the United Arab Emirates. *Front Microbiol* **9**.

Kelemen, G. H., Financsek, I. & JáRai, M. (1989). Efficient transformation of *Micromonospora purpurea* with PIJ702 plasmid. *J Antibiot (Tokyo)* **42**, 325–328.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. a. **(2000).** Practical Streptomyces Genetics. *John Innes Found Norwich,* UK ISBN 0-7084-0623-8.

Kojic, M., Topisirovic, L. & Vasiljevic, B. (1991). Efficient transformation of *Micromonospora melanosporea* protoplasts by *Streptomyces* plasmid. *Curr Microbiol* **23**, 343–345.

Lam, K. S., Hesler, G. A., Gustavson, D. R., Berry, R. L., Tomita, K., MacBeth, J. L., Ross, J., Miller, D. & Forenza, S. (1996). Pyrrolosporin A, a new antitumor antibiotic from *Micromonospora* sp. C39217-r109-7. I. Taxonomy of producing organism, fermentation and biological activity. *J Antibiot (Tokyo)* **49**, 860–864.

Landaburu-Aguirre, J., García, V., Pongrácz, E. & Keiski, R. L. (2009). The removal of zinc from synthetic wastewaters by micellar-enhanced ultrafiltration: statistical design of experiments. *Desalination* **240**, 262–269.

Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W. & Lu, Y. (2018). CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. *Appl Environ Microbiol* **84**.

Li, X., Long, C., Zhou, X. & Deng, Z. (2007). Study on electroporation of rare *Actinomycete Micromonospora* sp. 40027 with *Streptomyces* plasmid pSET152. *Wei sheng wu xue bao* **47**, 718–20.

Li, X., Zhou, X. & Deng, Z. (2003). Vector systems allowing efficient autonomous or integrative gene cloning in *Micromonospora* sp. Strain 40027. *Appl Environ Microbiol* **69**, 3144.

Löfblom, J., Kronqvist, N., Uhlén, M., Ståhl, S. & Wernérus, H. (2007). Optimization of electroporation-mediated transformation: *Staphylococcus carnosus* as model organism. *J Appl Microbiol* **102**, 736–747.

Love, S. F., Maiese, W. M. & Rothstein, D. M. (1992). Conditions for protoplasting, regenerating, and transforming the calicheamicin producer, *Micromonospora echinospora*. *Appl Environ Microbiol* **4**, 1376–1378.

Lubsanova, D. A., Zenova, G. M., Kozhevin, P. A., Manucharova, N. A. & Shvarov, A. P. (2014). Filamentous *Actinobacteria* of the saline soils of arid territories. *Moscow Univ Soil Sci Bull* **69**, 88–92.

Manivasagan, P., Venkatesan, J., Sivakumar, K. & Kim, S.-K. (2014). Pharmaceutically active secondary metabolites of marine *Actinobacteria*. *Microbiol Res* **169**, 262–278.

Marciset, O. & Mollet, B. (1994). Multifactorial experimental design for optimizing transformation: Electroporation of *Streptococcus thermophilus*. *Biotechnol Bioeng* **43**, 490–496.

Martínez-Hidalgo, P., Galindo-Villardón, P., Igual, J. M. & Martínez-Molina, E. (2014). *Micromonospora* from nitrogen fixing nodules of alfalfa (*Medicago sativa* L.). A new promising Plant Probiotic Bacteria. *Sci Rep* **4**.

Martínez-Hidalgo, P., García, J. M. & Pozo, M. J. (2015). Induced systemic resistance against *Botrytis cinerea* by *Micromonospora* strains isolated from root nodules. *Front Microbiol* **6**.

Matsushima, P. & Baltz, R. H. (1988). Genetic transformation of *Micromonospora rosaria* by the *Streptomyces* plasmid pIJ702. *J Antibiot (Tokyo)* **41**, 583–585.

Ni, X., Sun, Z., Zhang, H., He, H., Ji, Z. & Xia, H. (2014). Genetic engineering combined with random mutagenesis to enhance G418 production in *Micromonospora echinospora*. *J Ind Microbiol Biotechnol* **41**, 1383–1390.

Ni, X., Sun, Z., Gu, Y., Cui, H. & Xia, H. (2016). Assembly of a novel biosynthetic pathway for gentamicin B production in *Micromonospora echinospora*. *Microb Cell Fact* **15**, 1.

Norovsuren, Z., Oborotov, G. V., Zenova, G. M., Aliev, R. A. & Zvyagintsev, D. G. (2007). Haloalkaliphilic actinomycetes in soils of Mongolian desert steppes. *Biol Bull* **34**, 417–422.

Nouioui, I., Carro, L., García-López, M., Meier-Kolthoff, J. P., Woyke, T., Kyrpides, N. C., Pukall, R., Klenk, H.-P., Goodfellow, M. & Göker, M. (2018). Genome-based taxonomic classification of the phylum Actinobacteria. *Front Microbiol* **9**.

Okachi, R., Kawamoto, I., Takasawa, S., Yamamoto, M., Sato, S., Sato, T. & Nara, T. (1974). A new antibiotic XK-62-2 (Sagamicin). I. Isolation, physicochemical and antibacterial properties. *J Antibiot (Tokyo)* **27**, 793–800.

Oshida, T., Takeda, K., Yamaguchi, T., Ohshima, S. & Ito, Y. (1986). Isolation and characterization of plasmids from *Micromonospora zionensis* and *Micromonospora rosaria*. *Plasmid* **16**, 74–76.

Pandey, R. C., Mishra, S. K. & Putnam, A. R. (1992). Antifungal anthracycline antibiotics, spartanamicins A and B from *Micromonospora* Spp. *J Antibiot (Tokyo)* **45**, 1738–1745.

Parag, Y. & Goedeke, M. E. (1984). A plasmid of the sisomicin producer *Micromonospora inyoensis*. *J Antibiot (Tokyo)* **37**, 1082–1084.

Parte, A. C. (2014). LPSN—list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res* **42**, D613–D616.

Pigac, J. & Schrempf, H. (1995). A simple and rapid method of transformation of *Streptomyces rimosus* R6 and other streptomycetes by electroporation. *Appl Environ Microbiol* **61**, 352–356.

Prasanna, G. L. & Panda, T. (1997). Electroporation: Basic principles, practical considerations and applications in molecular biology. *Bioprocess Eng* **16**, 261–264.

Riter, L. S., Vitek, O., Gooding, K. M., Hodge, B. D. & Julian, R. K. (2005). Statistical design of experiments as a tool in mass spectrometry. *J Mass Spectrom* **40**, 565–579.

Rittich, B. & Španová, A. (1996). Electrotransformation of bacteria by plasmid DNAs: Statistical evaluation of a model quantitatively describing the relationship between the number of electrotransformants and DNA concentration. *Bioelectrochemistry Bioenerg*, 233–238.

Roh, H., Uguru, G. C., Ko, H. J., Kim, S., Kim, B. Y., Goodfellow, M., Bull, A. T., Kim, K. H., Bibb, M. J. & other authors. **(2011).** Genome sequence of the abyssomicin- and proximicin-producing marine actinomycete *Verrucosispora maris* AB-18-032. *J Bacteriol* **193**, 3391–3392.

Rose, K. & Steinbuchel, A. (2002). Construction and intergeneric conjugative transfer of a pSG5-based cosmid vector from *Escherichia coli* to the polyisoprene rubber degrading strain *Micromonospora aurantiaca* W2b. *FEMS Microbiol Lett* **211**, 129–132.

Ryan, E., Pollard, Z. A., Ha, Q. T., Roshandelpoor, A., Vakili, P. & Goldfarb, J. L. (2019). Designing heterogeneous hierarchical material systems: A holistic approach to structural and materials design. *MRS Commun* **9**, 628–636.

Satoi, S., Muto, N., Hayashi, M., Fujii, T. & Otani, M. (1980). Mycinamicins, new macrolide antibiotics. I. Taxonomy, production, isolation, characterization and properties. *J Antibiot (Tokyo)* **33**, 364–376.

Sherba, J. J., Hogquist, S., Lin, H., Shan, J. W., Shreiber, D. I. & Zahn, J. D. (2020). The effects of electroporation buffer composition on cell viability and electro-transfection efficiency. *Sci Rep* **10**, 1–9.

Solans, M. (2007). *Discaria trinervis* - *Frankia* symbiosis promotion by saprophytic actinomycetes. *J Basic Microbiol* **47**, 243–250.

Solans, M., Vobis, G. & Wall, L. G. (2009). Saprophytic actinomycetes promote nodulation in *Medicago sativa-Sinorhizobium meliloti* symbiosis in the presence of high N. *J Plant Growth Regul* **28**, 106–114.

Stanley, J. C. (1966). The influence of Fisher's 'The Design of Experiments' on educational research thirty years later. *Am Educ Res J* **3**, 223.

Supong, K., Suriyachadkun, C., Pittayakhajonwut, P., Suwanborirux, K. & Thawai, C. (2013). *Micromonospora spongicola* sp. nov., an actinomycete isolated from a marine sponge in the Gulf of Thailand. *J Antibiot (Tokyo)* **66**, 505–509.

Takada, Y., Inouye, M., Morohoshi, T., Muto, N., Kato, F., Kizuka, M., Tanaka, M. & Koyama, Y. (1994). Establishment of the host-vector system for *Micromonospora griseorubida*. *J Antibiot (Tokyo)* **47**, 1167–1170.

Tao, W., Yang, A., Deng, Z. & Sun, Y. (2018). CRISPR/Cas9-based editing of *Streptomyces* for discovery, characterization, and production of natural products. *Front Microbiol* **9**, 1660.

Tian, X., Cao, L., Tan, H., Han, W., Chen, M., Liu, Y. & Zhou, S. (2007). Diversity of cultivated and uncultivated actinobacterial endophytes in the stems and roots of rice. *Microb Ecol* **53**, 700–707.

Tiwari, K. & Gupta, R. K. (2012). Rare actinomycetes: A potential storehouse for novel antibiotics. *Crit Rev Biotechnol* **2**, 108-132.

Tong, Y., Charusanti, P., Zhang, L., Weber, T. & Lee, S. Y. (2015). CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS Synth Biol* **4**, 1020–1029.

Trujillo, M. E., Kroppenstedt, R. M., Schumann, P., Carro, L. & Martínez-Molina, E. (2006). *Micromonospora coriariae* sp. nov., isolated from root nodules of *Coriaria myrtifolia*. *Int J Syst Evol Microbiol* **56**, 2381–2385.

Trujillo, M. E., Kroppenstedt, R. M., Fernández-Molinero, C., Schumann, P. & Martínez-Molina, E. (2007). *Micromonospora lupini* sp. nov. and *Micromonospora saelicesensis* sp. nov., isolated from root nodules of *Lupinus angustifolius*. *Int J Syst Evol Microbiol* **57**, 2799–2804.

Trujillo, M. E., Bacigalupe, R., Pujic, P., Igarashi, Y., Benito, P., Riesco, R., Médigue, C. & Normand, P. (2014). Genome features of the endophytic *Actinobacterium Micromonospora lupini* strain Lupac 08: On the process of adaptation to an endophytic life style? *PLoS One* **9**, e108522.

Tye, H. (2004). Application of statistical 'design of experiments' methods in drug discovery. *Drug Discov Today* **11**, 485-491.

Valdés, M., Pérez, N. O., Estrada-De Los Santos, P., Caballero-Mellado, J., Peña-Cabriales, J. J., Normand, P. & Hirsch, A. M. (2005). Non-*Frankia* actinomycetes isolated from surface-sterilized roots of *Casuarina equisetifolia* fix nitrogen. *Appl Environ Microbiol* **71**, 460–466.

Voeykova, T., Emelyanova, L., Tabakov, V. & Mkrtumyan, N. (1998). Transfer of plasmid pTO1 from *Escherichia coli* to various representatives of the order *Actinomycetales* by intergeneric conjugation. *FEMS Microbiol Lett* **162**, 47–52.

Vukov, N. & Vasiljevic, B. (1998). Analysis of plasmid pMZ1 from *Micromonospora zionensis*. *FEMS Microbiol Lett* **162**, 317–323.

Wagner, J. R., Mount, E. M., & Giles, H. F. (2014). Chapter 25 - Design of experiments. In *Extrus Defin Process Guid Handb*, Second Edi., pp. 291–308. Plastics Design Library.

Wang, Q., Xie, F., Tong, Y., Habisch, R., Yang, B., Zhang, L., Müller, R. & Fu, C. (2020). Dual-function chromogenic screening-based CRISPR/Cas9 genome editing system for actinomycetes. *Appl Microbiol Biotechnol* **104**, 225–239.

Weinstein, M. J., Luedemann, G. M., Oden, E. M. & Wagman, G. H. (1964). Everninomicin, a new antibiotic complex from *Micromonospora carbonacea*. *Antimicrob Agents Chemother* **10**, 24–32.

Weinstein, M. J., Wagman, G. H., Marquez, J. A., Testa, R. T., Oden, E. & Waitz, J. A. (1969). Megalomicin, a new macrolide antibiotic complex produced by *Micromonospora*. *J Antibiot (Tokyo)* **22**, 253–258.

Weinstein, M. J., Marquez, J. A., Testa, R. T., Wagman, G. H., Oden, E. M. & Waitz, J. A. (1970). Antibiotic 6640, a new *Micromonospora*-produced aminoglycoside antibiotic. *J Antibiot (Tokyo)* **23**, 551–554.

Weinstein, M. J., Luedemann, G. M., Oden, E. M. & Wagman, G. H. (1967). Halomicin, a new *Micromonospora*-produced antibiotic. *Antimicrob Agents Chemother* **7**, 435–41.

Weinstein, M. J., Wagman, G. H., Marquez, J. A., Testa, R. T. & Waitz, J. A. (1975). Verdamicin, a new broad spectrum aminoglycoside antibiotic. *Antimicrob Agents Chemother* **7**, 246–9.

Weinstein, M. J., Luedemann, G. M., Oden, E. M., Wagman, G. H., Rosselet, J. P., Marquez, J. A., Coniglio, C. T., Charney, W., Herzog, H. L. & Black, J. (1963). Gentamicin, 1 a new antibiotic complex from *Micromonospora*. *J Med Chem* **6**, 463–464.

Wilkinson, B. & Micklefield, J. (2007). Mining and engineering natural-product biosynthetic pathways. *Nat Chem Biol* **7**, 379-86.

Yildirim, S., Thompson, M. G., Jacobs, A. C., Zurawski, D. V & Kirkup, B. C. (2016). Evaluation of parameters for high efficiency transformation of *Acinetobacter baumannii*. *Sci Rep* **6**.

Yoichi, H., Michiko, M., Kazuo, S. & Haruo, S. (1993). Quinolidomicins A1, A2 and BL9 novel 60-membered macrolide antibiotics I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activity. *J Antibiot (Tokyo)* **46**, 1557–1562.

Zeaiter, Z., Mapelli, F., Crotti, E. & Borin, S. (2018). Methods for the genetic manipulation of marine bacteria. *Electron J Biotechnol* **33**, 17-28.

Zeng, H., Wen, S., Xu, W., He, Z., Zhai, G., Liu, Y., Deng, Z. & Sun, Y. (2015). Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl Microbiol Biotechnol* **99**, 10575–10585.

Zhang, G. Q., Bao, P., Zhang, Y., Deng, A. H., Chen, N. & Wen, T. Y. (2011a). Enhancing electro-transformation competency of recalcitrant *Bacillus amyloliquefaciens* by combining cell-wall weakening and cell-membrane fluidity disturbing. *Anal Biochem* **409**, 130–137.

Zhang, H., Li, Y., Chen, X., Sheng, H. & An, L. (2011b). Optimization of electroporation conditions for *Arthrobacter* with plasmid PART2. *J Microbiol Methods* **84**, 114–120.

Zhang, Z., Ding, Z.-T., Shu, D., Luo, D. & Tan, H. (2015). Development of an efficient electroporation method for iturin A-producing *Bacillus subtilis* ZK. *Int J Mol Sci* **16**, 7334–7351.

Chapter 6. Conclusions and future perspectives.

6.1 The future of the natural products field: challenges and upcoming developments.

As reviewed extensively in this thesis, the methodologies used to find new antibiotics during the golden age of antimicrobial drug discovery became obsolete at the beginning of the 2000s (da Cunha *et al*., 2019). It was then, when the first bacterial genomes started to be sequenced, that the field as we know it today started its second generation. During these last two decades, approaches previously unthinkable have led not only to the discovery of new farmaceutically interesting natural products, but most importantly, to the generation of an unpredecented amount of data yet to be exploited. While we constantly keep generating more data, both experimentally and computationally, huge efforts are being made to develop tools that make the data accesible, so that it can be read and used efficiently by anyone without requiring experience in bioinformatics.

These are some of the challenges that the microbial natural products research field will face over the next decade.

6.1.1 Genome and metagenome mining.

The availability of quick and affordable sequencing platforms has resulted in thousands of genomes and metagenomes being sequenced every year by the scientific community. These advancements have come with their own set of problems. Next, some examples are outlined.

First, the recent development of hundreds of different algorithms to perform all kind of analysis of genomic and metagenomic data calls for an urgent need for standardised pipelines that integrate in a single user-friendly software all the tools needed to carry out specific analysis (Kumar & Dudley, 2007). This will ease the automation of data analysis and make bioinformatic pipelines more accesible to non-experts.

Second, a more specific issue of the field, is that, although there currently is an existing number of tools and algorithms that facilitate the discovery of BGCs through homology search, there is an increasing demand for tools that identify distant protein homologs – allowing the discovery

of unidentified BGCs – to overcome the main disadvantage of genome mining, this is, that only known or highly similar BGCs can be identified (Albarano *et al*., 2020).

Finally, a challenge that will determine how fast the field moves forward is data sharing. For example, the creation of a unique completely open-access database that integrates all the bibliographic, theoretical and experimental data of all the characterised and uncharacterised natural products (Van Santen *et al*., 2019).

6.1.2 Heterologous production.

The use of heterologous hosts for the production of secondary metabolites has been exploited for decades, however, the field is still far from establishing a "universal" platform that allows the quick exploitation of all the newly discovered and uncharacterized BGCs present in nature (Ahmed *et al*., 2020). To achieve this, understanding how BGCs are silenced or activated is still a major challenge. Despite many efforts have focused on the identification of biosynthetic regulatory mechanisms in *Actinobacteria*, complex multilevel dynamic cascades yet to be comprehensively studied seem to control the production potential of heterologous hosts. Specifically, the cross-talk between global regulators and pleiotropic regulators, and pathway-specific regulators from the native host can completely disrupt the production of the compound of interest (McLean *et al*., 2019). Since the study of individual regulatory mechanisms for each pathway is very time consuming, a high-throughput platform for biosynthetic regulatory mechanism analysis would definitely make the field advance more efficiently.

Following those lines, another limitation that will need to be overcome is the automation of high-throughput experimental platforms to efficiently screen microbial strains (Leavell *et al*., 2020). Some of these are already available in well-funded companies and public research institutions, however, access is still very limited for small research groups. Automating high-throughput strain isolation, growth media and production media testing, antimicrobial assays, solid and liquid state fermentations, organic extractions and chromatographic analysis would facilitate the analysis of a wide range of settings for the biosynthesis of secondary metabolites (Welch, 2019). For instance, this would allow the discovery of conditions that activate the production of criptic or silent BGCs.

Another challenge to address in the future is the genetic compatibility of the genes within the BGC of the native host and the characteristics specific to the heterologous host (Yan *et al*., 2018). That includes not only the codon usage but also the recognition of native promoter and RBS regions by the machinery of the hosts (Zhao *et al*., 2019). Until today, issues of this type have been addressed through the use of phylogenetically close species and through the replacement of native promoters by strong constitutive promoters already tested in the heterologous host (Myronovskyi & Luzhetskyy, 2019). However, the current number of tractable heterologous hosts is very limited, thus, future efforts need to focus on the development of new *Actinobacteria* chassis amenable to genetic manipulation.

6.1.3 Increasing production yield.

One of the goals of the biosynthesis research field is to improve the biological synthesis of secondary metabolites so that it would not be necessary to recur to chemical synthesis for their commercial exploitation, greatly reducing energy consumption and the use of petroleum derived solvents. Nevertheless, most natural products are synthesised in very low amounts both in nature and under standard laboratory conditions. Sometimes the production yield is so low that after extraction and purification the compound is lost or not present in enough quantity to be detected by chromatography, leading researchers to believe that the correspoing BGC is silent (Beutler, 2009). This leads them to focus on activating the BGC rather than on increasing the production of the target compound.

There are mainly two milestones that need to be fixed during the next few years. First, it is well-known among actinomycetes researchers that the instability of many *Actinobacteria* strains – which sometimes can grow differently under the same conditions – makes secondary metabolite production not robust nor maintainable enough for industrial exploitation (Huo *et al*., 2019). Second, future basic research will need to focus on establishing some measurement standarisation guidelines to assess production (Zhao & Medema, 2016). Indeed, it would be beneficial to monitor secondary metabolite biosynthesis in real time, for example, through the development and use of biosensors based on the regulatory genes of the pathway of interest (Hossain *et al*., 2020).

6.1.4 Exploitation of non-model microbial species.

Finally, a challenge yet to be addressed – while the discovery of novel microbial species continues – is the genetic manipulation of non-model organisms. Every time an organism with potential applications in any area is identified, researchers attempt to devise how the organism can be genetically manipulated. First, through the introduction of foreign DNA inside the cell, and second, through the modification of the strain's genome.

Cell transformation generally occurs either by electroporation or by conjugation and allows to carry out a wide variety of studies on the organism. This method is, in most cases, species-specific and finding the adequate transformation conditions can be done by modifying parameters of a previous protocol developed for a different species or it may need to be established from scratch. Tools such as statistical DOE have already facilitated the development of transformation protocols for non-model microbial species (Marciset & Mollet, 1994). However, the tedious work required to generate large amounts of data to build good statistical models makes it unscalable and calls for a more powerful approach. For example, combining DOE with the use of liquid handling robots for the automation of all the transformation conditions would allow the high-throughput and scalable development of protocols for any strain with potential biosynthetic capabilities.

Currently, there are broad host range vectors and genome editing systems that can be used in a wide variety of organisms that promise time and cost efficient genome editing, however, results are not guaranteed in any species until tested. When none of the previously developed genome editing tools work, variations need to be made to address the potential limitations of those systems in the target organism. During the following years we expect, not only present and new genome editing tools being applied to more natural product biofactories, but also, the implementation of different strategies such as the creation of MoClo CRISPR libraries – using a range of promoters, RBSs, CRISPR genes, terminators and backbones – to find the most efficient construct for a chosen organism. These strategies are also excellent candidates to be automated.

6.2 This thesis in a nutshell.

This thesis was dedicated to the study of the abyssomicin family of natural products. More specifically, it aimed at developing some basic groundwork that could facilitate the discovery of new abyssomicins with potentially interesting bioactivities in the future. For that, three different approaches were followed. First, an investigation on the environmental distribution and evolution of the abyssomicin BGC was carried out. Then, research was carried out to produce abyssomicins in a heterologous host based on the abyssomicin BGC from *M. maris*. Finally, efforts focused on increasing abyssomicin production in *M. maris* and developing tools that would facilitate engineering this strain.

6.2.1 Environmental distribution and evolution of the abyssomicin BGC.

This thesis started exploring nature's reservoir to identify microbial strains harbouring potential abyssomicin BGCs. At the same time, this comprehensive analysis allowed to study the habitat distribution and evolutionary history of the abyssomicin BGC. First, the mining of 3,027 publicly available metagenomes revealed that microorganisms harbouring Diels-Alderase homologs have a diverse environmental distribution, although they are mainly present in soil and plant-associated microbiomes. Since the abyssomicin producers known to date were equally isolated from aquatic and soil ecosystems, the absence of Diels-Alderase positive metagenomes from aquatic environments was unexpected, although a latter analysis showed that metagenomes of aquatic origin have a lower sequencing depth and that the abundance of *Bacteria* and *Actinobacteria* is lower, which could reduce the chances of sequencing Diels-Alderase homologs. Then, when mining the genomes present in NCBI database 74 Diels-Alderase homologs present in 66 different genomes were identified, and from these, it was possible to identify and annotate five complete and 12 partial new abyssomicin BGCs and 23 new potential abyssomicin BGCs. Interestingly, the isolation location all these 40 potential abyssomicin producers followed the same habitat distribution patterns found through the metagenome mining.

Apart from the environmental distribution, this study revealed that abyssomicin BGCs are limited to the phylum *Actinobacteria* and are mainly present in strains from the genus *Streptomyces*, *Frankia* and *Micromonospora*. Moreover, the fact that abyssomicin BGCs are limited to specific strains suggests that HGT events may have played an important role in their transmission. Finally, five of the strains discovered to harbour potential abyssomicin BGCs

were obtained and preliminary abyssomicin production screenings were carried out. Since none of them appeared to produce abyssomicins, it is proposed that future work could focus on investigating at the detail the regulation of abyssomicin BGCs, which seems essential to understand the expression of these clusters. Another potential research line could focus on unravelling the biased habitat distribution of strains with abyssomicin BGCs. It would be specially interesting to elucidate the ecological role that these molecules play in nature, as novel applications outside the biomedical field could emerge.

6.2.2 Heterologous production of abyssomicins.

After exploring the abyssomicin production potential of various environmental isolates, working with the characterised *aby* BGC from *M. maris* seemed the best opportunity to try to produce new abyssomicins. In order to do so, given the lack of molecular microbiology methods adapted to use with *M. maris*, the approach of using heterologous hosts was pursued. For that, a literature search looking at bacterial chassis used for the heterologous production of polyketide natural products from *Actinobacteria* was carried out to identify the strains most frequently used for this purpose. With the objective of trying to maximise the chances of success, the genomes of the most promising candidate strains were used to determine the codon usage of each strain and compare it to the codon usage of the *aby* BGC. After that, the *aby* BGC was moved into *E. coli* BAP1, *S. lividans* K4-144, *S. albus* J1074, *S. coelicolor* M145 and M1152 and *Streptomyces* sp. FR-008 LQ3, although none of the strains produced abyssomicins under the fermentation conditions tested. To induce abyssomicin production in *S. lividans* and *S. albus*, ribosome engineering was applied, followed by the introduction of a strong constitutive promoter upstream the main activator in the *aby* BGC, AbyI. These strains were rescreened for abyssomicin production, however, even after correct promoter engineering, no abyssomicin production was observed. To understand whether the lack of production was due to failed transcription of the heterologous *aby* cluster, a gene expression analysis was carried out. This revealed active gene expression, however, in a later experiment in which the genome of *S. lividans* was re-sequenced several mutations and deletions were found in *abyB1*, indicating that translational errors in one of the polyketide synthetase genes were likely the cause of failed production.

This piece of work constitutes a clear example of the unpredictability of biological systems and the need to include checkpoints during projects at which complete analysis are carried out, with special emphasis on examining, if possible, the whole genome of engineered strains after

editing them. Interestingly, *E. coli* strains generally used for cloning are *recA*$^-$ to prevent recombination of the plasmid with the host chromosome, scrambling of the plasmid, and/or mixing of various plasmids. However, this is not the case of the strain most frequently used as donor strain in *Streptomyces* conjugations, *E. coli* ET12567. To overcome this limitation and avoid, for example, homologous recombination to happen while transferring large vectors containing sequences of repetitive nature such as PKSs, other mobilisation systems such as the helper plasmid pTA-Mob have been developed. In this case, pTA-Mob is generally co-transformed into a *recA*$^-$ *E. coli* strain such as *E. coli* Top10 together with the vector that needs to be transferred into the receiver strain (Strand *et al.*, 2014). That was the approach followed in this thesis and through which *S. lividans* K4-144 and *S. albus* J1074 were transformed with the integrative vector pESAC13_*aby*. In this specific case, re-sequencing the genomes of the other heterologous hosts as well as pESAC13_*aby* from the *E. coli* strains used to clone and mobilise it would allow to identify at which point these mutations happened.

Another possibility to take into account is the rearrangement of the cluster in the heterologous hosts as a defense mechanism or even the selection of deletion mutants due to the associated toxicity of abyssomicin C. Previous work demonstrated that atrop-abyssomicin C irreversibly binds to the Cys263 of ADCS in *B. subtilis*, and that *R. erythropolis* and *S. aureus*, with a serine at that position, are sensitive to atrop-abyssomicin C too, whereas *M. maris* and *M. luteus* with a glycine are resistant (Figure 6.1) (Riedlinger *et al.*, 2004). That indicates that *S. albus* (Ser) and *S. koyangensis* (Ser), that produces abyssomicins but not abyssomicin C, would also be sensitive to this compound. On the other side, *S. lividans* presents the same amino acid as *M. maris* (Gly) and thus should not be affected by the biosynthesis of abyssomicin C.

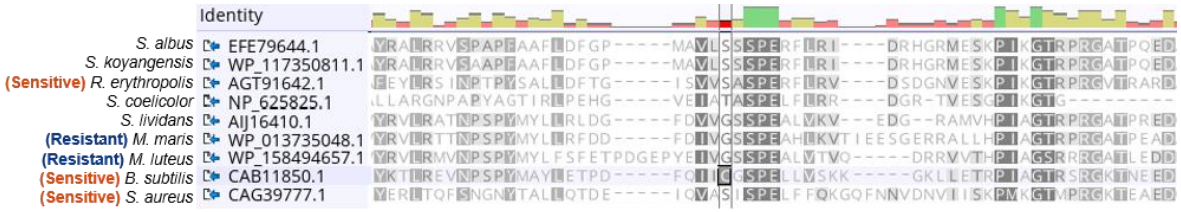

Figure 6.1. Partial amino acid sequences of the ADCS homologs in *S. albus*, *S. koyangensis*, *R. erythropolis*, *S. coelicolor*, *S. lividans*, *M. maris*, *M. luteus*, *S. subtilis* and *S. aureus*. Atrop-abyssomicin C irreversibly binds to the Cys263 of ADCS in *B. subtilis* and experimental data has shown before that *R. erythropolis* (Ser) and *S. aureus* (Ser) are sensitive to atrop-abyssomicin C, whereas *M. maris* (Gly) and *M. luteus* (Gly) are resistant.

Since in this work it was demonstrated that editing the *aby* BGC in *S. lividans* and *S. albus* is possible using the pCRISPomyces-2 system, future work could involve the repair of *abyB1* and

testing for abyssomicin production again. That would allow the project to continue towards the biosynthesis of other natural and non-natural abyssomicins by carrying out modifications in some of the PKS domains to accept different substrates (Figure 3.18A) and/or produce abyssomicins with various saturation/oxidation patterns (Figure 3.18B).

6.2.3 **M. maris** *as an abyssomicin production chassis.*

As very slow progress was being made through the heterologous expression approach, it was decided to start paving the way to use *M. maris* as a chassis to produce abyssomicins. For that, ribosome engineering was applied to select spontaneously developed drug-resistant mutants with enhanced abyssomicin production. This way it was possible to increase abyssomicin production in *M. maris* 3.4-fold, however, if at some point in the future this strain was to be used for industrial exploitation of the abyssomicins probably other approaches would need to be applied to further increase production. After that, considering that the only way of introducing DNA into *M. maris* is by conjugation, process that is very time consuming and inefficient, the creation of a tailored electroporation protocol was approached through the use of statistical DOE. That way, it was possible to optimise a protocol that yielded over 100 transformants per µg of DNA.

Future work could focus on testing and optimising current genome editing toolkits so that targeted genome edits could be made. Indeed, a *M. maris* quick editing toolkit could be made by combining various CRISPR systems and modular assembly approaches such as MoClo to assemble multiple promoters, ribosomal binding sites (RBSs), endonucleases such as SaCas9, SpCas9 and Cpf1, and terminators into backbones such as pCRISPomyces-2 or pQ-idgS to make hundreds of variants in one go. That whole library could then be directly electroporated into *M. maris*, preventing rearrangements, and transformants could be screened for the desired edits, which would also provide information of the most suitable combinations (Figure 6.2).

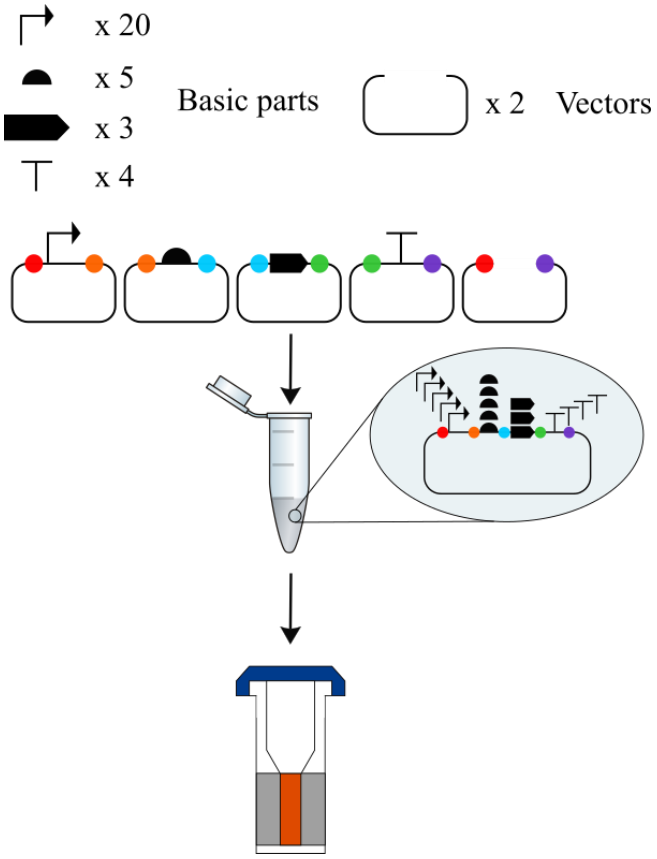

Figure 6.2. Schematic representation of a MoClo-like assembly of CRISPR vectors for genome editing *M. maris* through electroporation.

Although the abyssomicins are the main fermentation product of *M. maris*, future work could also include the introduction of a second copy of the *aby* BGC through transformation with the integrative pESAC13_*aby*, the deletion of other BGCs coupled to the introduction of more *attB* sites for multiple copies of the cluster (Li *et al.*, 2017), overexpression of the main pathway activator through cloning in an expression vector and transformation, building the strain's metabolic model and running flux balance analysis to make targeted genome edits that maximise abyssomicin production. Another approach that could be implemented could be the introduction of a promoter-less resistance gene in-frame after *abyB3*. This strain could then be subjected to UV - genotoxic agents and mutants selected based on the overexpression of the resistance gene (Qin *et al.*, 2017). As its expression would be reliant on the high expression of the *aby* BGC, increasing the antibiotic concentration through multiple rounds of exposure would end in the selection of mutants producing abyssomicins at a higher yield.

Overall, this work has explored three different approaches that will facilitate, in the future, the discovery of a plethora of novel natural and non-natural abyssomicins unknown to date.

6.3 References.

Ahmed, Y., Rebets, Y., Estévez, M. R., Zapp, J., Myronovskyi, M. & Luzhetskyy, A. (2020). Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb Cell Fact* **19**, 1–16.

Albarano, L., Esposito, R., Ruocco, N. & Costantini, M. (2020). Genome mining as new challenge in natural products discovery. *Mar Drugs* **18**, 1–17.

Beutler, J. A. (2009). Natural products as a foundation for drug discovery. *Curr Protoc Pharmacol* **46**, 1-21.

da Cunha, B. R., Fonseca, L. P. & Calado, C. R. C. (2019). Antibiotic discovery: Where have we come from, where do we go? *Antibiot (Tokyo)* **8**, 45.

Hossain, G. S., Saini, M., Miyake, R., Ling, H. & Chang, M. W. (2020). Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends Biotechnol* **38**, 797-810.

Huo, L., Hug, J. J., Fu, C., Bian, X., Zhang, Y. & Müller, R. (2019). Heterologous expression of bacterial natural product biosynthetic pathways. *Nat Prod Rep* **36**, 1412-1436.

Kumar, S. & Dudley, J. (2007). Bioinformatics software for biologists in the genomics era. *Bioinformatics* **23**, 1713-1717.

Leavell, M. D., Singh, A. H. & Kaufmann-Malaga, B. B. (2020). High-throughput screening for improved microbial cell factories, perspective and promise. *Curr Opin Biotechnol* **62**, 22-28.

Li, L., Zheng, G., Chen, J., Ge, M., Jiang, W. & Lu, Y. (2017). Multiplexed site-specific genome engineering for overproducing bioactive secondary metabolites in actinomycetes. Metab Eng **40**, 80–92.

Marciset, O. & Mollet, B. (1994). Multifactorial experimental design for optimizing transformation: Electroporation of *Streptococcus thermophilus*. *Biotechnol Bioeng* **43**, 490–496.

McLean, T. C., Wilkinson, B., Hutchings, M. I. & Devine, R. (2019). Dissolution of the disparate: Co-ordinate regulation in antibiotic biosynthesis. *Antibiot (Tokyo)* **8**, 83.

Myronovskyi, M. & Luzhetskyy, A. (2019). Heterologous production of small molecules in the optimized *Streptomyces* hosts. *Nat Prod Rep* **36**, 1281-1294.

Qin, R., Zhong, C., Zong, G., Fu, J., Pang, X. & Cao, G. (2017). Improvement of clavulanic acid production in *Streptomyces clavuligerus* F613-1 by using a claR-neo reporter strategy. *Electron J Biotechnol* **28**, 41–46.

Riedlinger, J., Reicke, A., Zähner, H., Krismer, B., Bull, A. T., Maldonado, L. A., Ward, A. C., Goodfellow, M., Bister, B. & other authors. **(2004).** Abyssomicins, inhibitors of the

para-aminobenzoic acid pathway produced by the marine *Verrucosispora* strain AB-18-032. *J Antibiot (Tokyo)* **57**, 271–279.

Strand, T. A., Lale, R., Degnes, K. F., Lando, M. & Valla, S. (2014). A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One* **9**, e90372.

Van Santen, J. A., Jacob, G., Singh, A. L., Aniebok, V., Balunas, M. J., Bunsko, D., Neto, F. C., Castaño-Espriu, L., Chang, C. & other authors. (2019). The Natural Products Atlas: An open access knowledge base for microbial natural products discovery. *ACS Cent Sci* **5**, 1824–1833.

Welch, C. J. (2019). High throughput analysis enables high throughput experimentation in pharmaceutical process research. *React Chem Eng* **4**, 1895-1911.

Yan, F., Burgard, C., Popoff, A., Zaburannyi, N., Zipf, G., Maier, J., Bernauer, H. S., Wenzel, S. C. & Müller, R. (2018). Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production. *Chem Sci* **9**, 7510–7519.

Zhao, H. & Medema, M. H. (2016). Standardization for natural product synthetic biology. *Nat Prod Rep* **33**, 920-924.

Zhao, M., Wang, S. L., Tao, X. Y., Zhao, G. L., Ren, Y. H., Wang, F. Q. & Wei, D. Z. (2019). Engineering diverse eubacteria promoters for robust gene expression in *Streptomyces lividans*. *J Biotechnol* **289**, 93–102.