Raul Vitor Ferreira de Oliveira

New molecular scaffolds for the development of antibiotics from the resistance exhibited by *Actinobacteria*

(Novas matrizes moleculares para o desenvolvimento de antibióticos a partir da resistência exibida por *Actinobactérias*)

Versão corrigida

São Paulo

2023

Raul Vitor Ferreira de Oliveira

New molecular scaffolds for the development of antibiotics from the resistance exhibited by *Actinobacteria*

(Novas matrizes moleculares para o desenvolvimento de antibióticos a partir da resistência exibida por *Actinobactérias*)

Dissertação apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Área de Concentração: Microbiologia

Orientador: Prof. Dr. Gabriel Padilla Maldonado

Versão corrigida

São Paulo

2023

CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e informação Biomédica
do Instituto de Ciências Biomédicas da Universidade de São Paulo

Ficha Catalográfica elaborada pelo(a) autor(a)

Ferreira de Oliveira, Raul Vitor
Novas matrizes moleculares para o desenvolvimento de antibióticos a partir da resistência exibida por Actinobactérias (New molecular scaffolds for the development of antibiotics from the resistance exhibited by Actinobacteria) / Raul Vitor Ferreira de Oliveira; orientador Gabriel Padilla Maldonado. -- São Paulo, 2023.
72 p.

Dissertação (Mestrado)) -- Universidade de São Paulo, Instituto de Ciências Biomédicas.

1. Resistência a antibióticos. 2. Mineração genômica. 3. Produtos Naturais. I. Maldonado, Gabriel Padilla, orientador. II. Título.

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Biomédicas
Secretaria de Pós-Graduação
Edifício Biomédicas III
Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415- CEP05508-900
Tel. (0xx11) 3091.7439 - Fax (0xx11) 3091.7423
E-mail: spgicb@usp.br

ICB

São Paulo,

Á
CPG do ICB/USP

Senhor(a) Presidente(a)

Encaminhamos a V.Sa. 01 (um) exemplar impresso, páginas 1 a 72, da Dissertação (x) Tese de Doutorado ( ), concluída no Programa de Pós-Graduação em Microbiologia do(a) aluno(a) **Raul Vitor Ferreira de Oliveira** ao qual foram incorporadas as correções sugeridas pelos membros da Comissão Julgadora.

Atenciosamente,

Gabriel Padilla

Orientador(a)

Raul Vitor Ferreira de Oliveira

Aluno(a)

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Biomédicas
Secretaria de Pós-Graduação
Edifício Biomédicas III
Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415- CEP05508-900
Tel. (0xx11) 3091.7439 - Fax (0xx11) 3091.7423
E-mail: spgicb@usp.br

ICB

Formulário/Declaração de Anuência do Orientador em relação ao Depósito da Dissertação/Tese

Nome do **Programa: Programa de Pós Graduação em Microbiologia**

Nome do **Orientador(a): Gabriel Padilla Maldonado**

Nome do **Aluno(a): Raul Vitor Ferreira de Oliveira**

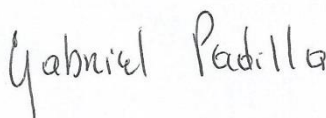
Nível: **(x) Mestrado ( ) Doutorado**

Titulo da Dissertação / Tese: New molecular scaffolds for the development of antibiotics from the resistance exhibited by Actinobacteria (Novas matrizes moleculares para o desenvolvimento de antibióticos a partir da resistência exibida por Actinobactérias)

D E C L A R A Ç Ã O
==================

De acordo com o Artigo 89 do Regimento da Pós-Graduação da Universidade de São Paulo, declaro minha anuência e autorizo o(a) aluno(a) Raul Vitor Ferreira de Oliveira a depositar na Seção de Pós-Graduação do ICB/USP, localizado no Edifício Biomédicas III, os exemplares da Dissertação/Tese sob o título mencionado nesta Declaração.

São Paulo, 13 de abril de 2023



____
Assinatura do Orientador

Cidade Universitária "Armando de Salles Oliveira", Butantã, São Paulo, SP - Av. Professor Lineu Prestes, 2415 - ICB III - 05508 000
Comissão de Ética em Pesquisa - Telefone (11) 3091-7733 - e-mail cep@icb.usp.br

CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB nº **1071/2019** referente ao projeto intitulado: **"Novas matrizes (scaffolds) moleculares para o desenvolvimento de antibióticos a partir da resistência exibida por Actinobactérias da coleção do ICB-USP"** sob a responsabilidade de ***Raul Vitor Ferreira de Oliveira*** e orientação do(a) Prof.(a) Dr.(a) ***Gabriel Padilla Maldonado,*** do Departamento de Microbiologia, foi analisado pela **CEUA** - Comissão de Ética no Uso de Animais e pelo **CEPSH** – Comitê de Ética em Pesquisa com Seres Humanos, tendo sido deliberado que o referido projeto não utilizará animais que estejam sob a égide da Lei nº 11.794, de 8 de outubro de 2008, nem envolverá procedimentos regulados pela Resolução CONEP nº 466/2012.

São Paulo, 18 de novembro de 2019.

Luciane Valéria Sita
Profa. Dra. **Luciane Valéria Sita**
Coordenadora CEUA ICB/USP

Profa. Dra. **Camila Squarzoni Dale**
Coordenadora CEPSH ICB/USP

Sumary

Acknowlegdments 3

Resumo 5

Abstract 6

Introduction.... 7

About the Golden Age of Antibiotics.... 7

The End of the Golden Age and “What Can We Do?”.... 8

1- Literature review 10

1.1 Actinobacteria and the genus Streptomyces 10

1.2 Bioprospecting of new antimicrobials in actinobacteria 11

1.3 Glycopeptides and Ansamycins 13

1.4 The resistance of the producer of glycopeptides and ansamycins 15

1.5 Genomic mining applied to the discovery of natural products 17

2. Objectives.... 18

3. Material and methods 19

3.1 Biological material 19

3.2 Resistance screening 19

3.3 Molecular screening 20

3.4 Biological activity assays 22

3.5 Genome sequencing and assembly 23

3.6 Genome analysis 24

4 Results and discussion 25

4.1 Screening for resistant bacteria.... 25

4.2 PCRs of resistance and biosynthetic genes 26

4.3 Antimicrobial activity assays 26

4.4 Genome Assembly 28

4.5 Phylogeny and second assembly 28

4.6 General characteristics of the genome 30
4.7 The secondary metabolism of BRB040 31
4.7.1 Overview 31
4.7.2 The BRB040 plasmids 36
4.7.3 Region 11: Nibomycin like 38
4.7.4 Region 14: Nenestatin like? 40
Conclusions 45
References 46
Appendix 1: Culture media used in this work 58
Appendix 2: Annotation and genomic sequence of the minor plasmid found in BRB040. 60
Appendix 3: Biosynthetic gene cluster found on the giant BRB040 plasmid (region 10) on the chromosome. 66
Appendix 4: Biosynthetic gene cluster from region 11 of BRB040 69
Appendix 5: Biosynthetic gene cluster from region 14 of BRB040 71

Acknowlegdments

O presente trabalho não seria possível sem uma rede de colaboradores que ofereceram suporte durante essa jornada de estudos e trabalho ao longo do caos que foram os últimos anos.

Um agradecimento especial a minha família, sem a qual eu jamais teria chegado tão longe.

Obrigado também a todos os meus amigos queridos, que proporcionaram momentos únicos e que fazem valer apena viver em São Paulo.

Obrigado a toda a equipe do laboratório 164, que acompanhou toda essa aventura de conhecimento oferecendo ajuda e insights importantes. Destaco aqui minha excepcional gratidão ao meu orientador, professor Dr. Gabriel Padilla Maldonado, que teve a paciência de me conduzir por esse complexo mundo das *Streptomyces*, com doses de café, boa conversa e muito conhecimento e experiência.

Não poderia deixar de mencionar Gisele, uma das pessoas mais sensacionais dessa pós-graduação em microbiologia.

E por fim, obrigado a CAPES, pela bolsa que tornou possível sobreviver e fazer essa pós-graduação.

I also dedicate here my sincere thanks to the SaMBio double degree program for the opportunity to carry out part of my master's studies in Germany, to the DAAD for the scholarship received and to all the incredible people I met during this period. Special thanks also to Professor Bodo Philip for having me in his laboratory and for his valuable teachings on scientific research.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

"*Anyway, he got so involved in reading that he spent his nights reading until it was light and the days until it was dark; and so, by sleeping little and reading a lot, he dried up his brain so that he came to lose his mind...*"

Miguel de Cervantes - Don Quixote

Resumo

A resistência aos antibióticos é um problema mundial que provocou em 2019 cerca de 1,27 milhões de mortes, mais do que a AIDS ou a malária. Apesar dessa ameaça a saúde global a adoção de novos antibióticos na clínica permanece estagnada desde o fim da Era de Ouro dos Antibióticos no final dos anos 70. Considerando essa crescente ameaça, torna-se urgente a descoberta de novas moléculas bioativas para tratar o número crescente de infecções por bactérias multirresistentes. Para isso pesquisas se voltam novamente para o estudo do metabolismo secundário de bactérias, especialmente as do filo Actinobactéria. O acesso a novas técnicas de sequenciamento e ferramentas de análise genômica, contribui bastante para acelerar a descoberta de novas moléculas a partir da informação genômica contida no DNA das bactérias. O presente trabalho buscou explorar a vasta riqueza da biodiversidade do ambiente marinho do litoral brasileiro para triar organismos com novos grupamentos de genes biossintéticos, capazes de produzir antibióticos de interesse clinico das classes dos glicopeptídeos e ansamicinas. Para tal foram utilizadas duas estratégias para triagem: a resistência do produtor a classe de antibiótico buscada e a verificação de atividade antimicrobiana contra patógenos clinicamente relevantes. Foram triadas 650 bactérias cultivadas em meio ISP2 com a adição dos antibióticos vancomicina (10µg/ml), teicoplanina (10µg/ml) e rifampicina (40µg/ml). Após essa etapa buscou-se identificar a presença dos genes fundamentais de biossíntese e resistência para esses antibióticos nas bactérias resistentes. Foram obtidas amplificações positivas para as cepas BRB014, BRB040 e BRB042 para o gene *rifK*, central para biossíntese de ansamicinas. Juntamente com os resultados de atividade antimicrobiana optou-se por sequenciar o genoma de BRB040. A partir da análise dos dados genômicos de BRB 040 foram encontradas evidências de que dois agrupamentos de genes de biossíntese ainda não descritos e candidatos a serem produtores de novas moléculas antimicrobianas.

Palavras chave: Resistência a antibióticos, Mineração Genômica, Produtos Naturais

Abstract
Antibiotic resistance is a worldwide problem that caused around 1.27 million deaths in 2019, more than AIDS or malaria. Despite this threat to global health, the adoption of new antibiotics in the clinic has remained stagnant since the end of the Golden Age of Antibiotics in the late 1970s. Considering this growing threat, it is urgent to discover new bioactive molecules to treat the growing number of multidrug-resistant bacterial infections. For this, research has turned again to the study of the secondary metabolism of bacteria, especially those of the phylum Actinobacteria. Access to new sequencing techniques and genomic analysis tools has greatly contributed to accelerating the discovery of new molecules from the genomic information contained in the DNA of bacteria. The present work sought to explore the vast biodiversity richness of the marine environment of the Brazilian coast to screen organisms with new groups of biosynthetic genes capable of producing antibiotics of clinical interest from the glycopeptides and ansamycins classes. For this, two screening strategies were used: the producer's resistance to the antibiotic class sought and the verification of antimicrobial activity against clinically relevant pathogens. 650 bacteria grown in ISP2 medium with the addition of the antibiotics vancomycin (10µg/ml), teicoplanin (10µg/ml) and rifampicin (40µg/ml) were screened. After this step, we sought to identify the presence of fundamental biosynthesis and resistance genes for these antibiotics in resistant bacteria. Positive amplifications were obtained for strains BRB014, BRB040 and BRB042 for the *rifK* gene, central to ansamycin biosynthesis. From the analysis of the genomic data of BRB 040, evidence was found that two clusters of biosynthetic genes not yet described and candidates to be producers of new antimicrobial molecules.

Key words: Antibiotic resistance, Genome mining, Natural products

Introduction

About the Golden Age of Antibiotics

Antibiotics are one of the most successful forms of chemotherapy ever discovered by mankind. In this context, research led by Paul Ehrlich and Alexander Fleming at the beginning of the 20th century established the bases for the search of bioactive compounds against pathogenic organisms (AMINOV, 2010). Ehrlich, in 1910, during his studies to discover a treatment for syphilis, a sexually transmitted disease caused by the bacterium *Treponema pallidium*, argued that chemical compounds capable of acting selectively against the target organism could be synthesized. In partnership with the bacteriologist Sahachiro Hata and the chemist Alfred Bertheim, hundreds of Atoxyl organoarsenic derivatives were synthesized and tested in rabbits infected with *Treponema pallidium*. Compound number 606 was able to cure the infected rabbits, but despite its promise as a possible treatment for syphilis, human trials have not progressed. (EHRLICH; HATA, 1910; VALENT et al., 2016). Despite the failure to establish this new compound as a treatment for syphilis, Ehrlich's work laid the groundwork for systematic compound screening, critical for drug research in the pharmaceutical industry.

Alexander Fleming's fortuitous sighting on September 3, 1928 (FLEMING, 1929) is undoubtedly the great milestone in the beginning of the so-called Golden Age of Antibiotics. Although the antibacterial properties of the *Penicillium* fungus had been known since antiquity, it was Fleming who was the first to suggest that the compound produced could be isolated and purified, which happened in 1940 in a study led by Howard Florey and Ernest Chain, who were able to develop an efficient protocol to produce sufficient amounts of penicillin to carry out clinical tests and later the massive use of this antibiotic in 1945 (CHAIN et al., 2005).

Another fundamental contribution in the search for bioactive molecules was the work by Selman Waksman in the 1930s, who began the systematic search for antimicrobials from microorganisms found in the soil. Waksman defined an antibiotic as “a compound made by a microbe to destroy other microbes” and was instrumental in identifying soil-dwelling filamentous *Actinomycetales* (*'actinomycetes'*) as abundant producers of antimicrobial

compounds, having discovered neomycin from these organisms, and streptomycin, the first active agent against tuberculosis (WAKSMAN; SCHATZ; REYNOLDS, 2010). This pioneering work identified in the genus *Streptomyces* a great source of natural products that, in addition to having antibacterial properties, are also active against fungi, viruses, nematodes, insects and can act as anticancer drugs and immunosuppressant.

Between the 1950s and 1970s, what is known as the Golden Age of Antibiotics took place. At that time, universities and pharmaceutical companies dedicated their efforts to the search for new bioactive molecules. During this period, most of the best known classes of antibiotics were discovered and characterized, which remain in clinical use until today; such as glycopeptides, tetracyclines, macrolides and other (OVERBYE; BARRETT, 2005). Thanks to the advent of vaccines and the discovery of antibiotics, which made other advances in medicine possible, there was a considerable increase in life expectancy, where only in the United States an increase was observed from 47 years in 1900 to 74 years for men and for 80 years for women in the year 2000 (MARTENS; DEMAIN, 2017).

The End of the Golden Age and “What Can We Do?”

Antibiotics were victims of their own success. In the 1940s, even before its mass use, it was observed that some bacteria were capable of destroying penicillin through enzymatic degradation (ABRAHAM; CHAIN, 1940).

Every antibiotic introduced in the clinic has a limited useful life, as it selects bacteria that have some intrinsic resistance mechanism that over generations and continuous use of the antibiotic become the most abundant in relation to antimicrobial-sensitive bacteria (WALSH, 2003).

Among the multiple factors that lead to the emergence of antibiotic-resistant bacteria are the frequent prescription of these drugs, the use of sub-therapeutic doses and the indiscriminate use of antibiotics in agriculture (CHANG et al., 2015; PALUMBI, 2001).

Another relevant fact is that the number of new classes of antibiotics discovered available in the clinic has stagnated. From the 1970s, the pharmaceutical industry began to reduce the priority in research and development of new antibiotics, many even withdrew completely from this field (OVERBYE; BARRETT, 2005).

Antibiotic resistance has become a major public health problem globally. In 2019 it is estimated that 4.95 million people died from diseases in which antibiotic resistance played a key role, and of these 1.27 million deaths were a direct result of infections with multidrug resistant bacteria - which means higher mortality than HIV/AIDS (864,000 deaths) or malaria (643,000 deaths) (MURRAY et al., 2022).

In 2017, the WHO published a list of antibiotic-resistant "priority pathogens" for the development of new antibiotics. There are 12 families of bacteria that represent the greatest threat to human health and must be prioritized in the development of new drugs (WHO, 2017). These pathogens are known by the acronym ESKAPE, opportunistic pathogens that have resistance to multiple drugs and great virulence, they are: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp. (MULANI et al., 2019).

Given this scenario, the search for antibiotics from bacterial natural products remains a promising source of new bioactive compounds for the treatment of infections caused by bacteria resistant to the antibiotics currently used in medicine. Searching for new antimicrobial compounds in extreme environments such as the seabed, deserts, polar and volcanic regions, where organisms are subjected to enormous evolutionary pressures, proves to be an exciting possibility in the search for new molecules, even more so when aligned with new sequencing techniques and genomic analyzes we were able to access a large number of biosynthetic gene clusters previously inaccessible by conventional techniques (ALBARANO et al., 2020; GIORDANO, 2020; SIVALINGAM et al., 2019).

Taking into account all the concepts mentioned above; the present work sought to identify possible new producers of bioactive molecules in the marine

microorganisms collection of the Laboratory of Bioproducts of the Institute of Biomedical Sciences at the University of São Paulo. For this, the producer's self-resistance principle was used, together with the presence of biosynthesis genes and bioinformatic analysis tools to identify groups of biosynthetic genes capable of producing molecules with antimicrobial potential.

1- Literature review

1.1 Actinobacteria and the genus Streptomyces

Actinobacteria form a defined clade of high-GC bacteria (72%), whose members can be found in a wide range of ecological niches. Some are important human and animal pathogens, such as some species of the genera *Mycobacterium* and *Nocardia,* while others are important sources of natural products as exemplified by the genera *Streptomyces* and *Micromonospora* (UL-HASSAN; WELLINGTON, 2009).

Many of the *Actinomycetes* morphologically resemble fungi because of their elongated cells that branch into filaments or hyphae. These hyphae can be distinguished from fungal hyphae based on size, as *Actinomycete* hyphae are much smaller than fungal hyphae and without internal fragmentation (PEPPER; GENTRY, 2015). Another notable feature of some genera of actinobacteria is the ability to reproduce by spores (KALAKOUTSKIIP; NINA AGRE, 1976).

Within the phylum *Actinobacteria*, the genus *Streptomyces* stands out. Initially proposed in 1943, this genus currently has 706 species listed (PARTE et al., 2020; WAKSMAN; HENRICI, 1943). These Gram-positive aerobic bacteria form aerial mycelium and have a complex life cycle, colonizing soils, waters, underwater sediments, plants, animals and extreme environments (DONALD et al., 2022).

The life cycle of *Streptomyces* consists of the following stages (BARKA et al., 2016; CHATER, 2013; UREM et al., 2016):

1- The spores germinate and the bacterium grows with the morphology of branched vegetative mycelia consisting of long multinucleoid hyphae;

2- When nutritional depletion occurs, the vegetative mycelium enters a phase of programmed cell death, providing nutrients for the formation of aerial mycelium. It is at this stage that the production and release of antibiotics occurs;

3- The newly formed aerial mycelium begins to produce chains of spores that will be released into the environment, restarting the cycle.

Streptomyces genomes have unique and interesting characteristics among bacterial genomes, in addition to the high GC content, they have highly compartmentalized linear chromosomes of 6 to 12 Mb, with a highly conserved central region and highly variable ends (HOFF et al., 2018). A very intriguing phenomenon that has been known for decades is the high genetic instability of this genus. Thus, during cultivation, the offspring of some *Streptomyces* species present between 0.1 and 1% of spontaneous mutants defective in morphological and physiological differentiation, including secondary metabolic pathways. (ZHOU et al., 2012).

The first complete *Streptomyces* genome was published in 2002 and at the time had the largest number of genes described for a bacterium. More than 20 clusters of genes encoding secondary metabolites were identified, several of them not yet characterized and recognized for their potential for various biotechnological applications (BENTLEY et al., 2002). This proposition opposed the thesis that a bacterium could synthesize only one natural product, demonstrating that the same species has the capacity to synthesize dozens of secondary metabolites, opposing the idea of exhaustion in the screening of new bacterial natural products.

1.2 Bioprospecting of new antimicrobials in actinobacteria

The antagonism platform, exploring the competitive relationships between soil microorganisms, was the basis of the agar overlay method developed by Selman Waksman and was employed for screening of more than 10,000 bacterial strains (WOODRUFF, 2014), discovering compounds like; actinomycin, streptotricin, fumigacin and clavacine, but only in 1944 it was

discovered a *Streptomyces griseus* strain which produced a non-toxic aminoglycoside antibiotic called streptomycin, an inhibitor of protein synthesis by binding to the 30S subunit of the bacterial ribosome (DA CUNHA; FONSECA; CALADO, 2019).

In 1947 chloramphenicol was isolated from *Streptomyces venezuelae*, developing a new class of molecules, the amphenicols. The antimicrobial activity of chloramphenicol derives from its reversible binding to the 50S ribosomal subunit, capable of inhibiting bacterial protein synthesis (EHRLICH et al., 1885). It was the first broad-spectrum antibiotic approved by the FDA as early as 1949 (WIEST; COCHRAN; TECKLENBURG, 2012). In 1948, chlorotetracycline introduced the class of tetracyclines, which also interrupt protein synthesis by acting on the 30S subunit of the ribosome, it was discovered from *Streptomyces aureofaciens* (LIU; MYERS, 2016).

The second most prescribed class of antibiotics, the macrolides, were introduced in 1949 with the discovery of erythromycin produced by *Saccharopolyspora erythrea*, this antibiotic also acts by binding to the 50s unit of the bacterial ribosome (CYPHERT et al., 2017).

In 1956, studies carried out using *Amycolatopsis orientalis* led to the discovery of vancomycin, creating a new class of molecules, the glycopeptides. Unlike previous molecules that acted to inhibit protein synthesis, glycopeptides interfere with the transpeptidation and transglycosylation stages of cell wall synthesis, thus inhibiting crosslinking and maturation (JAMES et al., 2012).

Also differing from protein synthesis inhibitors, ansamycins function by inhibiting prokaryotic DNA-dependent RNA polymerase. Rifamycin B was isolated in 1959 from *Streptomyces mediterranei* (reclassified as *Amycolatopsis mediterranei*), and despite its considerably low antimicrobial effect, this class affects a single metabolic target in bacteria, which made these molecules interesting for further modifications capable of improving their performance efficiency (FLOSS; YU, 2005).

After the period of stagnation in the discovery of new antibiotics, now in the 21st century in the face of the problems challenged by the growing bacterial resistance, and counting on the advances in sequencing techniques and genome

analysis, we have returned to the point of reestablishing the search for bioactive compounds from bacterial sources, among which the genus *Streptomyces* stands out due to its genetic capacity to produce an average of 30 secondary metabolites per strain (CHEVRETTE et al., 2019). Part of the challenge in exploring this vast wealth of natural products is that many of the clusters found are cryptic or silent, making the identification of these products through conventional means a challenge (HOSKISSON; SEIPKE, 2020).

The wide availability of information in public databases also revealed that the metabolic repertoire of *Streptomyces* is strain-specific and may vary between isolates of the same species (ANTONY-BABU et al., 2017; SOTTORFF et al., 2019).

One of the main challenges in the search for new antibiotic compounds is the number of samples to be screened both by conventional methods and by metagenomic approaches. In conventional phenotypic screening, it is estimated that screening of 10^5 strains is required to find a producer of a new antibiotic. For a screening based on a metagenomic library, screening is on the order of 10^7 clones to identify a producer (THAKER et al., 2013).

To overcome this great challenge in terms of time and resources, researchers adopted a hybrid approach based on the principle that the producer organism be resistant to the class of antibiotics it produces, selecting possible producers based on the resistance of these classes of interest for the study. (THAKER; WAGLECHNER; WRIGHT, 2014).

1.3 Glycopeptides and Ansamycins

Glycopeptides are drugs of last choice in the treatment of antibiotic-resistant Gram-positive bacterial infections such as methicillin-resistant *Staphylococcus aureus* and *Clostridium difficile* (BUTLER et al., 2014). Its structure is based on a conserved central heptapeptide domain in which five or seven amino acid residues common in all molecules of this class are coupled. (REYNOLDS, 1989). (Figure 1)

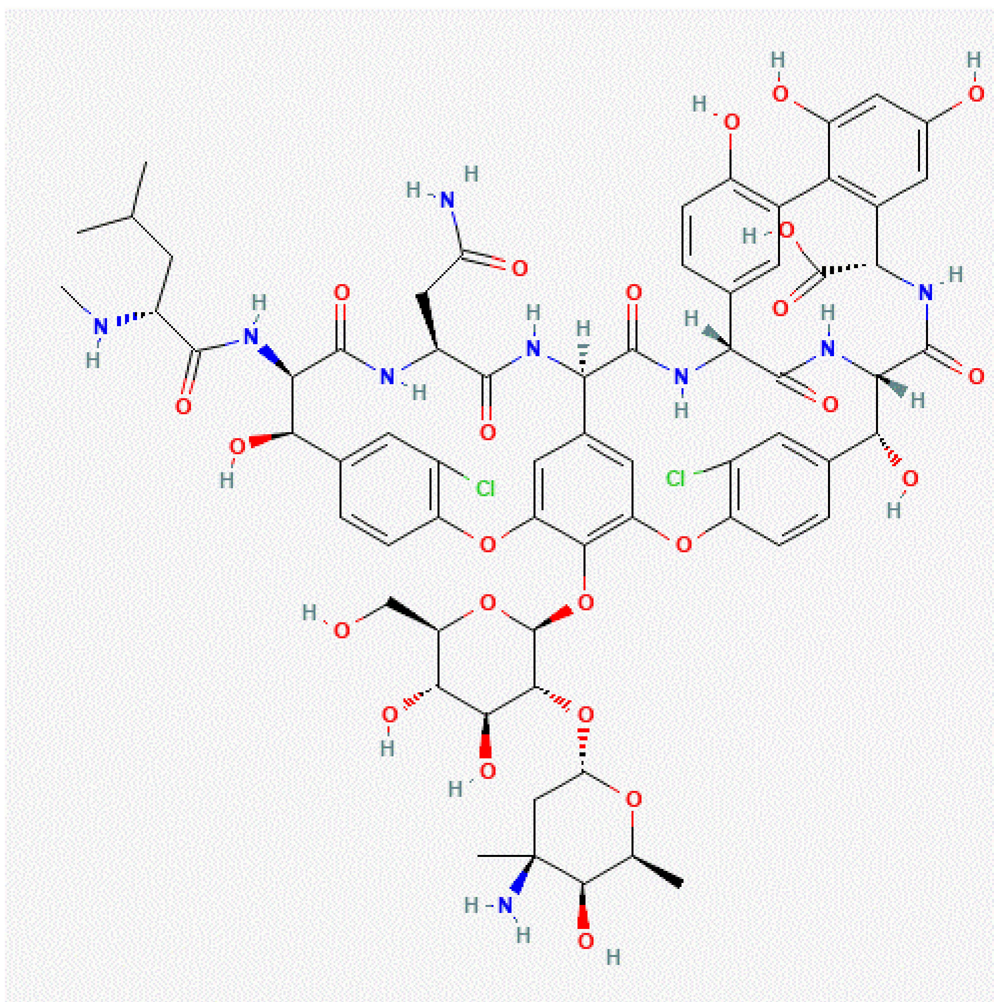


Figure 1. Chemical structure of vancomycin, first discovered antibiotic of the glycopeptide class, produced by *Amycolatopsis orientalis* ("Vancomycin | C66H75Cl2N9O24 - PubChem", [s.d.])

The synthesis of these molecules is complex, but can be divided into three basic stages; Firstly, the primary metabolism of the bacteria provides the necessary substrates for the construction of the molecule (metabolism of phenylalanine, tyrosine and tryptophan / metabolism of pyruvate / metabolism of amino sugars). Secondly, linking of these building blocks, interconnection of aromatic side chains and halogenation takes place, all accomplished by a multi-enzyme complex. In the third stage we have the modification of the heptapeptide by glycosyltransferases, a methyltransferase and in the case of type IV glycopeptides, such as teicoplanin, we have the action of acyltransferases and sulfotransferases (LI et al., 2004; STEGMANN; FRASCH; WOHLLEBEN, 2010).

Rifamycins belong to the ansamycin group and are characterized by a cyclic structure consisting of an aromatic part (naphthalene, naphthoquinone, benzene or benzoquinone rings) (BURY-MONÉ, 2014). Notably its derivatives, which have a naphthohydroquinone or naphthoquinone structure, are potent antibiotics against mycobacteria and are one of the most widely used drugs against tuberculosis (GROBBELAAR et al., 2019) (Figure 2)

This class of molecules is synthesized by PKS type 1 enzymes having as main precursor the AHBA synthase enzyme, in *Amycolatopsis mediterranei* synthesized by the *rifK* gene (AUGUST et al., 1998; FLOSS; YU; ARAKAWA, 2010). This gene serves as a useful tool in genetic screening for new ansamycins and other AHBA-derived natural products (HE et al., 2006; PEEK et al., 2018; WANG et al., 2013).

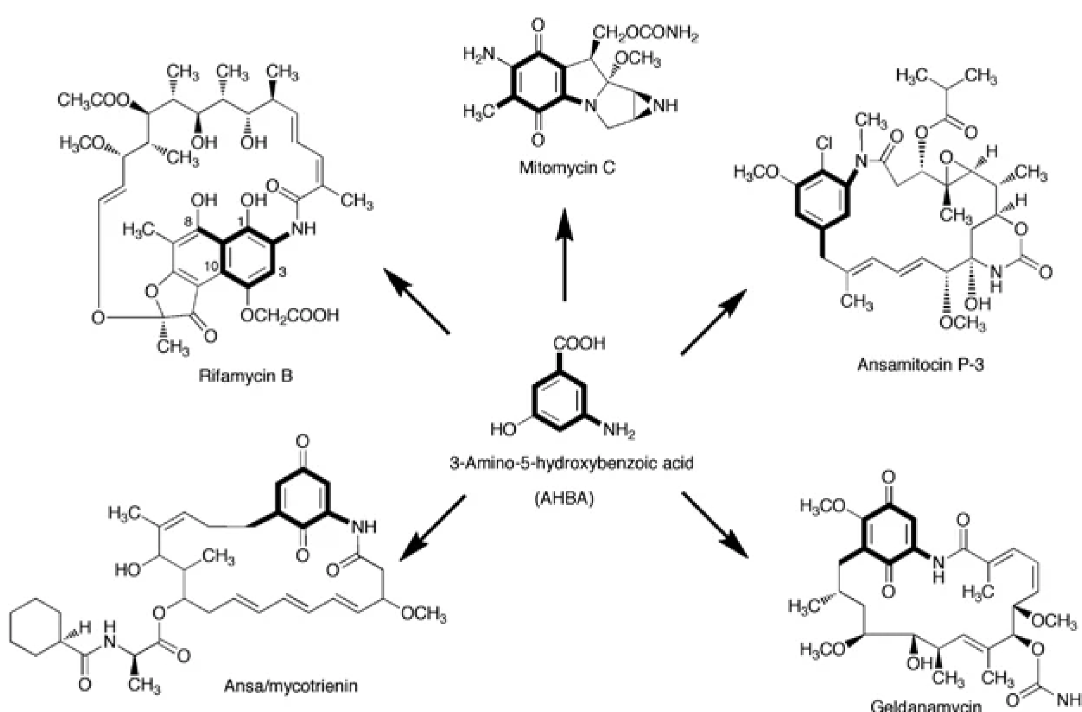


Figure 2. Rifamycin and others compounds with 3-amino-5-hydroxybenzoic acid (AHBA) as precursor (FLOSS; YU; ARAKAWA, 2010).

1.4 The resistance of the producer of glycopeptides and ansamycins

Antibiotic-producing organisms need to be protected from the lethal substances synthesized by themselves. This inherent ability of producing

organisms is called self-resistance (SUGIYAMA, 2015). In order to be able to effectively resist its own antibiotic product, the bacterium needs to have actives during the synthesis of these molecules; antibiotic-modifying and target-protecting enzymes or an antibiotic excretion system must be present in the producing organisms.

In the case of glycopeptides, both producing organisms and pathogens use very similar mechanisms to tolerate this class of antibiotics (BINDA; MARINELLI; MARCONE, 2014; COURVALIN, 2006). Resistant organisms change the target pentapeptide in Lipid II to end in D -Ala- D -Lactate (D -Lac) instead of the canonical D -Ala-D -Ala. This change turns an amide bond into an ester, leaving one less hydrogen bond for the glycopeptide, reducing its affinity for the binding site (BUGG et al., 1991; MCCOMAS; CROWLEY; BOGER, 2003).

In producing organisms this machinery of resistance is mediated by a set of fundamental core genes, two regulatory genes and some genes that contribute to the mechanism. The core resistance genes encode three enzymes that modify the target and confer resistance. *VanH* is a D-Lac dehydrogenase that reduces pyruvate to the necessary set of D-Lac precursors. *VanA* is a D-Ala-D-Lac ligase, a modified form of the traditional D-Ala-D-Ala ligase (Ddl), with a preference for D-Lac which forms the depsipeptide. *VanX* is a dipeptidase that selectively cleaves D-Ala-D-Ala, but does not recognize the modified depsipeptide as a substrate. With the exception of one, the *VanHAX* gene cluster has been found in all glycopeptide producing organisms (YIM et al., 2013). (Figure 3).

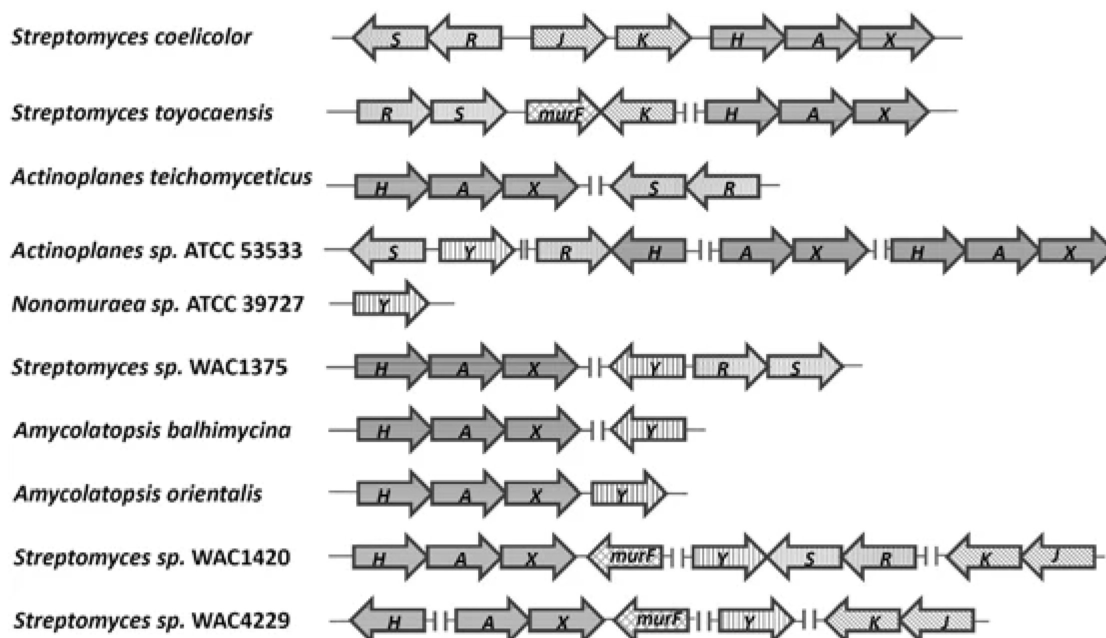


Figure 3. *Van*HAX cluster genes in different glycopeptide producing organisms (YIM et al., 2013).

As for ansamycins, producing organisms seem to adopt as main mechanisms for resistance an efflux system encoded by the *rifP* gene found *in Amycolatopsis mediterranei* (ABSALÓN et al., 2007) used to export the antibiotic produced inside the cell to the extracellular medium. It was reported that mutants with partial deletion of the *rifP* gene were unable to export the antibiotic to the extracellular medium.

Other strategies used by potential producers involve the production of glycosyltransferases capable of inactivating rifampicins (SPANOGIANNOPOULOS et al., 2012) and mechanisms to disassociate the antibiotic from RNA polymerase, nullifying its activity (SUDZINOVÁ et al., [s.d.]).

1.5 Genomic mining applied to the discovery of natural products

The exploration of genomic sequences in search of useful information for the discovery of new natural products is called “genomic mining” (CHALLIS, 2008). The availability of thousands of bacterial genomes and metagenomes on public platforms allows access to a massive amount of data and a source for discovery of new molecules (FOULSTON, 2019).

The genes involved in the biosynthesis, processing and transport of secondary metabolites are found physically grouped in the chromosome of prokaryotes (CIMERMANCIC et al., 2014). These gene clusters can be tracked and mapped using automated tools that work guided by identifying core biosynthetic genes, which are often highly conserved across different species. (ZIEMERT; ALANJARY; WEBER, 2016).

Genomic mining helps overcome challenges encountered in conventional screening, such as dependence on culture conditions, metabolite extraction techniques, and gene expression levels (BELKNAP et al., 2020; BIERMANN; HELFRICH, 2021), and aims to find new scaffolds with clusters of genes responsible for the production of new natural products.

2. Objectives

- Screen the BRB collection of marine actinobacteria to find possible new producers of glycopeptides and ansamycins.
- Identify new scaffolds of biosynthetic genes in the genome of selected organisms that may encode unpublished products of pharmacological interest.

3. Material and methods

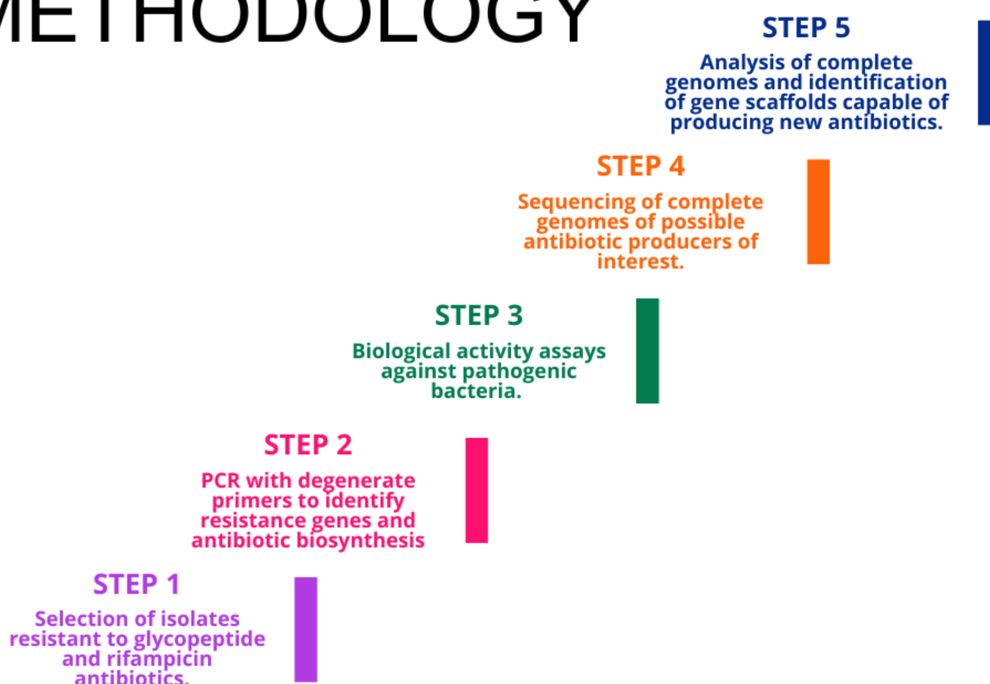


Figure 4. Overview of the methodology used in this work.

3.1 Biological material

The BRB collection used in the present study was obtained from collections of sediments and marine organisms along the Brazilian coast. The isolated organisms were preserved in 80% glycerol and stored in a freezer at – 80°C.

3.2 Resistance screening

Resistance screening followed the protocol developed by THAKER; WAGLECHNER; WRIGHT, 2014 with adaptations. Bacteria from the collection were grown in 6-well plates containing 3mL of ISP2 culture medium. The antibiotic concentration for the resistance filter was taken as being five times the highest value of the minimum inhibitory concentration within five different genera

of actinobacteria for the antibiotics: vancomycin, teicoplanin and rifampicin. For this purpose, strains were used:

- *Streptomyces rimosus* 2398;
- *Amycolatopsis mediterranei* S699;
- *Actinoplanes teichomyceticus* ATCC 31121;
- *Micromonospora* sp. BRA 006A;
- *Rhodococcus* sp. BRB 36;

From this, the following concentrations of antibiotics were determined for screening:

- Vancomycin and teicoplanin: 10µg/ml;
- Rifampicin: 40µg/ml;

After defining the resistance filter, the bacteria from the collection were cultivated in 6-well plates containing 3mL of the ISP2 culture medium with the respective antibiotic used in the screening. In these plates, a standard inoculum of 100 µl of $1x10^8$ spores per mL of actinobacteria standardized according to the scale at 0.5 McFarland nephelometric. As a control, the tested strains were also grown in ISP2 without the presence of antibiotics to identify resistant bacteria through contrast between the growth. The plates were incubated under agitation at 180 RPM for a period of 72 to 168 hours at a 28°C.

3.3 Molecular screening

The primers for amplification of biosynthetic genes and the vanHAX cluster responsible for resistance to the glycopeptide class of antibiotics were the same used in the work of THAKER et al., 2013 (Table 1). Such genes act in fundamental steps in the complex construction of a glycopeptide molecule (YIM et al., 2013).

Table 1. List of primers used to screen the bacterial collection for glycopeptides

Target	Sequence	Gene	Amplicon
oxyB	oxyBF: CTGGTCGGCAACCTGATGGAC oxyBR:CAGGTACCGGATCAGCTCGTC	P450 oxidase	730 bp
halL	halF:CGACGCSTCGGGCAACAAGAGC halR:GGHCACCAGBCCKCCGGGGAAA	Halogenase	950 bp
dpgC	dpgCF:GGGCDSGAGATCGAYCAGGG dpgCR:AGGTTCAGCATSCGCCGGTTG	Dpg synthase	800 bp
vanHAX	vanHF:CSCTGATGCTSATGYTGATGG vanXR:GGYTTGCCGGTGAARTTGTCC vanAR:CCTGCATCGCGCCGTCYTCRCCGAG	Resistance	1700 bp
oxyE	oxyEF:GGCCSCTGCACGAATACGACACC oxyER:CCACTCAASGGGAAGCTCCTC	P450 oxidase	1100 bp

For antibiotics of the ansamycin class to which rifampicin belongs, primers were chosen (Table 2):

Table 2. List of primers used to screen the bacterial collection for ansamycins.

Target	Sequence	Gene	Amplicon
rifK (Thaker et al. 2013)	AHBA-F: AGAGGATTCGAGCRSGAGTTCGC AHBA-R:GCAGGATCCGGAMCATSGCCATGTAG	AHBA synthase	750 bp
rifP	rifPF:TCGATCTTCCTSATCAASATCCC rifPR:GGTTGCGGAAKAAVVCCATGTC	Resistance	292 bp

The degenerate primer of the *rifP* resistance gene was designed from conserved sequences of this nucleotide sequence in different species of actinobacteria verified with the aid of the NCBI Blast tool (MADDEN, 2003), inserting the degenerations where they were needed.

The validation of the primers was performed using the strains *Amycolatopsis mediterranei* S699, *Actinoplanes teichomyceticus* ATCC 31121

and *Amycolatopsis orientalis* ATCC 19795, standard strains recognized as producers of the antibiotics rifampicin, teicoplanin and vancomycin respectively.

Total genomic DNA extractions were performed following the standard protocol for Gram positive bacteria from the Wizard® Genomic DNA Purification Kit (Promega). A 1% agarose gel electrophoresis was performed to verify the quality of the extraction and the quantification of the total extracted DNA and the quantification in ng/µL was performed in an Eppendorf BioPhotometer D30 nanodrop.

PCR reactions for biosynthetic and resistance genes were performed following the standard Phusion High-Fidelity DNA Polymerases enzyme kit protocol (Thermo) with the following specifications for 50µL reactions:

- Buffer GC: 10µL;
- 1mM dDNTPs: 1µL (10mM)
- Primer foward: 1µL (10mM)
- Primer reverse: 1µL
- DMSO: 1,5µL
- Phusion DNA polimerase: 0,5µL
- DNA template: 2µL (100 ng/µL)
- Milli-Q Water: 33µL

The cycle program followed the following steps: 98º (5 minutes) for initial denaturation. Followed by 30 cycles of 98º for 1 minute, 60º for 30 seconds and 72º for one minute, followed by one cycle at 72º for 5 minutes for final extension. For genes whose amplicon is smaller than 1kb, the extension time at 72º was reduced from 1 to 30 seconds.

3.4 Biological activity assays

Antimicrobial activity tests with actinobacteria from the collection were performed by 2 different methods;

Antagonism: Actinobacteria spores were seeded on petri dishes with the solid culture media ISP2, ISP3, ISP5, ISP6, Mannitol/Soybean Flour and R5

(KIESER et al., 2000) (Appendix 1), thus they grow over the entire surface like a homogeneous carpet. After an incubation period of 7 days, agar discs with a diameter of 7 mm were cut out and placed on a mat of pathogenic bacteria on Mueller-Hinton Agar to observe the formation of an inhibition halo.

Antibiogram: Possible bacteria producing antimicrobial metabolites were cultivated in Erlenmeyers with the liquid culture medium R5. After incubation for 7 days under agitation at 180 rpm. After this period, the metabolic liquid was separated from the cells by filtration and the solvent ethyl acetate was added in a 1/1 proportion to obtain the extract, the mixture was vigorously stirred and distributed in previously weighed amber eppendorf tubes. The solvent was then evaporated with the help of a speedvac concentrator. The dry extract was resuspended in DMSO and embedded in 7mm diameter filter paper discs. These discs were placed on petri dishes with Mueller-Hinton Agar previously seeded with pathogenic bacteria to observe the inhibition halo. Extracts that inhibited pathogens had their MIC calculated in a 96-well plate assay following CLSI recommendations, 2020 ("M100 Performance Standards for Antimicrobial Susceptibility Testing A CLSI supplement for global application", [s.d.]).

3.5 Genome sequencing and assembly

Sample sequencing was performed by the Macrogen laboratory (South Korea, Seoul) using the HiSeq System platform (Illumina Inc., San Diego, USA).

The quality of reads resulting from sequencing was evaluated using the FastQC tool ("Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data", [s.d.]). Low quality readings and adapters were removed with the Trimmomatic program (BOLGER; LOHSE; USADEL, 2014). The de novo montage was performed using the SPAdes algorithm (ANTIPOV et al., 2016a; BANKEVICH et al., 2012).

The assembly statistics in contigs were obtained by the QUAST-LG tool (MIKHEENKO et al., 2018) and genome completeness was verified with the Busco pipeline (SIMÃO et al., 2015).

3.6 Genome analysis

The phylogeny of BRB040 was inferred with the TYGS tool (MEIER-KOLTHOFF; GÖKER, 2019). The results were used for a second scaffold assembly of the genome using the MeDuSa server (BOSI et al., 2015). The functional annotation of the genome was performed using Prokka software (SEEMANN, 2014).

Identification of biosynthetic gene clusters was performed via antiSMASH 6.0, bacterial version (BLIN et al., 2021), with strict detection rigor in order to identify gene groups with all the necessary parts for the synthesis of a bioproduct. The functional annotation resulting from Prokka was used to guide and refine the search. In addition to the antiSMASH, the ARTS software was used to track antibiotic resistance targets that may be part of gene clusters for the biosynthesis of these molecules in the genome of BRB040 (MUNGAN et al., 2020).

Additional analyzes of the genomic data were performed with the softwares:

- ANI Calculator tool, to calculate the distance between different genomes, identifying their percentage of identity (YOON et al., 2017).
- plasmidSPAdes, for assembly of plasmid sequences from the sequencing reads (ANTIPOV et al., 2016b).
- pLannotate, for annotation of the sequences of smaller plasmids (MCGUFFIE; BARRICK, 2021).
- Mega 11, for alignment and phylogenetic analysis of specific genes (TAMURA; STECHER; KUMAR, 2021).

In Mega, sequence alignment was performed using the MUSCLE (MUltiple Sequence Comparison by Log- Expectation) (EDGAR, 2004). The phylogenetic trees of the genes were constructed using the Maximum-Likelikelihood statistical method and the Jukes-Cantor model (JUKES; CANTOR, 1969), with 1000 bootstrap repetitions with a matrix of distances paired with the two-parameter substitution model (KIMURA, 1980).

4 Results and discussion

4.1 Screening for resistant bacteria

650 isolates from the marine actinobacteria collection were screened for resistance to the antibiotics vancomycin, teicoplanin and rifampicin. Of these isolates, some showed resistance to more than one antibiotic (Table 3).

Table 3. Resistant isolates identified in the bacterial collection.

Antibiotic	Strain code	Total resistant
Rifampicin 40 µg/ml	*S.rimosus*, *S.rimosus* 2398, BRB 05, BRB 14, BRB 16, BRB040, BRB 42, BRB 48, BRB 51, BRB 52, BRB 88, BRB 89, BRB 105, BRB 107, BRB 110, BRB 121, BRB 173, BRB 219, BRB239, BRB242, BRB 243, BRB 246, BRB248, BRB 250, BRB 312, BRB 324, BRB 325, BRB 333, BRB 366, BRB 369, BRB0401, BRB0408, BRB 412, BRB 422, BRB 437, BRB 454, BRB 460, BRB 529, BRB 531, BRB 562, BRB 570, BRB 639A	52
Vancomycin 10 µg/ml	*S.rimosus*, BRB 05, BRB 14, BRB 43, BRB50, BRB 89, BRB 99, BRB 105, BRB124, BRB239, BRB248, BRB250, BRB 271, BRB 303, BRB 316, BRB 333, BRB 423, BRB 444, BRB 445, BRB 451, BRB 471, BRB 489, BRB 500, BRB 501, BRB 517, BRB 523, BRB 565, BRB 611	28
Teicoplanin 10 µg/ml	BRB 05, BRB 13, BRB 22, BRB 36, BRB 50, BRB 89, BRB 99, BRB 105, BRB 239, BRB 240, BRB 248, BRB 250, BRB 303, BRB 326, BRB 333, BRB 352, BRB 471, BRB 489, BRB 500, BRB 501, BRB 517, BRB 523, BRB 565, BRB 611	24
Resistant to 3 antibiotics	BRB 333, BRB 05, BRB 239, BRB 248, BRB 250, BRB 89, BRB 105	7
Vancomycin and teicoplanin resistant	BRB 523, BRB 471, BRB 303, BRB 500, BRB 565, BRB 501, BRB 50, BRB 99, BRB 489, BRB 517, BRB 611	11
Vancomycin and rifampicin resistant	S.rimosus, BRB 14	2

Note the percentage of differentiated resistance to different antibiotics among the bacteria in the collection of only 4% resistant to the glycopeptides vancomycin and teicoplanin, but get up to 8% of isolates resistant to rifampicin. The low frequency of glycopeptide-resistant isolates may suggest a possible low sampling of bacteria in the collection from the genera *Amycolatopsis* and *Actinoplanes*, where the presence of glycopeptide biosynthesis clusters is identified more frequently and, consequently, the set of self-resistance genes (YUSHCHUK; BINDA; MARINELLI, 2020).

In the case of rifampicin, the resistance mechanisms involved appear to be more varied than glycopeptide resistance. While resistance to glycopeptides is more closely linked to the vanHAX group described above, resistance to rifampicin can be mediated by different factors, such as efflux systems, enzymes capable of degrading the antibiotic and target modification. Although this resistance is not widespread to the point of making screening impossible, we still have a greater number of resistant isolates when compared to glycopeptides.

4.2 PCRs of resistance and biosynthetic genes

The isolates resistant to teicoplanin, vancomycin and rifampicin had their DNA extracted, and an attempt was made to amplify genes responsible for biosynthesis and resistance to the antibiotic classes wanted. In the case of glycopeptides, as the vanHAX cluster is always associated with the presence of a biosynthetic cluster in producing bacteria (YIM et al., 2013; YUSHCHUK; BINDA; MARINELLI, 2020), priority was given to finding isolates that amplified this sequence of genes to later search for the other genes responsible for different stages in the biosynthesis of a possible molecule of the glycopeptide class.

No amplifications were found for the vanHAX cluster other than those that occurred in the controls *Actinoplanes teichomyceticus* ATCC 31121 and *Amycolatopsis orientalis* ATCC 19795.

As for the class of ansamycins, priority was given to the amplification of the AHBA synthase gene (*rifK*). Of the tested isolates, 3 obtained a positive amplification for this gene: BRBs 014, 040 and 042. Despite this, it was not possible to identify the *rifP* gene in any of the bacteria. The reason for this may lie in the previously discussed fact that resistance to this class of molecules is linked to other mechanisms.

4.3 Antimicrobial activity assays

Antimicrobial activity assays by antagonism were performed to quickly obtain information regarding the ability of the evaluated bacteria to produce

metabolites capable of inhibiting the growth of clinically relevant pathogens before deciding to refer them to whole genome sequencing (Table 4).

Table 4. Results of antagonism biological activity assays and the MIC of crude extracts.

Pathogen	Strain that inhibited	Halo diameter (mm)	Culture medium	MIC with R5 extract (μg)
Klebsiella pneumoniae ATCC 4352	BRB 42	13,53	R5 e MS	100
Burkholderia cepacia ATCC 17759	BRB 42	10,13	ISP2	100
Staphylococcus aureus ATCC 14458	BRB 42	13,20	R5 e ISP2	50
E. coli ATCC 25922	BRB040	12,60	ISP2 e MS	200
E. coli KPC	BRBs 40 e 42	10 e 9 mm	R5 e MS	400

The MIC showed that only at the highest concentrations it was possible to observe some inhibition of the pathogens. This effect may be due both to the low number of bioactive metabolites present in the extracts and to a low effect of these metabolites on the pathogens. Producer bacteria often synthesize their antibiotics in cryptic amounts, and it is often challenging to obtain quantities of these compounds to observe any effect against other bacteria (BALAGURUNATHAN et al., 2020; HOSKISSON; SEIPKE, 2020; VAN DER HEUL et al., 2018). In this case, some strategies can be adopted to stimulate the production of antibiotic compounds, such as the use of elicitors, chemical compounds capable of binding to specific receptors and stimulating the production of secondary metabolites in bacteria. (HOSKISSON; SEIPKE, 2020; RUTLEDGE; CHALLIS, 2015). Also, the variation of cultivation conditions in order to generate a higher yield of the metabolite of interest (ROMANO et al., 2018; SCHWARZ et al., 2021).

In addition, the effect of these extracts on pathogenic bacteria must be viewed with caution, as it may be due to the synergistic effect of several different molecules produced by actinobacteria acting together against the pathogen

(ACAR, 2000; KOHANSKI; DWYER; COLLINS, 2010; LEEKHA; TERRELL; EDSON, 2011).

4.4 Genome Assembly

The first level of genome assembly of BRB040 generated a total of 68 contigs and a size and genome size of 8 326 214 base pairs. It presented a completeness of 99.5% verified via BUSCO, which searched the genome for homologues of 1575 genes in its database of 145 genomes of the *Streptomycetales* order.

4.5 Phylogeny and second assembly

The resulting dataset was analyzed on the TYGS platform, where a taxonomic analysis is performed comparing the genome of BRB040 with 18742 complete genomes of strains stored in its database.

The data set in contigs is insufficient to carry out genomic mining, since it is fragmented, so that genes belonging to the same cluster can be separated in this arrangement (Figure 4).

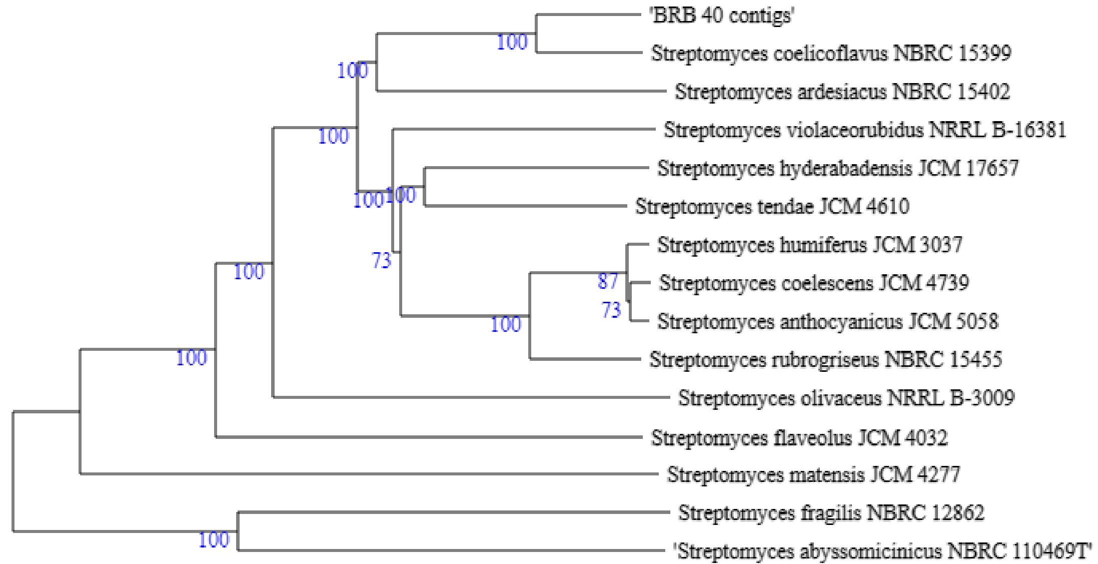


Figure 4. TYGS Generated Whole Genome Alignment Phylogenetic Tree for BRB040

In order to obtain a genome with best contiguity, a second level of assembly was adopted following the principle of syntenic conservation, where it is postulated that the genomic arrangements remain stable due to functional and evolutionary restrictions, thus the position of the genes in the chromosome tends to be preserved among the chromosomes of different related species (HILLIER et al., 2007).

To perform the second level of assembly using the MeDuSa server, the complete genomes of *Streptomyces coelicoflavus* NBRC 15399 and *Streptomyces ardesiacus* NBRC 15402 were chosen as a reference, which showed the greatest phylogenetic proximity to BRB040. The comparison between the two assemblies is shown in the table 5 below:

Table 5. Comparison between the first and second genome assembly of BRB040.

Statistics	BRB040 Contig Assembly	BRB040 Scaffold Assembly
Number of sequences	68	10
Largest sequence (bp)	1374921	8274393
Genome size (bp)	8326214	8278027
N50 (bp)	457246	6015735
L50	6	1
GC content (%)	73.1	72.30
Completeness (%)	99.8	99.2
Complete orthologs	1571	1561
Duplicate orthologs	4	5
Fragmented orthologs	1	2
Missing orthologs	3	11

The increase in contiguity of the second assembly is demonstrated by the reduction in the number of sequences between the two assemblies, with the largest of these sequences covering 99.9% of the genome. In addition, the metrics N50 and L50, which represent respectively the number of the smallest sequence that combined with progressively longer fragments covers 50% of the genome and the number of sequences necessary for this, indicate a more contiguous genome.

4.6 General characteristics of the genome

The BRB040 isolate has a linear chromosome of 8.2 Mb in size, compatible with that of *Streptomyces coelicoflavus* which has 8.7 Mb. The difference in size may be due to the trimming process at the ends of the readings, which may have removed some information that led to a reduction in the total genome size and the loss of some genes. The GC content of 72.30% also matches the average percentage of GC found in genomes of the genus *Streptomyces* and very close to the GC percentage of *Streptomyces coelicofavus*, 72.18%, the taxonomically closest genome.

To verify the proximity between BRB040 and *Streptomyces coelicoflavus*, the average nucleotide identity between the two genomes was

calculated using the ANI Calculator tool (YOON et al., 2017) which presented a value of 98.48%. It is suggested that the average nucleoid identity values for distinct species is less than 95 - 96% (RICHTER; ROSSELLÓ-MÓRA, 2009). Considering only genomic data, there is an indication that BRB040 belongs to the species *Streptomyces coelicoflavus*.

4.7 The secondary metabolism of BRB040

4.7.1 Overview

Genomic mining using automated tools has paved the way for exploring the vast richness present in bacterial genomes. In the present work, 23 regions containing gene clusters related to bacterial secondary metabolism were identified (Figure 5, Table 6). We observed that 12 of the 23 clusters of BRB040 biosynthesis genes are found at the left end of the chromosome and 7 at the right end. It is known that the *Streptomyces* chromosome presents great genetic instability, suffering large spontaneous deletions and mutations at its ends that affect different phenotypic properties, including morphological differentiation, production of secondary metabolites, such as pigments and antibiotics, resistance to antibiotics, secretion of extracellular enzymes and sometimes genes for primary metabolism (VOLFF; ALTENBUCHNER, 1998). We can assume that most of the BRB040 biosynthesis gene clusters are located in the extremity regions, which in turn are more easily affected by these conditions of chromosomal instability. Although chromosomal instability is more frequently observed at the ends, cases where this instability affects the central region of the chromosome have already been reported (CHEN et al., 2010).

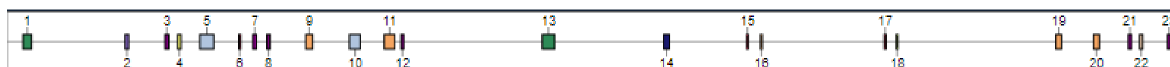


Figure 5. Overview of BRB040 secondary metabolite coding regions distributed along the chromosome.

Table 6. List of secondary metabolites produced by BRB040.

Region	Type of cluster	Location in genome	Similar compound	Identity (%)
1	NRP-metallophore,NRPS	114,005, 172,414	coelichelin	100
2	Butyrolactone	835,613, 859,106	methylenomycin A	23
3	Terpene	1,116,811, 1,142,704	hopene	100
4	lanthipeptide-class-iii	1,207,527, 1,230,223	SapB	100
5	transAT-PKS,NRPS	1,363,747, 1,460,457	inthomycin B	100
6	NI-siderophore	1,639,607, 1,652,060		
7	betalactone	1,736,192, 1,763,968	skyllamycin D/skyllamycin E	7
8	Terpene	1,836,632, 1,858,352	geosmin	100
9	T1PKS, prodigiosin	2,111,454, 2,158,404	undecylprodigiosin	95
10	butyrolactone, T2PKS, phenazine	2,419,619, 2,496,917	lactonamycin	15
11	T2PKS	2,666,187, 2,738,729	nybomycin	74
12	Terpene	2,784,456, 2,803,798	albaflavenone	100
13	NRPS	3,783,479, 3,868,526	sarpeptin A/sarpeptin B	91
14	Other	4,639,631, 4,680,047	nenestatin	30
15	NI-siderophore	5,225,866, 5,236,773	desferrioxamin B/desferrioxamine E	83
16	melanin	5,325,523, 5,336,155	melanin	60

17	NI-siderophore	6,197,904, 6,209,925		
18	ectoine	6,285,479, 6,295,877	ectoine	100
19	T3PKS	7,411,767, 7,450,696	flaviolin/1,3,6,8-tetrahydroxynaphthalene	100
20	T3PKS	7,677,730, 7,718,914	germicidin	100
21	terpene	7,921,851, 7,945,140	carotenoid	45
22	indole	8,002,976, 8,024,103	5-dimethylallylindole-3-acetonitrile	100
23	terpene	8,199,077, 8,220,294	2-methylisoborneol	100

Five of the BGCs found in the genome of BRB040, with 100% identity, are terpenes, a group of molecules that so far have more than 50,000 structures already identified (KUZUYAMA, 2017). These molecules have great industrial and pharmaceutical applications, such as herbicides, flavorings and biofuels. Although most terpenes are obtained from plants or fungi, bacteria can also be a prolific source of this class (REDDY et al., 2020). One of the identified terpenes (region 8) is geosmin, responsible for the characteristic damp earth odor of *Streptomyces* (JIANG; HE; CANE, 2007).

As an organism from a marine environment, BRB040 needs mechanisms to protect itself from the pressures of this environment. Ectoine plays the role of defending the cell from osmotic pressure caused by the saline environment (BERNARD et al., 1993). It was found that BRB040 has in its genome the ability to produce ectoine (region 18).

BRB040 also has 4 groups of genes that synthesize molecules with type II and III polyketide synthases, homodimeric multimodular enzymes involved in the synthesis of most of the bioactive compounds identified from actinobacteria, among which many are known antibiotics (BALTZ, 2006; PLATER; STROHL, 1994; RISDIAN; MOZEF; WINK, 2019; WEBER et al., 2003).

Two biosynthetic clusters responsible for the production of non-ribosomal peptides were also identified, secondary metabolites that are synthesized outside the ribosomal machinery and have a variety of properties, such as cytostatics, immunosuppressants or anticancer agents, antibiotics, pigments and toxins. (DINCER et al., 2022).

We also have hybrid molecules, which combine both polyketide and non-ribosomal peptide synthesis, as is the case of the intomycin B coding region, with 100% identity with the gene cluster of *Streptomyces* sp. YB104 (WU et al., 2018). Intomycins have been reported to possess many relevant biological properties, which include specific inhibition of cell biosynthesis, *in vitro* antimicrobial activity, and anticancer activity against human prostate cancer cell lines (SEYCHELL; BECK, 2021).

An almost complete correlation was observed between the secondary metabolites of BRB040 and *Streptomyces coelicoflavus* NBRC 15399, although a gene cluster does not present in NBRC 15399 was identified being exclusive to BRB040, the one identified by lactonamycin, although the identity is only 15%.

Blast search of the sequence of this cluster revealed 98% identity with plasmids found in *Streptomyces* sp. SYP-A7193 however the maximum coverage for the sequence was 56%.

Knowing this, the MOB-suite tool was used (ROBERTSON; NASH, 2018) to isolate this possible plasmid using the sequence found in SYP-A7193, but without success. For this reason, it was decided to carry out a genomic assembly for the plasmids using the sequencing readings with the plasmidSPAdes (ANTIPOV et al., 2016b). Only with the sequences obtained by this step was it possible to use the MOB-suite to isolate the plasmid sequences of the BRB040 assembly.

Horizontal gene transfers are common in the evolutionary history of *Streptomyces* (MCDONALD; CURRIEA, 2017). The strict parameter of antiSMASH only searches for gene groups where it is possible to identify all its functional parts, although it has been unable to strongly relate the hypothetical product of this plasmid with any known metabolite, the comparative analysis shows that there are several fragments of regions belonging to different

biosynthetic pathways (Figure 5). Some hypotheses can be pointed out for this apparently disordered of this group of genes. The first is the possibility of an error in the assembly of the genome. The other is the presence of a group of biosynthetic genes of a product not yet identified in the plasmid found in BRB040. Without a close match for sequence alignment, the genomic mining software found only a few related genes in different clusters of different biosynthesis products (Figure 5).

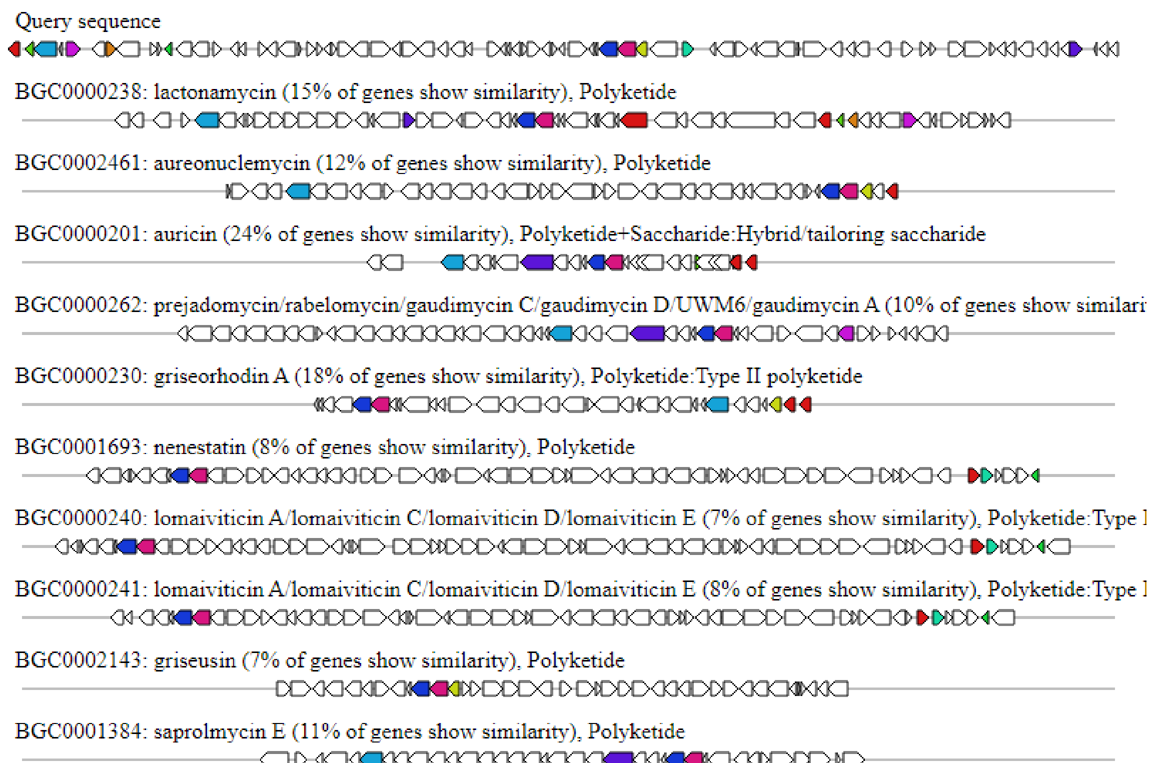


Figure 5: Similarity of known clusters with the BRB040 region 10 biosynthetic gene cluster.

Removal of the plasmid did not alter the configuration of the biosynthetic gene clusters except for the exclusion of region 10 which made the metabolic profile of BRB040 identical to that of *Streptomyces coelicoflavus*. After this final data processing, the BRB040 genome is ready to be deposited in public databases along with its plasmid sequences.

4.7.2 The BRB040 plasmids

Two sets of sequences reported as plasmids in BRB040 were identified with sizes of 123 kb and 14 kb, sizes are related with those reported for plasmids in the literature, which range from 5 to 500 kb, although plasmids even larger than 1 Mb have already been found and smaller than 2 kb (KOTHARI et al., 2019). In *Streptomyces*, there are linear plasmids of sizes as large as 1850 kb containing sequences capable of encoding different molecules (KINASHI, 2010) and also circular plasmid sequences of size between 8 to 31 kb (ZHANG et al., 2008).

One of the targets of the screening carried out in the present work was the *rifK* gene (AHBA synthase), fundamental for the biosynthesis of antibiotics of the ansamycin family. Identification of this gene via PCR together with verification of antimicrobial activity was responsible for choosing BRB040 for further studies staring for its genome sequence.

After plasmid annotation using the pLannotate tool (MCGUFFIE; BARRICK, 2021) (Figure 6) (Appendix 2), it was found that the homologue of the *rifK* gene found in BRB040 is not associated with any biosynthetic pathway found in this organism, but is found in a 14 kb plasmid sequence containing gene fragments without the complete assembly of genes to produce a bioproduct. This discovery reveals a limitation in the technique proposed by Thaker and his collaborators in 2013, the possibility of a false positive occurring due to the exchange of DNA by horizontal transfers of genetic material. In addition to the *rifK*, a *macB* homologue was found, an ABC transporter capable of conferring resistance against macrolides (XU et al., 2009), and perhaps cross-resistance against rifampicin. The *lolD* gene encodes an ABC transporter linked to the transport of lipids between biomembranes (DAVIDSON; CHEN, 2004; ITO et al., 2006).

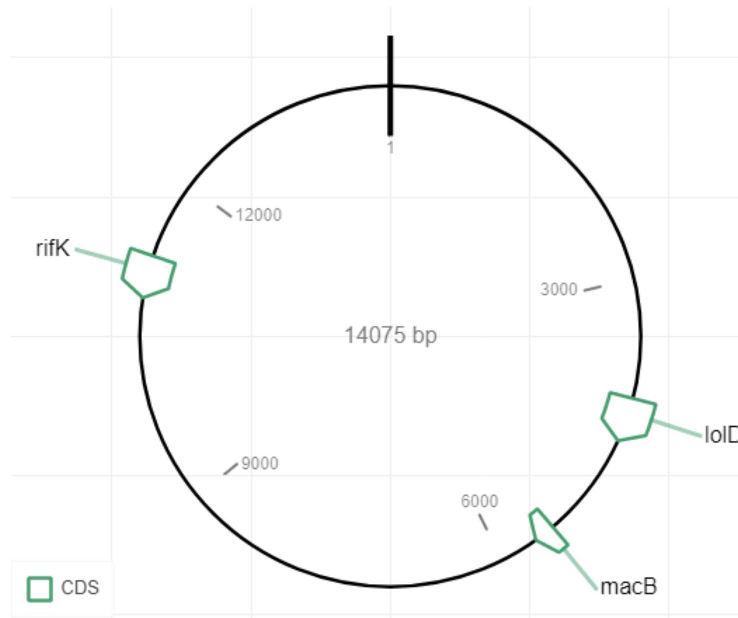


Figure 6. Illustration of the smaller plasmid provided by pLannotate found in BRB040 with the *rifK* gene and the ABC transporter genes *macB* and *lolD* are highlighted.

Using the ARTS tool to identify possible models of resistance in BRB040 larger plasmid (123 kb) (Appendix 3), a total of 123 genes were identified with two possible models of antibiotic resistance. The first of these is a mutation in the gene encoding acetyl-coenzyme-A carboxylase that confers resistance to the herbicides diclofop and haloxyfop (ZHANG; TWEEL; TONG, 2004). The other pattern of resistance detected was an ABC efflux system like the one found in the smaller plasmid. In addition to being able to confer drug resistance, these transporters mediate the transmembrane transport of a diverse variety of substrates, including drugs, sugars, ions, amino acids, and proteins (GREENE et al., 2018).

The presence of these giant linear plasmids is highly distributed in *Streptomyces* and these mobile elements may or not contribute directly to the transfer of capability to produce antibiotics between different strains (HINES et al., 2017). Initially, these linear plasmids are transferred to compatible strains, granting them some capacity to produce antibiotics (RAVEL; WELLINGTON; HILL, 2000). Subsequently, the crossover between the plasmid and the chromosome occurs, which transfers the scaffold of the antibiotic to the

chromosome, creating chimeric genomes (PANDZA et al., 1998; YAMASAKI; KINASHI, 2004).

More investigations are needed to determine what function the giant plasmid plays in the genome of BRB040, either because of the possibility of producing some molecule of interest or perhaps because it plays some other role in the synthesis of other BRB compounds.

4.7.3 Region 11: Nibomycin like

The group of 64 genes found in region 11 of BRB040 (Figure 7 / Appendix 4), classified as PKS type 2, presents 74% identity with the set of genes responsible for the synthesis of the antibiotic nibomycin, a compound capable of inhibiting the growth of *E. coli* strains resistant to fluoroquinolones (SHIRIAEV et al., 2021). The group of genes responsible for the synthesis of this antibiotic in *Streptomyces albus* deposited in the MIBIG database (TERLOUW et al., 2023) has a total size of 30490 nucleotides, while the cluster identified in BRB040 has a total of 72543 nucleotides, indicating the presence of more genes and the possibility that the final product is different than nibomycin.

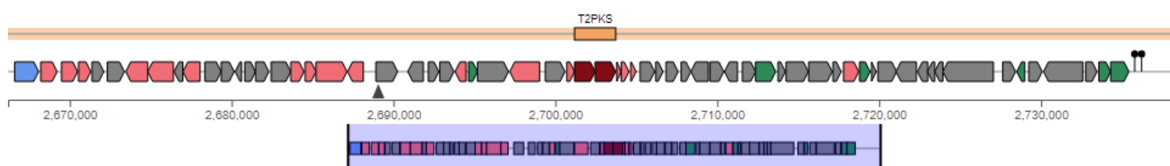


Figure 7. Overview of the region 11 biosynthetic gene cluster. In green: regulators / Blue: genes encoding transport systems / Pink: other biosynthetic genes / Red: core biosynthetic genes / Gray: Other genes.

It is observed that part of the gene cluster of region 11 in BRB040 has 66% identity with some type of pigment (Figure 8). Some antibiotics have the characteristic of being pigmented, such as actinorodine produced by *Streptomyces coelicolor* and cosmomycin D produced by *Streptomyces olindensis* (BYSTRYKH et al., 1996; FURLAN et al., 2004; GARRIDO et al., 2006), which give respectively a blue or purple pigment to the medium in which

them are produced. This characteristic may help to identify the production of this metabolite in future investigations regarding its biosynthetic pathway.

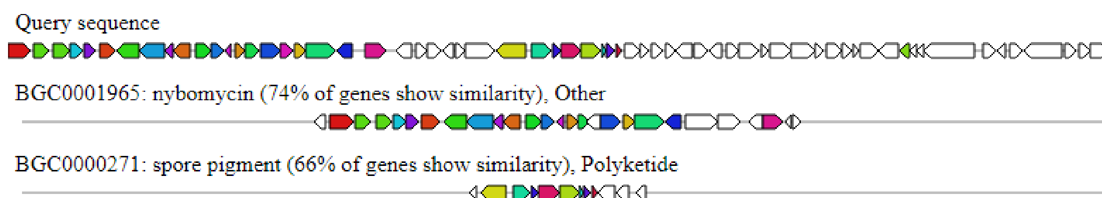


Figure 8. Comparative identity of know clusters and region 11

Another relevant evidence that should be considered is that BRB040 has a sequence that encodes a gene homologous to the nibomycin transporter in *Streptomyces albus*, the presence of this transporter in the group of biosynthesis genes may indicate the mechanism of self-resistance of the bacteria to this compound.

Phylogenetic analysis of the genes related to the hypothetical transporter MFS-type EfpA (ORF1) (Figure 9) performed with the Mega 11 software (TAMURA; STECHER; KUMAR, 2021) showed that the gene identified in BRB040 and the transporter present in the nibomycin biosynthetic cluster belong to distinct clades. This evidence reinforces the hypothesis that we are facing a different product than nibomycin.

Figure 9. Phylogenetic tree of the putative MFS-type transporter EfpA found in the BRB040 region 11 gene cluster and their homologues in other species related.

The MFS-type are essential for the movement of different substrates across biomembranes, in bacteria they are mainly used for nutrient absorption

and extrusion of harmful compounds (QUISTGAARD et al., 2016). The *efpA* gene that encodes this transporter is also present in a range of pathogens that are also part of the *Actinobacteria* phylum, such as *Mycobacterium leprae* and *Mycobacterium tuberculosis* (DORAN et al., 1997). Evidence indicates that a number of antibiotic resistance mechanisms used by antibiotic-producing bacteria are also conserved in pathogenic organisms such as *Mycobacterium* (GYGLI et al., 2017; MORRIS et al., 2005).

Another strong evidence is the presence of two central biosynthesis genes that encode Actinorhodin polyketide beta-ketoacyl synthase enzymes in this cluster (ORFs 31, 32).

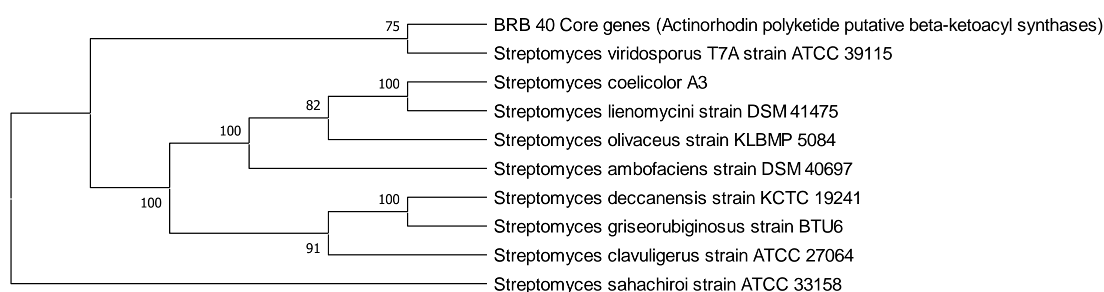


Figure 10. Phylogenetic tree with concatenated Actinorhodin polyketide beta-ketoacyl synthase 1 and 2 enzyme coding genes and their homologues related in different *Streptomyces* species.

A phylogeny was also carried out with the Actinorhodin polyketide beta-ketoacyl genes and their homologues and it was verified that the gene present in BRB040 belongs to a clade distinct from that present in *Streptomyces coelicolor*. The search by Blast did not find sequences for these genes in *Streptomyces coelicoflavus*, which may indicate that the cluster present in BRB040 may have another gene configuration. Such a fact would not be unusual considering that the repertoire of natural products in *Streptomyces* is sometimes strain specific.

4.7.4 Region 14: Nenestatin like?

The cluster found by antiSMASH in region 14 shows 30% identity with the nenestatin antibiotic biosynthesis scaffold was also observed. We also

observed a similar degree of identity with the lomaiviticin synthesis gene cluster (Figure 11). Nenestatin is a benzofluorene with broad-spectrum antibacterial activity and cytotoxicity produced by the deep sea *actinomycete Micromonospora echinospora* (JIANG et al., 2021). Lomaivitcins are genotoxic metabolites and potent cytotoxins, with half-maximal inhibitory potency (IC 50) values in the nanomolar-picomolar range against many cancer cell lines (HERZON, 2017; KIM et al., 2021).

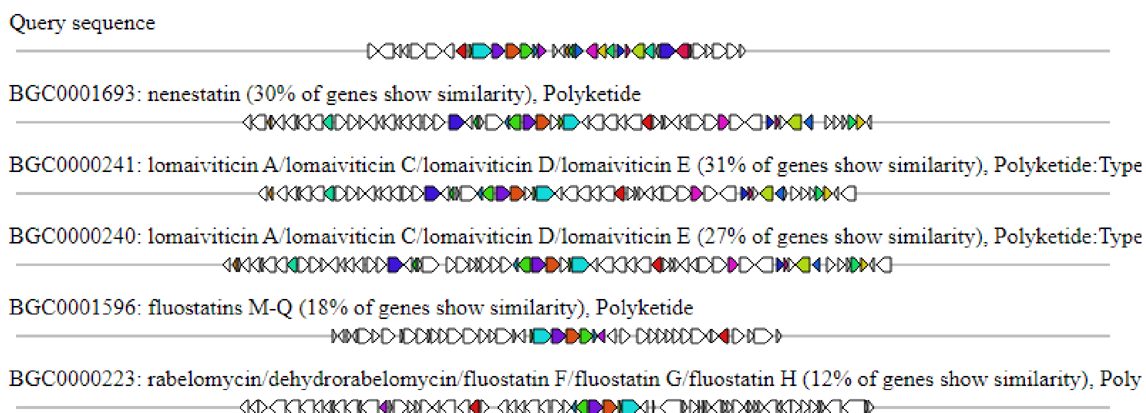


Figure 11 Comparative identity of know clusters and region 14

The region of this grouping contains 40417 nucleotides and 36 genes (Figure 12), being identified a central biosynthesis gene homologous to antitumor antibiotic C-1027 apoprotein. C1027, a macromolecular peptide antitumor antibiotic produced by *Streptomyces globisporus* C1027 (ZHANG et al., 2020).

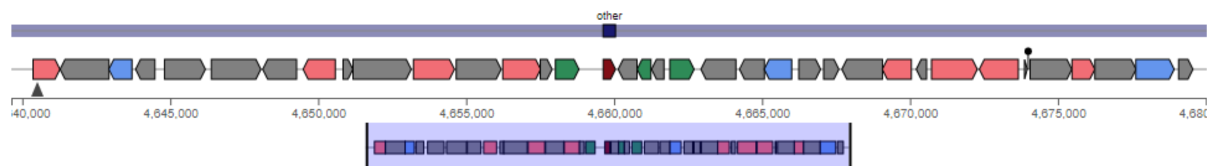


Figure 12. Overview of the region 14 biosynthetic gene cluster. In green: regulators / Blue: genes encoding transport systems / Pink: other biosynthetic genes / Red: core biosynthetic genes / Gray: Other genes.

A search of the amino acid sequence encoded by the central biosynthesis gene of cluster 14 found 35 matches with 50 to 100% identities with the enediyne chromoproteins, a class of potent antitumor antibiotic, comprising a 1:1 complex

of a protein and a non-covalently linked chromophore (BAKER et al., 2007; IWASAKI, 1999).

The search for matches of the nucleotide sequence of the gene found only 5 matches, with only 2 hits having 100% coverage while the other 3 having 7% coverage (Figure 13).

Sequences producing significant alignments Download Select columns Show 100

select all 5 sequences selected GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Streptomyces sp. SYP-A7193 plasmid unnamed2, complete sequence	Streptomyces sp. SYP-A7193	760	760	100%	0.0	99.52%	120680	CP045549.1
Streptomyces pactum strain ACT12, complete genome	Streptomyces pactum	577	1154	100%	7e-160	91.61%	8550793	CP019724.1
Monoraphidium neglectum hypothetical protein mRNA	Monoraphidium neglectum	56.5	56.5	7%	0.004	96.97%	405	XM_014040127.1
Clavibacter phaseoli strain LPPA 982 chromosome, complete genome	Clavibacter phaseoli	54.7	54.7	7%	0.014	96.88%	3103693	CP040786.1
Clavibacter phaseoli strain CFBP 8217 chromosome, complete genome	Clavibacter phaseoli	54.7	54.7	7%	0.014	96.88%	3103717	CP040795.1

Figure 13. Blast hits of the nucleotide sequence of the central Antitumor antibiotic C-1027 apoprotein gene found in region 14 in BRB040.

Despite the promising pharmaceutical potential, research with Enediyne Antitumor Antibiotic C-1027 has not advanced due to the low titer of the original producer (YAN et al., 2018). Analysis show that the clustering genes in different bacteria shows that are all located on giant plasmids of varying sizes. The high nucleotide sequence similarity among the C-1027 clusters implies that they most likely have evolved from a common ancestor.

It is possible that the identification of the product encoded by region 14 as nenestatin is due to the lack of biosynthetic gene clusters in the databases for enedyine producers. Due to the low availability of genomic data to perform a phylogenetic analysis, it was decided to perform a sequence similarity network follows by a genome-wide neighborhood analysis using the protein sequence encoded by the core biosynthesis gene was performed using the algorithms in the Enzyme Function Initiative tools to support this hypothesis (OBERG; ZALLOT; GERLT, 2023; ZALLOT; OBERG; GERLT, 2019). The sequence similarity network built using proteins deposited in the UniProt database revealed only 15 homologous proteins with sequences with e-values less than $1x10^{-5}$

(Figure 14). The created network showed a set of 10 highly conserved sequences and another 5 external protein sequences.

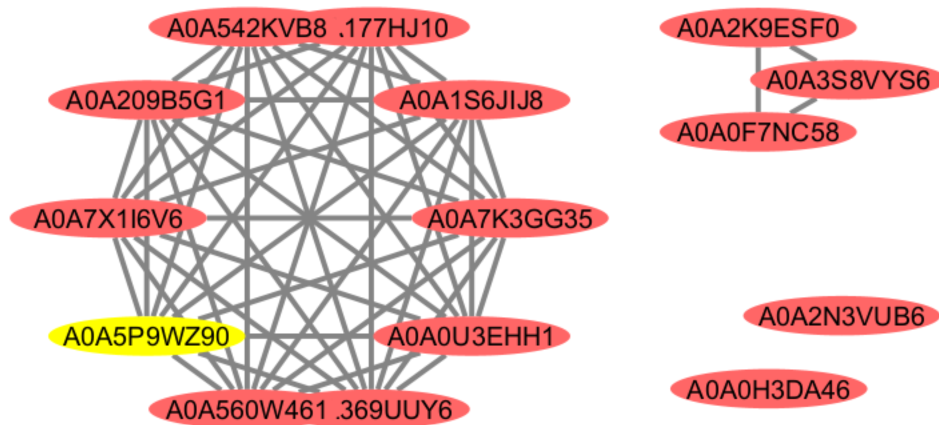


Figure 14. Sequence similarity network of the Antitumor Antibiotic C-1027 gene. Here, each network member is represented by a oval label and is connected to the other members by lines. What determines the distance between the different members is the similarity threshold. For this network, a threshold of 40% identity was used. Highlighted in yellow, the protein present in BRB040.

A genomic neighborhood analysis was performed to identify the closest proteins found in genomic clusters where Antitumor Antibiotic C-1027 protein is present (Figure 15).

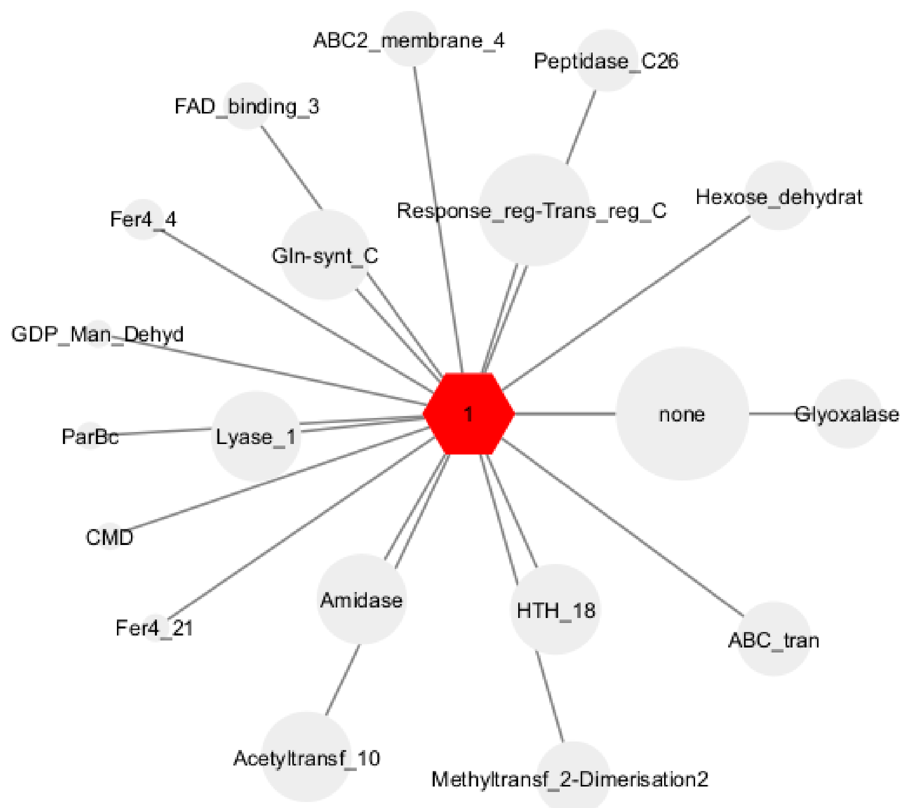


Figure 15. Neighborhood of genomes showing the closest protein families found in gene clusters where we find Antitumor Antibiotic C-1027 protein (red). The larger circles mean more frequently these proteins are found in different clusters. The lines indicate the distance of these protein families from the central gene.

This analysis shows that the protein identified as the central biosynthesis gene found in region 14 of BRA040 has in its vicinity the same genes coding for proteins found in other enedyine producers such as amidase (ORF 11), lyase (ORF 13), ABC transporters of different types (ORF 3, ORF 23, ORF 35), transcriptional regulator (ORF 15) (Look at Appendix 5).

These results suggest that the gene cluster in region 14 of BRB040 has good similarity with other clusters similar to enedyine producers, but the search in the NCBI sequence database revealed little correspondence with the nucleotide sequence of the gene that encodes the protein. Antitumor Antibiotic C-1027. It is interesting evidence that the gene scaffold found in BRB040 could be a new producer of this class of molecules.

Conclusions

The present work identified that the method of Thaker and his collaborators for screening organisms producing bioactive compounds, although promising, can be susceptible to failures. The appearance of false positive, however, could be mitigated with the association of the resistance screening method with the traditional screening technique, which was able to identify the biological activity of the tested organisms.

Directing the focus to a specific set of interesting bacteria for genomic mining. The assembly of the genomic draft of BRB040 was a fundamental step in the pre-analysis of the data. Genome mining using antiSMASH revealed the presence of 23 biosynthesis gene clusters capable of encoding secondary metabolites.

Two clusters of genes were found with promising evidence to be considered new potential scaffolds for the production of new antibiotics Reliable evidences suggests that the clusters identified in regions 11 and 14 may be a possible new scaffold for the production of a new antibiotic.

References

ABRAHAM, E. P.; CHAIN, E. An Enzyme from Bacteria able to Destroy Penicillin. **Nature 1940 146:3713**, v. 146, n. 3713, p. 837–837, 1940.

ABSALÓN, A. E. et al. RifP; a membrane protein involved in rifamycin export in Amycolatopsis mediterranei. **Biotechnology Letters**, v. 29, n. 6, p. 951–958, 10 jun. 2007.

ACAR, J. F. ANTIBIOTIC SYNERGY AND ANTAGONISM. **Medical Clinics of North America**, v. 84, n. 6, p. 1391–1406, 1 nov. 2000.

ALBARANO, L. et al. Genome Mining as New Challenge in Natural Products Discovery. **Marine Drugs 2020, Vol. 18, Page 199**, v. 18, n. 4, p. 199, 9 abr. 2020.

AMINOV, R. I. A brief history of the antibiotic era: Lessons learned and challenges for the future. **Frontiers in Microbiology**, v. 1, n. DEC, p. 134, 8 dez. 2010.

ANTIPOV, D. et al. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. **Bioinformatics**, v. 32, n. 7, p. 1009–1015, 1 abr. 2016a.

ANTIPOV, D. et al. plasmidSPAdes: assembling plasmids from whole genome sequencing data. **Bioinformatics**, v. 32, n. 22, p. 3380–3387, 15 nov. 2016b.

ANTONY-BABU, S. et al. Multiple Streptomyces species with distinct secondary metabolomes have identical 16S rRNA gene sequences. **Scientific Reports**, v. 7, n. 1, 1 dez. 2017.

AUGUST, P. R. et al. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of Amycolatopsis mediterranei S699. **Chemistry & biology**, v. 5, n. 2, p. 69–79, 1998.

Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Disponível em: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Acesso em: 19 mar. 2023.

BAKER, J. R. et al. Protein–Small Molecule Interactions in Neocarzinostatin, the Prototypical Enediyne Chromoprotein Antibiotic. **ChemBioChem**, v. 8, n. 7, p. 704–717, 7 maio 2007.

BALAGURUNATHAN, R. et al. Eliciting Cryptic Metabolite Production from Actinobacteria. p. 129–137, 2020.

BALTZ, R. H. Molecular engineering approaches to peptide, polyketide and other antibiotics. **Nature Biotechnology 2006 24:12**, v. 24, n. 12, p. 1533–1540, 11 dez. 2006.

BANKEVICH, A. et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, 1 maio 2012.

BARKA, E. A. et al. Taxonomy, Physiology, and Natural Products of Actinobacteria. **Microbiology and Molecular Biology Reviews**, v. 80, n. 1, p. 1–43, 25 mar. 2016.

BELKNAP, K. C. et al. Genome mining of biosynthetic and chemotherapeutic gene clusters in Streptomyces bacteria. **Scientific Reports 2020 10:1**, v. 10, n. 1, p. 1–9, 6 fev. 2020.

BENTLEY, S. D. et al. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). **Nature 2002 417:6885**, v. 417, n. 6885, p. 141–147, 9 maio 2002.

BERNARD, T. et al. Ectoine accumulation and osmotic regulation in Brevibacterium linens. **Journal of General Microbiology**, v. 139, n. 1, p. 129–136, 1 jan. 1993.

BIERMANN, F.; HELFRICH, E. J. N. Hidden Treasures: Microbial Natural Product Biosynthesis off the Beaten Path. **mSystems**, v. 6, n. 4, 31 ago. 2021.

BINDA, E.; MARINELLI, F.; MARCONE, G. L. Old and New Glycopeptide Antibiotics: Action and Resistance. **Antibiotics**, v. 3, n. 4, p. 572, 4 nov. 2014.

BLIN, K. et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. **Nucleic Acids Research**, v. 49, n. W1, p. W29–W35, 2 jul. 2021.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

BOSI, E. et al. MeDuSa: a multi-draft based scaffolder. **Bioinformatics**, v. 31, n. 15, p. 2443–2451, 1 ago. 2015.

BUGG, T. D. H. et al. Molecular Basis for Vancomycin Resistance in Enterococcus faecium BM4147: Biosynthesis of a Depsipeptide Peptidoglycan Precursor by Vancomycin Resistance Proteins VanH and VanA. **Biochemistry**, v. 30, n. 43, p. 10408–10415, 1 out. 1991.

BURY-MONÉ, S. Antibacterial Therapeutic Agents: Antibiotics and Bacteriophages. **Reference Module in Biomedical Sciences**, 1 jan. 2014.

BUTLER, M. S. et al. Glycopeptide antibiotics: Back to the future. **The Journal of Antibiotics 2014 67:9**, v. 67, n. 9, p. 631–644, 13 ago. 2014.

BYSTRYKH, L. V. et al. Production of actinorhodin-related “blue pigments” by Streptomyces coelicolor A3(2). **Journal of Bacteriology**, v. 178, n. 8, p. 2238–2244, 1996.

CHAIN, E. et al. THE CLASSIC: penicillin as a chemotherapeutic agent. 1940. **Clinical orthopaedics and related research**, v. 439, p. 23–26, 2005.

CHALLIS, G. L. Genome mining for novel natural product discovery. **Journal of Medicinal Chemistry**, v. 51, n. 9, p. 2618–2628, 8 maio 2008.

CHANG, Q. et al. Antibiotics in agriculture and the risk to human health: how worried should we be? **Evolutionary Applications**, v. 8, n. 3, p. 240–247, 1 mar. 2015.

CHATER, K. F. Streptomyces. **Brenner's Encyclopedia of Genetics: Second Edition**, p. 565–567, 1 jan. 2013.

CHEN, W. et al. Chromosomal instability in Streptomyces avermitilis: Major deletion in the central region and stable circularized chromosome. **BMC Microbiology**, v. 10, n. 1, p. 1–15, 26 jul. 2010.

CHEVRETTE, M. G. et al. The antimicrobial potential of Streptomyces from insect microbiomes. **Nature Communications 2019 10:1**, v. 10, n. 1, p. 1–11, 31 jan. 2019.

CIMERMANCIC, P. et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. **Cell**, v. 158, n. 2, p. 412–421, 17 jul. 2014.

COURVALIN, P. Vancomycin Resistance in Gram-Positive Cocci. **Clinical Infectious Diseases**, v. 42, n. Supplement_1, p. S25–S34, 1 jan. 2006.

CYPHERT, E. L. et al. Erythromycin Modification That Improves Its Acidic Stability while Optimizing It for Local Drug Delivery. **Antibiotics 2017, Vol. 6, Page 11**, v. 6, n. 2, p. 11, 25 abr. 2017.

DA CUNHA, B. R.; FONSECA, L. P.; CALADO, C. R. C. Antibiotic Discovery: Where Have We Come from, Where Do We Go? **Antibiotics 2019, Vol. 8, Page 45**, v. 8, n. 2, p. 45, 24 abr. 2019.

DAVIDSON, A. L.; CHEN, J. ATP-Binding Cassette Transporters in Bacteria. **https://doi.org/10.1146/annurev.biochem.73.011303.073626**, v. 73, p. 241–268, 9 jun. 2004.

DINCER, S. et al. Nonribosomal Peptide Synthesis. **Scientist**, v. 17, n. 6, p. 12, 30 abr. 2022.

DONALD, L. et al. Streptomyces: Still the Biggest Producer of New Natural Secondary Metabolites, a Current Perspective. **Microbiology Research**, v. 13, n. 3, p. 418–465, 1 jul. 2022.

DORAN, J. L. et al. Mycobacterium tuberculosis efpA encodes an efflux protein of the QacA transporter family. **Clinical and Diagnostic Laboratory Immunology**, v. 4, n. 1, p. 23–32, 1 jan. 1997.

DURAND, G. A.; RAOULT, D.; DUBOURG, G. Antibiotic discovery: history, methods and perspectives. **International Journal of Antimicrobial Agents**, v. 53, n. 4, p. 371–382, 1 abr. 2019.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792–1797, 1 mar. 2004.

EHRLICH, J. et al. Chloromycetin, a new antibiotic from a soil actinomycete. **Science**, v. 106, n. 2757, p. 417, 31 out. 1885.

EHRLICH, P.; HATA, S. Die experimentelle Chemotherapie der Spirillosen. **Die experimentelle Chemotherapie der Spirillosen**, 1910.

FLEMING, A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. **British journal of experimental pathology**, v. 10, n. 3, p. 226, 1929.

FLOSS, H. G.; YU, T. W. Rifamycin - Mode of action, resistance, and biosynthesis. **Chemical Reviews**, v. 105, n. 2, p. 621–632, fev. 2005.

FLOSS, H. G.; YU, T. W.; ARAKAWA, K. The biosynthesis of 3-amino-5-hydroxybenzoic acid (AHBA), the precursor of mC7N units in ansamycin and mitomycin antibiotics: a review. **The Journal of Antibiotics 2011 64:1**, v. 64, n. 1, p. 35–44, 17 nov. 2010.

FOULSTON, L. Genome mining and prospects for antibiotic discovery. **Current Opinion in Microbiology**, v. 51, p. 1–8, 1 out. 2019.

FURLAN, R. L. A. et al. DNA-Binding Properties of Cosmomycin D, an Anthracycline with Two Trisaccharide Chains. **The Journal of Antibiotics**, v. 57, n. 10, p. 647–654, 25 out. 2004.

GARRIDO, L. M. et al. Insights in the glycosylation steps during biosynthesis of the antitumor anthracycline cosmomycin: Characterization of two glycosyltransferase genes. **Applied Microbiology and Biotechnology**, v. 73, n. 1, p. 122–131, 30 nov. 2006.

GIORDANO, D. Bioactive Molecules from Extreme Environments. **Marine Drugs 2020, Vol. 18, Page 640**, v. 18, n. 12, p. 640, 14 dez. 2020.

GREENE, N. P. et al. Antibiotic resistance mediated by the MacB ABC transporter family: A structural and functional perspective. **Frontiers in Microbiology**, v. 9, n. MAY, p. 950, 28 maio 2018.

GROBBELAAR, M. et al. Evolution of rifampicin treatment for tuberculosis. **Infection, Genetics and Evolution**, v. 74, p. 103937, 1 out. 2019.

GYGLI, S. M. et al. Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. **FEMS Microbiology Reviews**, v. 41, n. 3, p. 354–373, 1 maio 2017.

HARRISON, J.; STUDHOLME, D. J. Recently published Streptomyces genome sequences. **Microbial Biotechnology**, v. 7, n. 5, p. 373–380, 1 set. 2014.

HE, W. et al. Identification of AHBA biosynthetic genes related to geldanamycin biosynthesis in Streptomyces hygroscopicus 17997. **Current Microbiology**, v. 52, n. 3, p. 197–203, mar. 2006.

HERZON, S. B. The Mechanism of Action of (-)-Lomaiviticin A. **Accounts of Chemical Research**, v. 50, n. 10, p. 2577–2588, 17 out. 2017.

HILLIER, L. D. W. et al. Comparison of C. elegans and C. briggsae Genome Sequences Reveals Extensive Conservation of Chromosome Organization and Synteny. **PLOS Biology**, v. 5, n. 7, p. e167, jul. 2007.

HINES, K. M. et al. Large-Scale Structural Characterization of Drug and Drug-Like Compounds by High-Throughput Ion Mobility-Mass Spectrometry. **Analytical Chemistry**, v. 89, n. 17, p. 9023–9030, 5 set. 2017.

HOFF, G. et al. Genome plasticity is governed by double strand break DNA repair in Streptomyces. **Scientific Reports 2018 8:1**, v. 8, n. 1, p. 1–11, 27 mar. 2018.

HOSKISSON, P. A.; SEIPKE, R. F. Cryptic or silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. **mBio**, v. 11, n. 5, p. 1–5, 1 set. 2020.

ITO, Y. et al. Genetic Analysis of the Mode of Interplay between an ATPase Subunit and Membrane Subunits of the Lipoprotein-Releasing ATP-Binding Cassette Transporter LolCDE. **Journal of Bacteriology**, v. 188, n. 8, p. 2856, abr. 2006.

IWASAKI, S. Biosynthesis of Enediyne Antibiotics. **Comprehensive Natural Products Chemistry**, p. 557–571, 1 jan. 1999.

JAMES, R. C. et al. Redesign of glycopeptide antibiotics: Back to the future. **ACS Chemical Biology**, v. 7, n. 5, p. 797–804, 18 maio 2012.

JIANG, J.; HE, X.; CANE, D. E. Biosynthesis of the earthy odorant geosmin by a bifunctional Streptomyces coelicolor enzyme. **Nature Chemical Biology 2007 3:11**, v. 3, n. 11, p. 711–715, 16 set. 2007.

JIANG, X. et al. Discovery of a new asymmetric dimer nenestatin B and implications of a dimerizing enzyme in a deep sea actinomycete. **Organic & Biomolecular Chemistry**, v. 19, n. 19, p. 4243–4247, 19 maio 2021.

JUKES, T. H.; CANTOR, C. R. Evolution of Protein Molecules. **Mammalian Protein Metabolism**, p. 21–132, 1 jan. 1969.

KALAKOUTSKIIP, L. V; NINA AGRE, A. S. **Comparative Aspects of Development and Differentiation in ActinomycetesBACTJmOLOGICAL REVIEWS**. [s.l: s.n.]. Disponível em: <https://journals.asm.org/journal/br>.

KIESER, T. et al. **Practical Streptomyces genetics**. Norwich: The John Innes Foundation, 2000.

KIM, L. J. et al. Structure Revision of the Lomaiviticins. **Journal of the American Chemical Society**, v. 143, n. 17, p. 6578–6585, 5 maio 2021.

KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **Journal of Molecular Evolution**, v. 16, n. 2, p. 111–120, jun. 1980.

KINASHI, H. Giant linear plasmids in Streptomyces: a treasure trove of antibiotic biosynthetic clusters. **The Journal of Antibiotics 2011 64:1**, v. 64, n. 1, p. 19–25, 8 dez. 2010.

KOHANSKI, M. A.; DWYER, D. J.; COLLINS, J. J. How antibiotics kill bacteria: from targets to networks. **Nature Reviews Microbiology 2010 8:6**, v. 8, n. 6, p. 423–435, 4 maio 2010.

KOOMSIRI, W. et al. Sarpeptins A and B, Lipopeptides Produced by Streptomyces sp. KO-7888 Overexpressing a Specific SARP Regulator. **Journal of Natural Products**, v. 82, n. 8, p. 2144–2151, 23 ago. 2019.

KOTHARI, A. et al. Large circular plasmids from groundwater plasmidomes span multiple incompatibility groups and are enriched in multimetal resistance genes. **mBio**, v. 10, n. 1, 1 jan. 2019.

KUZUYAMA, T. Biosynthetic studies on terpenoids produced by Streptomyces. **The Journal of Antibiotics 2017 70:7**, v. 70, n. 7, p. 811–818, 15 fev. 2017.

LEE, N. et al. Thirty complete Streptomyces genome sequences for mining novel secondary metabolite biosynthetic gene clusters. **Scientific Data 2020 7:1**, v. 7, n. 1, p. 1–9, 13 fev. 2020.

LEEKHA, S.; TERRELL, C. L.; EDSON, R. S. General principles of antimicrobial therapy. **Mayo Clinic Proceedings**, v. 86, n. 2, p. 156–167, 1 fev. 2011.

LI, T.-L. et al. Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase. **Chemistry & Biology**, v. 11, n. 1, p. 107–119, 1 jan. 2004.

LIU, F.; MYERS, A. G. Development of a platform for the discovery and practical synthesis of new tetracycline antibiotics. **Current Opinion in Chemical Biology**, v. 32, p. 48–57, 1 jun. 2016.

LIU, W. et al. The Neocarzinostatin Biosynthetic Gene Cluster from Streptomyces carzinostaticus ATCC 15944 Involving Two Iterative Type I Polyketide Synthases. **Chemistry & Biology**, v. 12, n. 3, p. 293–302, 1 mar. 2005.

M100 Performance Standards for Antimicrobial Susceptibility Testing A CLSI supplement for global application. [s.d.].

MADDEN, T. The BLAST Sequence Analysis Tool. 13 ago. 2003.

MARTENS, E.; DEMAIN, A. L. The antibiotic resistance crisis, with a focus on the United States. **The Journal of Antibiotics 2017 70:5**, v. 70, n. 5, p. 520–526, 1 mar. 2017.

MCCOMAS, C. C.; CROWLEY, B. M.; BOGER, D. L. Partitioning the loss in vancomycin binding affinity for D-Ala-D-Lac into lost H-bond and repulsive lone pair contributions. **Journal of the American Chemical Society**, v. 125, n. 31, p. 9314–9315, 6 ago. 2003.

MCDONALD, B. R.; CURRIEA, C. R. Lateral gene transfer dynamics in the ancient bacterial genus Streptomyces. **mBio**, v. 8, n. 3, 1 maio 2017.

MCGUFFIE, M. J.; BARRICK, J. E. pLannotate: engineered plasmid annotation. **Nucleic Acids Research**, v. 49, n. W1, p. W516–W522, 2 jul. 2021.

MEIER-KOLTHOFF, J. P.; GÖKER, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. **Nature Communications 2019 10:1**, v. 10, n. 1, p. 1–10, 16 maio 2019.

MIKHEENKO, A. et al. Versatile genome assembly evaluation with QUAST-LG. **Bioinformatics**, v. 34, n. 13, p. i142–i150, 1 jul. 2018.

MORRIS, R. P. et al. Ancestral antibiotic resistance in Mycobacterium tuberculosis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 34, p. 12200–12205, 23 ago. 2005.

MULANI, M. S. et al. Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review. **Frontiers in Microbiology**, v. 10, n. APR, p. 539, 1 abr. 2019.

MUNGAN, M. D. et al. ARTS 2.0: feature updates and expansion of the Antibiotic Resistant Target Seeker for comparative genome mining. **Nucleic Acids Research**, v. 48, n. W1, p. W546–W552, 2 jul. 2020.

MURRAY, C. J. et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. **The Lancet**, v. 399, n. 10325, p. 629–655, 12 fev. 2022.

OBERG, N.; ZALLOT, R.; GERLT, J. A. EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools. **Journal of Molecular Biology**, p. 168018, 17 fev. 2023.

OVERBYE, K. M.; BARRETT, J. F. Antibiotics: Where did we go wrong? **Drug Discovery Today**, v. 10, n. 1, p. 45–52, 1 jan. 2005.

PALUMBI, S. R. Humans as the world's greatest evolutionary force. **Science**, v. 293, n. 5536, p. 1786–1790, 7 set. 2001.

PANDZA, S. et al. Recombination between the linear plasmid pPZG101 and the linear chromosome of Streptomyces rimosus can lead to exchange of ends. **Molecular Microbiology**, v. 28, n. 6, p. 1165–1176, 1 jun. 1998.

PARTE, A. C. et al. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. **International Journal of Systematic and Evolutionary Microbiology**, v. 70, n. 11, p. 5607–5612, 23 jul. 2020.

PEEK, J. et al. Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. **Nature Communications**, v. 9, n. 1, 1 dez. 2018.

PEPPER, I. L.; GENTRY, T. J. Earth Environments. **Environmental Microbiology: Third Edition**, p. 59–88, 1 jan. 2015.

PLATER, R.; STROHL, W. R. Polyketide Biosynthesis: Antibiotics in Streptomyces. **Genetic Engineering of Plant Secondary Metabolism**, p. 61–91, 1994.

QUISTGAARD, E. M. et al. Understanding transport by the major facilitator superfamily (MFS): structures pave the way. **Nature Reviews Molecular Cell Biology 2016 17:2**, v. 17, n. 2, p. 123–132, 13 jan. 2016.

RAVEL, J.; WELLINGTON, E. M. H.; HILL, R. T. Interspecific transfer of Streptomyces giant linear plasmids in sterile amended soil microcosms. **Applied and Environmental Microbiology**, v. 66, n. 2, p. 529–534, fev. 2000.

REDDY, G. K. et al. Exploring novel bacterial terpene synthases. **PLOS ONE**, v. 15, n. 4, p. e0232220, 1 abr. 2020.

REYNOLDS, P. E. Structure, biochemistry and mechanism of action of glycopeptide antibiotics. **European Journal of Clinical Microbiology & Infectious Diseases**, v. 8, n. 11, p. 943–950, nov. 1989.

RICHTER, M.; ROSSELLÓ-MÓRA, R. Shifting the genomic gold standard for the prokaryotic species definition. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 45, p. 19126–19131, 10 nov. 2009.

RISDIAN, C.; MOZEF, T.; WINK, J. Biosynthesis of Polyketides in Streptomyces. **Microorganisms 2019, Vol. 7, Page 124**, v. 7, n. 5, p. 124, 6 maio 2019.

ROBERTSON, J.; NASH, J. H. E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. **Microbial genomics**, v. 4, n. 8, p. e000206, 1 ago. 2018.

ROMANO, S. et al. Extending the “One Strain Many Compounds” (OSMAC) Principle to Marine Microorganisms. **Marine Drugs**, v. 16, n. 7, 1 jul. 2018.

RUTLEDGE, P. J.; CHALLIS, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. **Nature Reviews Microbiology 2015 13:8**, v. 13, n. 8, p. 509–523, 29 jun. 2015.

SCHWARZ, J. et al. Triaging of Culture Conditions for Enhanced Secondary Metabolite Diversity from Different Bacteria. **Biomolecules 2021, Vol. 11, Page 193**, v. 11, n. 2, p. 193, 30 jan. 2021.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068–2069, 15 jul. 2014.

SEYCHELL, B. C.; BECK, T. Progress in the total synthesis of inthomycins. **Beilstein Journal of Organic Chemistry**, v. 17, p. 58, 7 jan. 2021.

SHEN, B. et al. Enediynes: Exploration of microbial genomics to discover new anticancer drug leads. **Bioorganic & Medicinal Chemistry Letters**, v. 25, n. 1, p. 9–15, 1 jan. 2015.

SHIRIAEV, D. I. et al. Nybomycin inhibits both fluoroquinolone-sensitive and fluoroquinolone-resistant escherichia coli DNA gyrase. **Antimicrobial Agents and Chemotherapy**, v. 65, n. 5, 1 maio 2021.

SIMÃO, F. A. et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, v. 31, n. 19, p. 3210–3212, 1 out. 2015.

SIVALINGAM, P. et al. Extreme environment streptomyces: Potential sources for new antibacterial and anticancer drug leads? **International Journal of Microbiology**, v. 2019, 2019.

SOTTORFF, I. et al. Different Secondary Metabolite Profiles of Phylogenetically almost Identical Streptomyces griseus Strains Originating from Geographically Remote Locations. **Microorganisms 2019, Vol. 7, Page 166**, v. 7, n. 6, p. 166, 6 jun. 2019.

SPANOGIANNOPOULOS, P. et al. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. **Antimicrobial Agents and Chemotherapy**, v. 56, n. 10, p. 5061–5069, out. 2012.

STEGMANN, E.; FRASCH, H. J.; WOHLLEBEN, W. Glycopeptide biosynthesis in the context of basic cellular functions. **Current Opinion in Microbiology**, v. 13, n. 5, p. 595–602, 1 out. 2010.

SUDZINOVÁ, P. et al. What the Hel: recent advances in understanding rifampicin resistance in bacteria. **FEMS Microbiology Reviews**, [s.d.].

SUGIYAMA, M. Structural biological study of self-resistance determinants in antibiotic-producing actinomycetes. **The Journal of Antibiotics 2015 68:9**, v. 68, n. 9, p. 543–550, 15 abr. 2015.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, v. 38, n. 7, p. 3022–3027, 25 jun. 2021.

TERLOUW, B. R. et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. **Nucleic Acids Research**, v. 51, n. D1, p. D603–D610, 6 jan. 2023.

THAKER, M. N. et al. Identifying producers of antibacterial compounds by screening for antibiotic resistance. **Nature Biotechnology**, v. 31, n. 10, p. 922–927, out. 2013.

THAKER, M. N.; WAGLECHNER, N.; WRIGHT, G. D. Antibiotic resistance-mediated isolation of scaffold-specific natural product producers. **Nature Protocols**, v. 9, n. 6, p. 1469–1479, 2014.

UL-HASSAN, A.; WELLINGTON, E. M. Actinobacteria. **Encyclopedia of Microbiology**, p. 25–44, 1 jan. 2009.

UREM, M. et al. Intertwining nutrient-sensory networks and the control of antibiotic production in Streptomyces. **Molecular Microbiology**, v. 102, n. 2, p. 183–195, 1 out. 2016.

VALENT, P. et al. Paul Ehrlich (1854-1915) and His Contributions to the Foundation and Birth of Translational Medicine. **Journal of Innate Immunity**, v. 8, n. 2, p. 111, 1 mar. 2016.

VAN DER HEUL, H. U. et al. Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. **Natural Product Reports**, v. 35, n. 6, p. 575–604, 20 jun. 2018.

Vancomycin | C66H75Cl2N9O24 - PubChem. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/14969>. Acesso em: 30 mar. 2023.

VOLFF, J. N.; ALTENBUCHNER, J. Genetic instability of the Streptomyces chromosome. **Molecular Microbiology**, v. 27, n. 2, p. 239–246, 1 fev. 1998.

WAKSMAN, S. A.; HENRICI, A. T. The Nomenclature and Classification of the Actinomycetes. **Journal of Bacteriology**, v. 46, n. 4, p. 337–341, out. 1943.

WAKSMAN, S. A.; SCHATZ, A.; REYNOLDS, D. M. PRODUCTION OF ANTIBIOTIC SUBSTANCES BY ACTINOMYCETES* †. **Ann. N.Y. Acad. Sci. 112 Ann. N.Y. Acad. Sci**, v. 1213, p. 112–124, 2010.

WALSH, C. Where will new antibiotics come from? **Nature Reviews Microbiology 2003 1:1**, v. 1, n. 1, p. 65–70, 2003.

WANG, H. X. et al. PCR screening reveals considerable unexploited biosynthetic potential of ansamycins and a mysterious family of AHBA-containing natural products in actinomycetes. **Journal of Applied Microbiology**, v. 115, n. 1, p. 77–85, jul. 2013.

WEBER, T. et al. Exploiting the genetic potential of polyketide producing streptomycetes. **Journal of Biotechnology**, v. 106, n. 2–3, p. 221–232, 19 dez. 2003.

WHO publishes list of bacteria for which new antibiotics are urgently needed. Disponível em: <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>. Acesso em: 8 mar. 2023.

WIEST, D. B.; COCHRAN, J. B.; TECKLENBURG, F. W. Chloramphenicol Toxicity Revisited: A 12-Year-Old Patient With a Brain Abscess. **The Journal of Pediatric Pharmacology and Therapeutics**, v. 17, n. 2, p. 182–188, 1 out. 2012.

WOODRUFF, H. B. Selman A. Waksman, winner of the 1952 nobel prize for physiology or medicine. **Applied and Environmental Microbiology**, v. 80, n. 1, p. 2–8, jan. 2014.

WU, Q. et al. Production and Identification of Inthomycin B Produced by a Deep-Sea Sediment-Derived Streptomyces sp. YB104 Based on Cultivation-Dependent Approach. **Current Microbiology**, v. 75, n. 7, p. 942–951, 1 jul. 2018.

XU, Y. et al. Crystal structure of the periplasmic region of MacB, a noncanonic ABC transporter. **Biochemistry**, v. 48, n. 23, p. 5218–5225, 16 jun. 2009.

YAMASAKI, M.; KINASHI, H. Two chimeric chromosomes of Streptomyces coelicolor A3 (2) generated by single crossover of the wild-type chromosome and linear plasmid SCP1. **Journal of Bacteriology**, v. 186, n. 19, p. 6553–6559, out. 2004.

YAN, X. et al. Discovery of Alternative Producers of the Enediyne Antitumor Antibiotic C-1027 with High Titers. **Journal of Natural Products**, v. 81, n. 3, p. 594–599, 23 mar. 2018.

YIM, G. et al. Glycopeptide antibiotic biosynthesis. **The Journal of Antibiotics 2014 67:1**, v. 67, n. 1, p. 31–41, 13 nov. 2013.

YOON, S. H. et al. A large-scale evaluation of algorithms to calculate average nucleotide identity. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 110, n. 10, p. 1281–1286, 1 out. 2017.

YUSHCHUK, O.; BINDA, E.; MARINELLI, F. Glycopeptide Antibiotic Resistance Genes: Distribution and Function in the Producer Actinomycetes. **Frontiers in Microbiology**, v. 11, p. 1173, 17 jun. 2020.

ZALLOT, R.; OBERG, N.; GERLT, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. **Biochemistry**, v. 58, n. 41, p. 4169–4182, 15 out. 2019.

ZHANG, H.; TWEEL, B.; TONG, L. Molecular basis for the inhibition of the carboxyltransferase domain of acetyl-coenzyme-A carboxylase by haloxyfop and diclofop. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 16, p. 5910–5915, 20 abr. 2004.

ZHANG, J. et al. Recent Progress in Unusual Carbohydrate-Containing Natural Products Biosynthesis. **Comprehensive Natural Products III**, p. 336–392, 1 jan. 2020.

ZHANG, R. et al. Characterization of replication and conjugation of Streptomyces circular plasmids pFP1 and pFP11 and their ability to propagate in linear mode with artificially attached telomeres. **Applied and Environmental Microbiology**, v. 74, n. 11, p. 3368–3376, jun. 2008.

ZHOU, Z. et al. Genome plasticity and systems evolution in Streptomyces. **BMC bioinformatics**, v. 13 Suppl 10, n. 10, p. 1–17, 25 jun. 2012.

ZIEMERT, N.; ALANJARY, M.; WEBER, T. The evolution of genome mining in microbes – a review. **Natural Product Reports**, v. 33, n. 8, p. 988–1005, 27 jul. 2016.

Appendix 1: Culture media used in this work

Trace salts solution: (0.1 g FeSO4 x $7H_2O$ / 0.1 g $MnCl_2$ x $4H_2O$ / $ZnSO_4$ x $7H_2O$ / 100 mL dH_2O).

ISP2 (Cultivation, screening of resistant isolates and bioactivity assays)

Per Litre:

- 4 g Yeast extract powder
- 10 g Malt extract powder
- 4 g Dextrose
- 1 L dH_2O
- 20 g Agar (for solid media)

ISP3 (Bioactivity assays)

Per Litre:

- 20 g White Oats
- 1 mL Trace salts solution
- 20 g Agar

ISP5 (Bioactivity assays)

Per Litre:

- 1 g L-asparagine (anhydrous)
- 10 g Glycerol
- 1 g K_2HPO_4
- 1 L dH_2O
- 1 mL Trace salts solution
- 20 g Agar

ISP6 (Bioactivity assays)

Per Litre:

- 15 g Bacto-Peptone
- 5 g Proteose-peptone
- 0.5 g Ferric ammonium citrate
- 1g K_22HPO_4
- 0.08 g Sodium thiosulfate
- 15 g Bacto agar

- 1 g Bacto-yeast extract

Mannitol Soya Flour (Bioactivity assays)

Per Litre:

- 20 g Mannitol
- 20 g Soya Flour
- 20 g Agar

R5 (Bioactivity assays)

Per Litre:

- 103 g Sucrose
- 0.25 g K_2SO_4
- 10.12 g $MgCl_2.6H_2O$
- 10 g Glucose
- 0.1 g Difco Casaminoacids
- ml Trace element solution
- 5 g Difco yeast extract
- 5.73 g TES buffer
- ml KH_2PO_4 (0.5%)
- 0.4 ml $CaCl_2.2H_2O$ (5M, 3.68%)
- 1.5 ml L-proline (20%)
- 0.7 ml NaOH (1N)
- 0.75 ml Required growth factors for auxotrophs

YEME (For biomass formation for genomic DNA extraction when required)

Per Litre:

- 3g Yeast Extract
- 3g Malt Extract (Oxoid)
- 5g Peptone
- 10g Glucose
- 340 g sucrose (34% final concentration)

Appendix 2: Annotation and genomic sequence of the minor plasmid found in BRB040.

```
COMMENT     Annotated with pLannotate v1.2.0
FEATURES             Location/Qualifiers
     CDS             complement(10902..11288)
                     /note="pLannotate"
                     /label="rifK (fragment)"
                     /database="swissprot"
                     /identity="69.0"
                     /match_length="33.2"
                     /fragment="True"
                     /other="CDS"
     CDS             4082..4483
                     /note="pLannotate"
                     /label="lolD (fragment)"
                     /database="swissprot"
                     /identity="50.4"
                     /match_length="58.8"
                     /fragment="True"
                     /other="CDS"
     CDS             5458..5649
                     /note="pLannotate"
                     /label="macB (fragment)"
                     /database="swissprot"
                     /identity="50.0"
                     /match_length="9.9"
                     /fragment="True"
                     /other="CDS"
ORIGIN
        1 cggttctcgt gaacaaagcc acaccagcca ggccagggtc aacgagctgg cttcttcgtc
       61 cactggctgg cccgaggact cttgcgagtg aggcgtcggg tcatcgtaat gatgagtgct
      121 cagttcaggt gtgcttcgga atgctgcggg aggcgtgcgt tgcggcgggc gttcatgcac
      181 cagccgatcg tgcgctcgac ttgccaacgg ccgggcggga tgacgaagcc cttggtgccc
      241 ttggggcggg agacgatgcg caggtgagtc ggagacggtc gcgggcccag tccacgaggt
      301 cgctgaacgg atagccgccg cctctcgtcc aggtgatcgg cgtcaggttc tccgcggcga
      361 cgacgtcgcg gacgtgccgg atgccccggc cgacctgcga gcgggtcagg cgggtggcgg
      421 ccatgagctg gtaggtgtgc aggccggcgg ggcgtgcttc catcagggcg acccggacca
      481 ggtcgccgta gtgcttggcc aggggccggg cctcccggca ccgcggcacg gctactcctc
      541 accctggagt agctggacca gctgctcgtc cagggcgaac tcatcgttgt cgaggactcc
      601 ttcgagcccc tcgtgcttgt cgaccaggat ggtgagcagt tcgttccaga cctggtcgcg
      661 ggcgtcgttg tgcggggcgc cctcggccgc gtcgaaggcc tcggcccagt cggcggcggt
      721 cagccggatg tcggaccagt gggtggtgac ggtcatgccc tcggtgggcg gctcctcgta
      781 gaagcggacg gccctctcga tgcccttcac caggtcggcg gaggacttca gacgggccac
      841 ctcggggtcg ctctccaccc cggccgccgc cccctcggcc acgaagtccc gcaggacgca
      901 ggtctggacg ccctcctcgc ccaggctggg cagcacgtcg gagacgtagg ccaggtacgg
```

961 ccggtgcggg ccgacgaaca gcacgccgcc gcgccggtgg ccgaggcgcg ggtcggagtg
1021 caggaggtag gcggagcggt gcagggcgac gacggtcttg ccggtgccgg ggccgccgtc
1081 gacgacgagc gcgccgcggg agcccgcgcg gatgatggcg tcctggtcgg cctggatggt
1141 gccgaggacg tcccgcatcc gctccgagcg gtcgctgccg aggctggcga tgaaggcgga
1201 ctggtcgtcg agcgaggcgt ggtgcccggc gaagtcgtcg gcgacgaaca cctcgtccca
1261 gtagtcgccg atacgaccgt cgttccagcg gtacctgcgg cggctcgcca ggcccatcgg
1321 gttggcgtgg gtggccccga agaacggctc ggcggcaggg gagcgccagt cgacgagcag
1381 tctgcggccc tcgctgtcgg tgagtccgag gcgtccgacg tacacgggct cggggtcgtc
1441 ggcggtgacc atgtggccga ggcacaggtc caggccgaag cggcgcaggg tgcgcagccg
1501 ggcgctgagg cggtggatct ccgtgtcccg gtccatcgcc tcctggccga tgccgccggg
1561 cgccctgagg gtggcggcga gggtggcgga cagttcggcg atcgtctgct ccaggcatgc
1621 cgcgatggcc gcgaagtgcc gctcgtcggc ggcgatcagg gccgggtcgg ccttggcggc
1681 gaggcggtcg gggaggtcgg gaccggggtg ccggtgctgc ggatcgacgt ggggacgttg
1741 gcgtcgggga gggcggcgcg gatgaccggc tcggcctcct tcatgatctc ttccggggtg
1801 cggtagttga tgctcaggga ggccatctcg atgcggtcga ggcccacccg ccgcagcctc
1861 tcctgccagg actccgtgaa gccgtgccgc gcctgggcgc ggtcgccgac gatcgtgaag
1921 ctgcgggagg ggcagcgcag cagcagcatc tgccactccg cgtcggtcag ttcctgcgcc
1981 tcgtcgacga cgacatgggc gaagggtccg gccagcaggt ccgggtcggt cacctgcaag
2041 gcgctctcgt cgaccagggc gtggtggaag tcctcgccgc tgagcatcgt caccaggccc
2101 atgccgtact cgtcgtcggc gaccgcgctc aggttctcca cgacggtggc catgcgttcg
2161 cgttcggccg cggcggtggc ctcgttgcgg cgcttgcgca gggaggcctt cgggtcgccg
2221 aggcgctggc gggccgcgtc caggaggggc aggtcggaca ccgtccaggc ctggccgtcc
2281 gcgcgctgga gtttccgtac ctcgtcgcgg tccagccagg gggcgcacag ccgcagatag
2341 gccgggaccg tccacaggtc cgagaccagg tcggccgcct ccagcagggg ccacgcccgg
2401 ttgaacgcgg tgagcagctc cctgttctgg cgcagggacc tgcgcagagt ctgctgccgc
2461 gcggaccctg ccgatctacc gtcggacagt ctcgcgctcg tcctgggtga actcctgccc
2521 gcgcagctgc ggaacgagcc ggcccagggt ggccccgaag tcttggcagg agccggtcag
2581 gtcgaggtcc tcgatggcgt gctcgatctt ccggacggcc ggtgtgcgct cccggatccg
2641 ttcgcccgtc tcgttggagt tgtcgaactg cgcacggttg acgagcatcc gggtggtgtc
2701 cgtcctggtc aactctgtgg cgacaccgag gccaggatgc gatggaccgt gaacgacacc
2761 ccgtcctttc ggcgttcctt gggccgccgg ttggcggtgt agctccagtc ctccaccgtc
2821 cgccgttcga caccgatgtc gcccgcgaat atctgcagcg actccgtcac cgtgaacagg
2881 tcgtccctgc cgttcggcat cgatccgccc accgaacgca tcggctcgat ctccagcgcc
2941 atcctgatcg ccctcgttac aagccccgac tcgtactcgt actcgtgccc agccaccctc
3001 gaccaccgga ccaccacgtg aactcaatgt cgcggtagac gggttgacgg ggtcagttgg
3061 ggggatggga gcgcggccag agcgttgtgg gcgggcggga cgttgtagac gccaccgccg
3121 atagccgtgt ccgttgtcgc cttgaccgag cgcggctccc atgagagggg caggggaatc
3181 ggggggtgcc tctatgtctt tcgtcaagcg gcgcgcggcg tttcgggttc gtagaaggtg
3241 ctgtcgcgga gcatcgcgac agcacattgg tgcggtggcc gaccaggcaa agggggctgg
3301 ggtgtgggtc tttccgaaga ctgcgctacg ccacccgctc agcgatcgga ccatctccgc
3361 gagacctgcc ccggcaccag ccaggtcgca actcgggggg cgggactcag gaaggggtgc
3421 acgcctcggc gatcgatacg ctgttctcgt agaggtgatg ccgggtgatc cggccctcct
3481 cgacggtgag gcgcagcgcg aaggggccct cgaaggactt ccccgtcgcg cgcacggtcc
3541 ccgagaggtg ccccgtcagg acggcgtcgg tgccgtcgac gaggaacgtg tcgacggagg
3601 cccgtgcatc ctcggccacc gtgtgctcgg ccagctccgt gaactgggcg gcgcattcgg
3661 ctccggtgga gcgtggccgg atccacggga cggcagggtt ctcggcgagt agccagtcga

3721 cgtcgtccgc gaagagcgcg acgagccgct cgacgtcccc ggccatccgg gcggcgagga
3781 acgactggac gacgtggcgg gtggtgtcgg cgacggaact cgcggtgatg tcagcgttct
3841 tgatgggcat ggttgctcct cgcgccttag atgccggaat ccgcgtcggg atctcgttga
3901 catgtatcac gctgccgcag caggggagcg tggtcgatta cggcagaggt aatggccggg
3961 tggagggctt ctacgtccgt cgtcaagccc tggcctgtcg gctgtgttgg gtgtgtttgt
4021 gttgtcggtg gtggtgggtg tccggctcac cgttcacgct gtcaggcccg tgtcggcggt
4081 atcggttagt gaccggcagc gcacgtcgct gcgcggaagc agcatcggct tcgtcttcca
4141 ggccttccac ctgcttccct atcggaccgt ggaggagaac gtgatgctgg cggaggtcta
4201 ccgcaggcag cgcacgggcc aagggcgccg ccaccgtgcc gaacggccg gagcggccct
4261 ggaccgggtg ggcctgtccc accgcaacgg tttccggccc gaccgtctct ccggcggcga
4321 gcgccagcgc gtcgccatcg cacgggccct gatgagcgaa ccggccctgc tgctgtgcga
4381 cgaaccgacc ggcaacctcg acagcgagaa cacgcactcg gtcctggaac tcttcgacca
4441 gctcggcgcc cagggcatga ccctcgtcgt gatcacccac gacgatgccg tcagccgacg
4501 tgccggccga cgggtccgta tccgcgacgg acggctcatg gaggaacccg catgaaacgc
4561 cgtcaccgcg ccgccgctgc cggccccggc aagcgggtgg accgtccccg aacccatccg
4621 aaggacctgt ggacggaggc actgtcggga gtactggccc gcccagtccg ctccgcgctg
4681 accaccttgg gcacggtgct gggtatcggc acgctggtcg tcacggtcgg ggtggcgtcg
4741 acggccgctc accagatcgt cggccggttc gatgcgctca ccgccacgtc ggtcaccgtc
4801 gtggtcccgc cgccgcctcc ggcgtcggag accgtgccgc tggtggactg ggacgacgtc
4861 ggcgccgtcg gccggctggc cggggtggag tcgtccgctg cgctcgcgca aacggacgaa
4921 ctggctgccg tgcaggtgcg ggccaacgac gtggtcacgc ccggtggcct gtccggccag
4981 acgctggcgg tcgtcgcggc ttcgtcgtcg ctgccggcgg cggtgcgcgg tgaggtcacg
5041 gccggccggt tctacgaccg gggtgacatc gcccgcagtg ccagagtggc ggttctgggc
5101 gaccaggccg ctcgccttct cggtgtcacc cgggtggagg acgcgcccgc ggtcttcttc
5161 aacggccttt ctttcaccgt gatcgggatt ctcggcggaa tcgagcgtga gcagcgcctg
5221 tccacctcgg tgatcctgcc gcccacgacg gccagggacc gtctgggcct ggacacggtg
5281 aaccgcgtcc tgatcaacac ctccctcggt gccgcgcagc aggtcgccag gcagtcaccc
5341 gtcgccctgg cacccaacag cgcggaggca ctcaccgtcc tcgctccgcc cgatctgtcg
5401 aaggcacgca agggtgtcca gggcgacgtc aacggtctgt tcctggttct cgggctcgtc
5461 tcgctgattg tgggtgctct cggtatcgcc aacgtcacgc tggtcaccgt gatggaacgg
5521 gtcggagaga tcgggttgcg gcgtgccctg ggcgcctcgc gccggcagat cgccgctcag
5581 ttcctcctgg agtcgacgac gatcgggctg ctggggggcg tgatcggtgc cgcgctgggg
5641 gttgtcgtcg tggtggccgt gtcggcggtc aaggagtgga ctcccgtact cgacggccgg
5701 cttgccgtcg gcgctccggt cgtcggcgcg gtggtcggcc tgctggcggg tgtgtatccg
5761 tcgctgcggg cttcccggat ggaaccggtc gaggcgttgc gagccccgtc gtgaagacgg
5821 ccctcgggga ccggcgcggg ggcgtactgg acggcccgcc tgccgcccgc aaccgcgttt
5881 cccgtgccat gagtttcgct gatcgacggt gaccggggac tgatacggcc gcggccgcgg
5941 gaggacgaag ggggacatcg ttctcggcgt cttcgggtgg gaacggcggc ccgtgtgaga
6001 gtgaggtgga tcactggaac gcgtccgggg gcgtggagag aacaaggtgt gaggaactca
6061 gacaacgacc tgcggcgtcg tgtgcgcgcc ggtgaccgtg aggcgttcgc cgagctgtac
6121 gaggcgttcg ccggcgcggt gtacaaccat ggtctgcggc tgaccggtga ctggtcggcg
6181 gcggaggaca tcatgtccga gactttcctg gccgcgtggg gcaaccgtgc cgcggtgcac
6241 gaggaggggg gcactctgct gccgtggctg ctgggcatcg ccacccacaa ggcggacaac
6301 gcacggcgca gcgtgcgccg gcggcagctg ttcctgacgc gccggccgca gcccacggcg
6361 gtggaggact tcgcaccgga ggcggccggg cgtctggacg acgtgcgccg gctgcgtgcc
6421 gttcaggcct ccctgggccg gctgcggcgc cacgagcggg aggtgctcgc cctgtgcgtg

6481 ttctccggtc tggattaccg gcagacggct caggcgctgg gcatctgcgt cggcaccgta
6541 cggtcccgtc tgtcccgcgc gcgcaagaag ctcgcccggc ttgccgagga gaacatggaa
6601 ccgcccaccg tccgcggaga gatgacaggt gcggctgcac tcgcggccgt gcccgtgcgg
6661 aaggaactca catgaacggc aggccgaacc cggcttcttt cgagccctcg gacgagcgcg
6721 aggagattgc ccgtctgctg ccggccccgg ccgactggga ccttccgcgc gggcggcacc
6781 tccatcacaa ggacctcttg atgcgacaca tcgaccgcga ccaggccacc accgccaccc
6841 ccgccactgc ccccgccgtc actgcccccg cccccgccgc cgccacggcc cccgccgccc
6901 ccgccgttgc cgcggttgcc cggccggggc gccgacggct gcgccccgta ctcctggcgc
6961 ccgtgaccgc gctggccctg gccggtatcg tgaccgcggc cttcgccctg aacgacggcg
7021 accgcaagcc cgccgccggc acgtcgtccg acgcggccgc ggacatgcag cccgtctccg
7081 ctctcctcgg gcagatctcc gacgccgcgg gcaagcgtga cactcccacg gtgcgggacg
7141 accagttctc ctatgtcagg gagaaggtcc gcgaagccga cctgtccagt gggaaggccg
7201 tcgtcggccc gctgcacgac tacgagctgt ggctctccca ggaccccgag cccgtgcaga
7261 agcagggcct gatccgcagc gacggcaaga ccatgcccct caacgccgag ttcggcgaca
7321 cgaacggcac accggccggc ttcagccgcc ccacctacca atggctggcc tcgctgccga
7381 ccgaccccgg ccggctgctg gaccacctgt acgccaaggc accgcaggcc gagaacaggg
7441 aacgcgacca ggccgtcttc gagcagatcg gtgacctcct cgaccaggtg atgccggcac
7501 gtacggccgc cgcgctctac cgtgccgcag cgaagatccc cggcgtcacc gagacccctc
7561 gggcccagga cgccgtcggc cgccggggcc tgggcatcgc ccgcagcgac acgaagaacg
7621 gcgtcgtcac cgaatgggtc ttcagcacac aggacttcac ctacctcggc cactccaccc
7681 gcctgaccaa ggacacgccc tacgccgagg cgggaacgct cctggacagc tcggccctcc
7741 tcaagagcgc ggtcgtcgac aagcccggcc ggcttcccgg cagcgagaag aagtaggaac
7801 ggccggactg gtcggtcctg cggctgtcgc ccggtccgcc catgctccgc acgtggggct
7861 ggaggaccgc ggtgttctgc cggacgcgga gcgggagcgg gaacggcccc ggcccgtccc
7921 gccggagcgg tacgggccgc ggaaggacgg cgggtccacg cgtgccgtcg gcgctggcgg
7981 ccgacgccgt cggcgagata gatcgacggg ggcatcccgg tgagccccac cagtgcgcgc
8041 cgctctgccc cgcagccgcc gcaccggccg gcctgttcgc ggcatacggg ccggccctca
8101 ctgtgccgcc tgcgccccca ccggcatccg ctcctgcacc gccgccccgg ggcccgcctg
8161 ccggccgtgc cgcccacagc gcgtggaacg gcgcgggaaa cacaaccagc cgaactcccg
8221 gccccgcgta cgccctccca cggcggcgcg ggctcggcga acgcctcctg cagggcggca
8281 gccggcaacc gaccaccgtg gtgccgcggg tgccggtgtc cggccagcca ggccacatcg
8341 cgtcggtaga tcgggtcggc aggcccaagc ccccacgacc gccggcccgc cgccgtacag
8401 atctctccga ctaccgcagc gagagaccgc tgccaccgag tcagctctcc gctccctgtg
8461 cagtcggcga gcaccccttg gccatcggcc ggccggagca tcggctgagg agcatgcgcc
8521 cgcaccccgc tcggcacatg ccaccgcagt ccaagcggcc gggcgttccg gcccgctgac
8581 gcgtggattc gcggtcgagg agcaccccat gcagtcacat cgctgcggac ccgtagtgca
8641 ccagcggtcc ggtggtggcc gaccaccacc gtcccggcgc ccgccgcttg ccgttgcggg
8701 ccgggaaggc ttcagcggcg gacagttttc gaagaccacg tgggcggccg cccatgacca
8761 cctcggacac tcagtcgccc cgcccgcagt ccggaacagc acgtccaccg aatcggcttc
8821 ggaccgccac gccctcgctc cgccgcgccc tcgctccgcc cacccagcat cctgcataca
8881 gaaccaaacg gcttatcggc ccgtccggta ctgggcgcgt gccgcatcgg cgaacatcag
8941 ctgcgcccgg accgttctgc agcccaggtg tagcccagct ctgcgagggc ggcccgctga
9001 tccgcatcca gccggtcacg cctggccttt tggttactca agaacacacc cagcctgatc
9061 gccgaaccgt caggcagctc ctcaacgtgt gccctcccga cgacggttct gccttctcgc
9121 tggatgtact gggtcagggc tgtcaggccg cgttggaacg cgtgctgcgc cgtggtcggg
9181 cggcccttcg tcgcacgctg gggcgccggg gcgggtgccg gcgtctcagc gggctgtacg

9241 cccagcgctg tcaaccgctg ctgctgtccc tcggagagct ctacccaact gcggcgttgc
9301 cgctgaagcc atttgccgat gtcgtcgccg tccatgagga cgccgggtgc gatgtcgggc
9361 aggactccgc cgggttcgtc ggcggcgagg tgggcgaggt ggcggtagtg gcgttgccag
9421 ttgagtggcc aggggcagtt ccagtcgggg tcaatcgagg ccagctgcgc gtcccgtgtc
9481 tgcgcccgtt cggggttctt gccgaggccg cccttacgcc tgaggttcgc cagtagctga
9541 ccgatgggga caggcttgcc gtcggcggga tcgccccaca gggcgtcctg cctcggggcc
9601 atgtgcccgt gcgcacggtg gtacgagcgc agcaccgcga gcttggtctc ccaggcttct
9661 tcgccgggct cccagaccgt gccggcctcc ggcgcgtcga gcagcgtctt gcgccgctcc
9721 tcgagctccc cggcccgccg cgcttttcgc tgctggtgca cccaccgccc cagcggaaag
9781 gccctcgtga cgccgacctc gacctcggtg tcgtagggga cggcgtgcag gccggtgatc
9841 cggtgctcct tccgccaccg cagcagcgcc tgatagccct ccagccagac cagcgactcc
9901 ggccggtaga cccgggtgcg caagaacgcc gcgatcgtcg cggcgtcgcg cggggtggag
9961 aagtgcagca gggccgactc gcccgcaccg tcggtgccgt cctgctcctg gtcgcctccg
10021 ccctctccgc cgtcggccgc gatgatccgc ccgtcctcgt cgcgccggat gtggaccctg
10081 cgcttgccgc tggtgagcgc acgcgaggcg agctgctcga ccaggcgctc gtcatgcgag
10141 cgcaggccct gcagcacggc cacgagcggc ttgaagctcg cgctggcgac catgtcggcc
10201 gggtcctcgc cgggctccag gaacaccggc acgatgatcc gcgcgacctt cgtggaaccg
10261 tcccggttga gccgaagagc ccggccgatg ttctgcacga tctcgacctg cgagccgcgg
10321 gtgtcggcga agcagacggc ctcgactccc cgttcgccgg tgatgtccac gccctcccca
10381 agcacgcgac agctggcgag gaacgcccgg tgcacgcgtc ggctctcggc gtcgatgccg
10441 ttggcgaact ggtgcagggc ttcgcgtcgt tcggtgacga ggtggtcgcc gcacagccac
10501 gccgaccaca cccgctccgg cggcacgtgg cgtcctgctt ccagctcgta gaacccagcg
10561 ccggtcgacg acgcgggcag cttctccgcg gccgccaggt cctcgccgga ggcgtcgcgc
10621 gcgtacagcg cggctgcggt cttcggcagt ttctccgcga acgcggcggc ctcctcgacc
10681 ttctggtgga acgtcatgac cgtacgcagg ttgtacgccg ccgcgtgctc caaaagcgcg
10741 gtctgcagca acgccaggcg ccggccccgc cgcgcctcct ccgactcccc gacgatgggg
10801 gaggggtcgc ggatctccag gacgtcgatc tcgaacccgg cgaggatgcc ccgctcgatc
10861 gcctcggaga gcccgagctc ggcgccgggg agccaggcct tctggaagct gaacgccgcg
10921 atggagccca gctcaccggc cttgcggccg gcccactcgg cgccgtgcgc gtgcgcggcg
10981 tcctggatca gcggcacccc gctgtccgcg gagaccttgc cgagcgcgtc catgtcggcg
11041 aagtgacccg ccatgtgcac cggcatgatc gcgctggtcc gcggggtgat cgcggcctgg
11101 accgccgccg ggtccaggca gtacgtctcc gggtcgacgt cgaccggtat cgccaccgcg
11161 cccaggcgct gggcggcctg cgaggaggag atgaaggtga aggcggggac gatgacctcg
11221 ctgcccccga agatgccgtc cgggtcgtcg gccatgctcg cgagctccac ctccatcccg
11281 cccgcgcccc tctgcgggcg gggcggggcg aggatccgcg gggtggcggt gaggtagagc
11341 cggaaggcgg cggggatgcg ggcgttgtcg tggatcgccg cccacggacg gccaagatca
11401 cccgcggttg agtgggcctc gtccacgatg gcgaggtcgc gctccctgcc cgccagagcg
11461 gcttccaccg gcccgcgaac cgtctgctgg cccgtcgggt cctccgggtc ctcgcggtct
11521 accagggagg cgtaggtggc gaacacgacc acgggcccgg acccggccca gagcgcgagc
11581 cggatcgcat tcgtggtggc gcgcaccccc agcgcgttca gcaccggatc gttctccagc
11641 gagcacaccg ccaccatcgg cgcccgatga ccgaccgccc gccacgcctg cgcggtctgt
11701 gcgagcaggt ccaaggtggg cacggtcacc aggatccgcc cgcccgggaa gcactccagc
11761 gcgctcgcgg cagccgtgat cgtcttaccg gatcccgtgg cggacaccag cgtgccccgg
11821 gccccctgcc ccggcacaga agatcttgca taaattctga cccattctcg aatggccgtt
11881 ctctgggtca cctgatgttc ccggagacta atcacggatt tcgacctttc ctttttccgg
11941 gcgacatcag gttcttccac cacgtgcccg tgcccgaagt gaccgtgctc tcgtcgaggg

12001 ccggagcggg gcaggggtac ctgtcgatcc cgccgcgcgc gggcagctcg tcctcgaagt
12061 cacaggcggc tacagctttg aagccgccac gccacagcgc ttcgagcacg gccgtgatgg
12121 cctgcaccac ctcgtccggt gttccggcgt gggcctcgca ctccagcatg gtcggccact
12181 gaaggaagtc ggaggagtgg gcgtgccggc cggtggcgta gtcgttgtgc tcgccctcca
12241 cgcgcaccgg cccgacggtc agctccgcac cgccgtcctc aacgaggccg agggcgtctc
12301 gaagccgacc cacgacccgc ggggcggaca cggcatccac gtagatcgag gcgtcgctgt
12361 aagcggagtt cgggccgagg ggcgaagacc cgcgatcatt cgtcacgcgg cgctccctgc
12421 ggatctggcc ggggtgaggg agagggagcg ctggagcagc cgcttggccg gttcgtggtg
12481 gtagtactcg gcgatctgca gggggagtct cgccgtcgat ggccggcagg cgcggatcgg
12541 caccgtaggc cagcaggatc gccgtggagg cggtgttcag cgggtaatcc gcctgcacat
12601 ggccgtcgcc ttccaggtcg atcgcatgag tcagcagtgt gtgcccgaag cacacctcgc
12661 ccgggtccgc cccagcgtcc aggagcacgg cgagggtctc gtagtcactc atctccacgg
12721 ccctgtgcac gggcgtccac gactccacct cagcggaaga cccgcgatcg cgtccacgcc
12781 cgctccctgg aatttctgcc ggagtaaggg agaggaagcg atggagcagc cgcatgggcg
12841 gctcgtggtc gtagctctcg gcgaagctgc agcggagttt ccccgctggc agtcggcagg
12901 tgcgggtcgg ctccataggc cagacgatcg ccgtgctggc cgtgtgcagg ggataccccg
12961 actgcaggtg actgtcccct tccatgtcga tcgcgtgagt cagcaacgtg agctcgaaga
13021 acacctcgtc cggatccgcc ccggagtcca ggagcatgga gagggtctcg tagtcactga
13081 tctccacggc ctgatgcgcc ggtgtccacg gctccacctc aggccaagtg ctacgccaca
13141 gcgcatcgga acccatagcg cccttccctt ctccaccggc cttgctgggc cgtggctatc
13201 acatcagtgg acggaccacg ggcataagac gccgaagctg cagcggcccc cgaccctgct
13261 atcggtgtcc tcgcattacg ccaattctac atggtgcatc acatcgagag acggactcgc
13321 gaaaccctcg acagagcccc cactcaagat cccgtccgac ccccggtgcg catcgcgcat
13381 cgcgggtaca cggtgccgca tggcaccgaa gaccgaacgc ctccagttcc tgcgggccac
13441 gcggcaactg cgccgtatcc gctccttcta cgctgcggcc atcctcctgt ggggcagggt
13501 ccacggcctg gaccggctgg cagaccccaa ggagccggca aatgtgcccg ggccgtcctc
13561 ctcctggtcg tcttcaccgc cctgctcgtc accgcgaccc tcgccctgcg gcgcctcgcc
13621 gttcccgcaa caggccggcc cgggcaccac gctgccccgc acaagctgcc gagcacccgc
13681 cgttacgctc ccgcctgagg gcccatcact cttcccacga gtcaatctgg gcccggactc
13741 cgcccgcagg gggcggagtc cacggtctgc tgacagcccg ctaccggagc actacaccct
13801 gcccctgcag gcgctcgggc cgaaggcccg ttgacgggcc cgcgcagcgg gtccgtcctg
13861 tggcgcacag cgccactagg tggctggagc acagcgcagg ccaccgcagc tccgggagcg
13921 cagcggacgg agcgtcacca acccggcgca gccgggttgg tgagctggac gcgcagcgtc
13981 cagccagcgc ttccgcggag cggaagcgca acacccgcgc gcagcgcggg tgttcgcttg
14041 cccatcgcgc agcgatgggg tacccgctcc gcggg

Appendix 3: Biosynthetic gene cluster found on the giant BRB040 plasmid (region 10) on the chromosome.

Identifier	Product	Length		Function
		Nucleotides	Aminoacids	
ORF 1	Transcriptional regulatory protein EmbR	831	276	regulatory
ORF 2	A-factor receptor protein	627	208	regulatory
ORF 3	Propionyl-CoA carboxylase beta chain	1593	530	biosynthetic-additional
ORF 4	hypothetical protein	210	69	other
ORF 5	hypothetical protein	987	328	biosynthetic
ORF 6	NAD(P)H azoreductase	870	289	other
ORF 7	A-factor receptor protein	669	222	regulatory
ORF 8	12-dehydrotetracycline 5-monooxygenase/anhydrotetracycline 6-monooxygenase	1542	513	biosynthetic-additional
ORF 9	hypothetical protein	453	150	other
ORF 10	hypothetical protein	312	103	other
ORF 11	hypothetical protein	459	152	other
ORF 12	Nucleoid occlusion protein	1149	382	other
ORF 13	hypothetical protein	1164	387	other
ORF 14	hypothetical protein	627	208	other
ORF 15	hypothetical protein	549	182	other
ORF 16	hypothetical protein	546	181	other
ORF 17	hypothetical protein	570	189	other
ORF 18	Isochorismatase family protein YecD	600	199	biosynthetic-additional
ORF 19	hypothetical protein	1245	414	transport
ORF 20	hypothetical protein	180	59	other
ORF 21	hypothetical protein	552	183	other
ORF 22	hypothetical protein	384	127	other
ORF 23	Redox-sensitive transcriptional activator SoxR	441	146	regulatory
ORF 24	hypothetical protein	354	117	other
ORF 25	hypothetical protein	906	301	other
ORF 26	hypothetical protein	1095	364	other
ORF 27	hypothetical protein	1371	456	other
ORF 28	putative HTH-type transcriptional regulator YfiR	624	207	regulatory
ORF 29	hypothetical protein	1017	338	other
ORF 30	ECF RNA polymerase sigma factor SigG	1029	342	regulatory
ORF 31	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial	768	255	biosynthetic-additional
ORF 32	hypothetical protein	972	323	other
ORF 33	hypothetical protein	477	158	other
ORF 34	hypothetical protein	894	297	other
ORF 35	hypothetical protein	375	124	other

ORF 36	hypothetical protein	321	106	other
ORF 37	hypothetical protein	615	204	regulatory
ORF 38	putative protein YesE	426	141	biosynthetic
ORF 39	NAD(P)H azoreductase	876	291	other
ORF 40	Nucleoid occlusion factor SlmA	579	192	regulatory
ORF 41	hypothetical protein	447	148	other
ORF 42	hypothetical protein	426	141	other
ORF 43	hypothetical protein	1341	446	other
ORF 44	hypothetical protein	387	128	other
ORF 45	hypothetical protein	261	86	other
ORF 46	Actinorhodin polyketide putative beta-ketoacyl synthase 2	1248	415	biosynthetic
ORF 47	Actinorhodin polyketide putative beta-ketoacyl synthase 1	1266	421	biosynthetic
ORF 48	Linear gramicidin dehydrogenase LgrE	807	268	biosynthetic-additional
ORF 49	Linear gramicidin synthase subunit D	1890	629	biosynthetic-additional
ORF 50	hypothetical protein	777	258	biosynthetic-additional
ORF 51	hypothetical protein	360	119	other
ORF 52	hypothetical protein	1143	380	other
ORF 53	putative HTH-type transcriptional regulator	903	300	regulatory
ORF 54	hypothetical protein	402	133	other
ORF 55	hypothetical protein	1524	507	other
ORF 56	hypothetical protein	1113	370	other
ORF 57	hypothetical protein	120	39	other
ORF 58	putative D,D-dipeptide-binding periplasmic protein DdpA	1548	515	transport
ORF 59	Pyrethroid hydrolase	852	283	biosynthetic-additional
ORF 60	hypothetical protein	651	216	other
ORF 61	Iron-sulfur cluster carrier protein	927	308	other
ORF 62	hypothetical protein	864	287	other
ORF 63	hypothetical protein	780	259	other
ORF 64	hypothetical protein	576	191	other
ORF 65	Membrane protein YdfJ	351	116	other
ORF 66	hypothetical protein	1080	359	other
ORF 67	hypothetical protein	1593	530	other
ORF 68	hypothetical protein	381	126	other
ORF 69	hypothetical protein	630	209	other
ORF 70	hypothetical protein	630	209	regulatory
ORF 71	hypothetical protein	972	323	biosynthetic-additional
ORF 72	Isonitrile hydratase	684	227	other
ORF 73	hypothetical protein	585	194	other

ORF 74	Putative mycofactocin biosynthesis transcriptional regulator MftR	633	210	regulatory
ORF 75	3-oxoacyl-[acyl-carrier-protein] reductase FabG	798	265	biosynthetic-additional
ORF 76	hypothetical protein	189	62	other
ORF 77	hypothetical protein	630	209	other
ORF 78	Putative ribosomal N-acetyltransferase YdaF	585	194	biosynthetic-additional

Appendix 4: Biosynthetic gene cluster from region 11 of BRB040

Identifier	Product	Length Nucleotides	Aminoacids	Function
ORF 1	putative MFS-type transporter EfpA	1485	494	transport
ORF 2	Validamycin A dioxygenase	1008	335	biosynthetic-additional
ORF 3	Validamycin A dioxygenase	999	332	biosynthetic-additional
ORF 4	Ubiquinone biosynthesis O-methyltransferase, mitochondrial	762	253	biosynthetic-additional
ORF 5	putative protein	729	242	other
ORF 6	hypothetical protein	1050	349	other
ORF 7	3-hydroxybenzoate 6-hydroxylase 1	1356	451	biosynthetic-additional
ORF 8	2-succinylbenzoate--CoA ligase	1590	529	biosynthetic-additional
ORF 9	hypothetical protein	540	179	other
ORF 10	Acetoacetyl CoA synthase NphT7	1062	353	biosynthetic-additional
ORF 11	hypothetical protein	1002	333	other
ORF 12	Arylamine N-acetyltransferase	801	266	other
ORF 13	hypothetical protein	396	131	other
ORF 14	NAD(P)H dehydrogenase (quinone)	669	222	other
ORF 15	Putative glyoxylase CFP32	822	273	other
ORF 16	Phospho-2-dehydro-3-deoxyheptonate aldolase	1209	402	other
ORF 17	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	819	272	biosynthetic-additional
ORF 18	Phenazine biosynthesis protein PhzD2	675	224	biosynthetic-additional
ORF 19	Isochorismate synthase MenF	1917	638	biosynthetic-additional
ORF 20	Acrylyl-CoA reductase AcuI	978	325	biosynthetic-additional
ORF 21	hypothetical protein	1341	446	other
ORF 22	hypothetical protein	960	319	other
ORF 23	hypothetical protein	654	217	other
ORF 24	hypothetical protein	876	291	other
ORF 25	Bifunctional F420 biosynthesis protein FbiB	681	226	biosynthetic-additional
ORF 26	hypothetical protein	552	183	regulatory
ORF 27	hypothetical protein	1914	637	other
ORF 28	3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase	1842	613	biosynthetic-additional
ORF 29	hypothetical protein	1218	405	other
ORF 30	hypothetical protein	477	158	biosynthetic-additional

ORF 31	Actinorhodin polyketide putative beta-ketoacyl synthase 1	1272	423	biosynthetic
ORF 32	Actinorhodin polyketide putative beta-ketoacyl synthase 2	1275	424	biosynthetic
ORF 33	hypothetical protein	270	89	biosynthetic-additional
ORF 34	Putative polyketide cyclase	480	159	biosynthetic-additional
ORF 35	Tetracenomycin F2 cyclase	336	111	biosynthetic-additional
ORF 36	hypothetical protein	900	299	other
ORF 37	F420H(2)-dependent reductase	447	148	other
ORF 38	hypothetical protein	747	248	other
ORF 39	hypothetical protein	597	198	other
ORF 40	Non-homologous end joining protein Ku	1083	360	other
ORF 41	Multifunctional non-homologous end joining protein LigD	849	282	other
ORF 42	hypothetical protein	804	267	other
ORF 43	hypothetical protein	816	271	other
ORF 44	Sensor protein CseC	1245	414	regulatory
ORF 45	hypothetical protein	459	152	other
ORF 46	putative FtsW-like protein	1374	457	other
ORF 47	Penicillin-binding protein A	1458	485	other
ORF 48	hypothetical protein	510	169	other
ORF 49	hypothetical protein	939	312	biosynthetic-additional
ORF 50	HTH-type transcriptional regulator BetI	654	217	regulatory
ORF 51	hypothetical protein	288	95	other
ORF 52	hypothetical protein	1158	385	other
ORF 53	Styrene monooxygenase StyA	1263	420	other
ORF 54	hypothetical protein	642	213	other
ORF 55	hypothetical protein	453	150	other
ORF 56	hypothetical protein	504	167	other
ORF 57	Adaptive-response sensory-kinase SasA	3123	1040	other
ORF 58	hypothetical protein	792	263	other
ORF 59	hypothetical protein	495	164	regulatory
ORF 60	Lysozyme M1	828	275	other
ORF 61	Lon protease	2424	807	other
ORF 62	Polyamine aminopropyltransferase	699	232	other
ORF 63	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	735	244	regulatory
ORF 64	Adaptive-response sensory-kinase SasA	1125	374	regulatory

Appendix 5: Biosynthetic gene cluster from region 14 of BRB040

Identifier	Product	Length		Function
		Nucleotides	Aminoacids	
ORF 1	2-(acetamidomethylene)succinate hydrolase	909	302	biosynthetic-additional
ORF 2	hypothetical protein	1650	549	other
ORF 3	Linearmycin resistance ATP-binding protein LnrL	777	258	transport
ORF 4	hypothetical protein	654	217	other
ORF 5	hypothetical protein	1386	461	other
ORF 6	hypothetical protein	1677	558	other
ORF 7	hypothetical protein	1140	379	other
ORF 8	Mitomycin biosynthesis 6-O-methyltransferase	1083	360	biosynthetic-additional
ORF 9	Ion-translocating oxidoreductase complex subunit B	330	109	other
ORF 10	hypothetical protein	1974	657	other
ORF 11	Acylamidase	1389	462	biosynthetic-additional
ORF 12	Gamma-glutamylanilide synthase	1521	506	other
ORF 13	Adenylosuccinate lyase	1287	428	biosynthetic-additional
ORF 14	hypothetical protein	405	134	other
ORF 15	Transcriptional regulatory protein WalR	789	262	regulatory
ORF 16	Antitumor antibiotic C-1027 apoprotein	417	138	biosynthetic
ORF 17	hypothetical protein	633	210	other
ORF 18	HTH-type transcriptional activator RhaR	471	156	regulatory
ORF 19	hypothetical protein	423	140	other
ORF 20	HTH-type transcriptional activator RhaR	825	274	regulatory
ORF 21	Transcriptional regulator NovG	1188	395	other
ORF 22	hypothetical protein	840	279	other
ORF 23	Bacitracin transport ATP-binding protein BcrA	927	308	transport
ORF 24	Putative glutamine amidotransferase	738	245	other
ORF 25	hypothetical protein	516	171	other
ORF 26	dTDP-4-dehydro-6-deoxy-alpha-D-glucopyranose 2,3-dehydratase	1374	457	other
ORF 27	dTDP-glucose 4,6-dehydratase	981	326	biosynthetic-additional
ORF 28	Polyketide putative beta-ketoacyl synthase 1	348	115	biosynthetic
ORF 29	Tripeptidyl aminopeptidase	1548	515	biosynthetic-additional
ORF 30	10-deoxymethynolide desosaminyltransferase	1317	438	biosynthetic-additional
ORF 31	hypothetical protein	90	29	other
ORF 32	Hercynine oxygenase	1410	469	other

ORF 33	hypothetical protein	780	259	biosynthetic-additional
ORF 34	Hercynine oxygenase	1395	464	other
ORF 35	Enterobactin exporter EntS	1296	431	transport
ORF 36	hypothetical protein	483	160	other