



**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
FACULDADE DE BIOTECNOLOGIA
CURSO DE BACHARELADO EM BIOTECNOLOGIA**

DAVI JOSUÉ MARCON

**ANÁLISE GENÔMICA DE BACTÉRIAS ISOLADAS DO PARQUE
ESTADUAL UTINGA - PARÁ**

**Belém
2022**



**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
FACULDADE DE BIOTECNOLOGIA
CURSO DE BACHARELADO EM BIOTECNOLOGIA**

DAVI JOSUÉ MARCON

**ANÁLISE GENÔMICA DE BACTÉRIAS ISOLADAS DO PARQUE
ESTADUAL UTINGA - PARÁ**

Trabalho de Conclusão de Curso apresentado
para obtenção do grau de Bacharel em Biotecno-
logia.

Orientador: Prof. Dr. Rafael Azevedo Baraúna

**Belém
2022**

RESUMO

Microrganismos ambientais são fontes interessantes de genes e metabólitos com uso biotecnológico, esses genes são comumente encontrados em clusters de genes biossintéticos (BGC) que são aglomerados gênicos contendo genes que codificam diversas enzimas importantes para o processo de síntese dos metabólitos. Nem sempre BGC's são expressos e permanecem silenciosos em condições laboratoriais, sendo necessário o uso de técnicas de eliciação da expressão ou o estudo genômico para analisar e descrever a presença de vias biossintéticas.

Este trabalho utilizou dois microorganismos da coleção microbiológica do Centro de Genômica e biologia de sistemas, e através de sequenciamento e análise genômica, buscou prever a capacidade desses microorganismo para produzir compostos de interesse biotecnológico.

Como principais resultados obtivemos que a cepa 094 possui a maquinaria capaz de produzir ao menos três compostos com possível atividade antimicrobiana, além de possuir vários genes de resistência associados a produção desses compostos.

Palavras-chave: Bactérias, Potencial biotecnológico, Genômica, Predição computacional

SUMÁRIO

1	INTRODUÇÃO	4
1.1	Contexto	4
1.2	Justificativa	4
2	OBJETIVOS	5
2.1	Objetivo Geral	5
2.2	Objetivos Específicos	5
3	REFERENCIAIS TEÓRICOS	6
3.1	Metabólitos secundários	6
3.2	Resistência a antimicrobianos	6
3.3	Actinomicetos	7
3.3.1	<i>Rhodococcus</i>	7
3.4	<i>Brevibacillus brevis</i>	8
3.5	Estudo MIB's	9
4	METODOLOGIA	11
4.1	Seleção de amostras	11
4.2	Extração de DNA e sequenciamento	11
4.3	Análise Genômica	11
5	RESULTADOS E DISCUSSÃO	13
5.1	Sequenciamento e controle de qualidade das leituras	13
5.2	Montagem das <i>contigs</i>	14
5.3	Predição de espécies e montagem de genoma	16
5.4	Predição de <i>BGCs</i> e resistência a antimicrobianos	16
5.5	Estrutura final dos genomas	19
6	CONCLUSÃO	21
	REFERÊNCIAS	22

1 INTRODUÇÃO

1.1 Contexto

A OMS relata uma crise na saúde pública derivada do crescente número de microorganismos resistentes e representam uma grande ameaça global, que somente pode ser resolvida se todos aqueles envolvidos no controle e na distribuição dos medicamentos antimicrobianos cumprissem com suas responsabilidades e deveres quanto a esse problema (ABADI et al., 2019).

A exploração moderna de genes vias metabólicas silencia em microorganismos através de técnicas do estado da arte em genômica e metabolômica, podem revelar novos tesouros metabólicos para solucionar problemas como a crise dos antibióticos (ZHANG et al., 2021).

Os ambientes amazônicos são um reservatório de biodiversidade muito importantes, a riqueza de espécies e o uso da biotecnologia como ferramenta para a solução de problemas industriais e relacionados a saúde humana e animal, seu patrimônio genético é fonte interessante para o desenvolvimento sustentável baseado no uso de tecnologia de ponta para a formulação de novas tecnologias. micro-organismos do solo amazônico podem ser a fonte de novos fármacos para doenças já conhecidas, a cura para doenças emergentes, biofábricas para novos processos industriais e biorremediadores de impactos ambientais.

1.2 Justificativa

Bactérias ambientais são interessantes alvos para a descoberta de compostos de relevância biotecnológica, especialmente como solução para os crescentes níveis de resistência a antimicrobianos encontrados em microorganismos patogênicos. A caracterização genômica e prospecção de genes de interesse desses microorganismos, especialmente do ambiente amazônico, são passos importantes em busca de genes para produção de compostos de potencial farmacológico e industrial.

2 OBJETIVOS

2.1 Objetivo Geral

Predizer o potencial biotecnológico de bactérias ambientais utilizando ferramentas *in silico*

2.2 Objetivos Específicos

1. Caracterizar os organismos sequenciados utilizando seus genomas
2. Predizer as características metabólicas dos organismos
3. Analisar os microorganismos quanto ao potencial de produção de compostos de interesse biotecnológico

3 REFERENCIAIS TEÓRICOS

3.1 Metabolitos secundários

O metabolismo celular bacteriano é o conjunto de processos bioquímicos anabólicos e catabólicos no qual as células bacterianas produzem novas substâncias a partir de substrato ou outras substâncias, os produtos dessas reações são conhecidos como metabólitos. Podendo ser classificados como primários ou secundários, sendo os primários o conjunto de substâncias essenciais para a sobrevivência do organismo, relacionadas à produção de energia e as funções vitais da célula, já os secundários não estão relacionados à sobrevivência da célula, mas sim sua perpetuação no ambiente utilizando estratégias de resistência a situações adversas (GOKULAN; KHARE; CERNIGLIA, 2014).

A maquinaria responsável pela produção desses compostos, normalmente está relacionada a aglomerados de genes biossintéticos (*Biosynthetic Genes Cluster - BGC*) que são dois ou mais genes agrupados codificam a via biosintética para a produção de um metabólito, sendo capazes de produzir compostos das seguintes classes: alcalóides, carboidratos, esteróides, lipídeos, peptídeos (com ou sem modificações pós-traducionais), policetídeos e terpenóides (MEDEMA et al., 2015).

Esses metabólitos possuem uma diversa gama de funções, seja como metodologia de "guerra química" com outros microorganismos, mediadores de atividade mutualística entre espécies ou simbiose química (O'BRIEN; WRIGHT, 2011). Apesar de não serem considerados essenciais para a vida desses organismos (DEMAIN; SANCHEZ, 2009) são de grande importância para sua dispersão e adaptação em ambientes hostis e escassos de nutrientes.

É importante ressaltar que a produção de metabólitos de ação antimicrobiana, está relacionada com a resistência a antimicrobianos, uma vez que, microorganismos produtores de substâncias antimicrobianas precisam resistir à sua ação de forma a evitar o suicídio causado pelas suas próprias substâncias (CUNDLIFFE; DEMAIN, 2010).

3.2 Resistência a antimicrobianos

Bactérias possuem diversos mecanismos para proteção contra agentes antimicrobianos como: desativação do fármaco, mutação no sítio de ligação do fármaco, expressão de bombas de efluxo e desvios metabólicos. Esses mecanismos, podem estar associados a elementos genéticos móveis permitindo a transferência entre indivíduos da mesma espécie ou não (MADIGAN et al., 2021, p. 150). A resistência a antibióticos é um problema emergente que está associado à mortalidade em patógenos bacterianos e sua solução é complexa e permeia a necessidade de políticas públicas, vigilância e controle do uso de antibióticos, medidas de prevenção e o desenvolvimento de novas opções de tratamento (FRIERI; KUMAR; BOUTIN, 2017).

3.3 Actinomicetos

Actinomicetos são um filo de microorganismos gram-positivos de alto conteúdo guanina e citosina que contém as classes: Acidimicrobiia, Actinobacteria, Coriobacteriia, Nitriliruptoria, Rubrobacteria, e Thermoleophilia (YADAV et al., 2018). Dentre suas principais características podemos ressaltar a presença de micélios e a produção de hifas filamentosas (CHATER, 2016). Sua dispersão ambiental é enorme e já foram isolados de ambientes diversos como: lagos salinos, mar profundo e solo (CLAVO et al., 2021; FELÍCIO et al., 2021; SAPKOTA et al., 2020). Além da simbose com animais, fungos, insetos, líquens e plantas (HEI et al., 2021; MEIJ et al., 2017). A capacidade de se adaptar a diversos ambientes está intimamente relacionada com a capacidade de produzir substâncias bioativas com funções igualmente diversas (BERGEIJK et al., 2020)

Essas bactérias foram uma fonte importante para o desenvolvimento de compostos de funções diversas como: antibactericidas, antifúngicos, antihelmínticos, antitumorais, anticancerígenos, anti-inflamatórios, antivirais, imunossupressores, inseticidas e herbicidas (DEMAIN; SANCHEZ, 2009; JOSE; MAHARSHI; JHA, 2021). 64% dos antibióticos derivados de produtos naturais foram obtidos a partir de actinomicetos filamentosos, especialmente durante a era de ouro dos antibióticos (1940-1960) sendo 20 utilizados clinicamente (HUTCHINGS; TRUMAN; WILKINSON, 2019). Segundo Genilloud (2017), continuam sendo uma fonte relevante para o isolamento e caracterização de compostos de interesse biotecnológicos, e com o emprego de metodologias modernas de investigação podem continuar a fornecer substâncias relevantes para mercado.

3.3.1 *Rhodococcus*

O gênero *Rhodococcus* contém actinomicetos de diversidade genômica e fisiológica, contendo alguns membros patogênicos para humanos, animais e plantas. Sua importância biotecnológica é encontrada principalmente por conter algumas cepas com capacidade de degradar compostos orgânicos. Devido seu grande tamanho genômico (8.5-10Mb) esses microorganismos possuem grande liberdade para modificar seu genoma com recombinações, translocações e inserções, contendo diversas vias catabólicas, e mantendo múltiplas funções metabólicas (CAPPELLETTI et al., 2019).

Rhodococcus são os microorganismos mais adequados para o desenvolvimento de tecnologias de remediação de ambientes por serem capazes de degradar poluentes persistentes e por terem sido isolados de ambientes contaminados com hidrocarbonetos (inclusive em forma gasosa) (KUYUKINA; IVSHINA, 2019). Sua resistência a intempéries como frio, calor, acidez, salinidade, pode ser explorada para o desenvolvimento de biorremediadores de derramamento de derivados de petróleo.

Além da capacidade de remediação biológica, podemos ressaltar o potencial de produção de diversas moléculas como: biosurfactantes, biofloclulantes, carotenoides, ácidos graxos poli-

insaturados, poli-hidroxi-alcalóides e triacil-glicerois (CAPPELLETTI et al., 2020). Essas estruturas especialmente as mais complexas como os Carotenoides e os Ácidos graxos Poli-Insaturados são de grande interesse industrial pois sua síntese é complexa e custosa, o uso de microorganismos pode facilitar e reduzir os custos nesses processos.

Dentre as possibilidades para o uso biotecnológico de *Rhodococcus* temos o uso como biofábricas para óleo, biocatálise em processos industriais e valorização de rejeitos (ALVA-REZ et al., 2021; KRIVORUCHKO; KUYUKINA; IVSHINA, 2019; ANTHONY et al., 2019; CHATTERJEE et al., 2020).

3.4 *Brevibacillus brevis*

Brevibacillus (anteriormente *Bacillus brevis*) é um gênero com grande potencial para uso organismo para uso como organismo de expressão heteróloga por ter crescimento rápido, baixa produção de proteases extracelulares e boa eficiência de transformação por eletroporação, além disso diversos membros do gênero produzem substâncias com atividades larvicidas e antimicrobianas e tem grande importância agroecológica por sua relação mutualística com plantas promovendo seu crescimento, protegendo de doenças e removendo metais pesados do solo (PANDA et al., 2014; RAY; PATEL; AMIN, 2020).

Yao, Zhang e Wu (2020) ressaltam capacidade prolífica de *Brevibacillus* para expressão heteróloga especialmente sua capacidade de produzir moléculas com eficiência ao ser mediada por promotores endógenos com repetição em tandem e peptídeos sinal, sugerindo a importância do uso de estratégias eficazes de otimização do hospedeiro, do vetor, do processo fermentativo e o estudo detalhado dos promotores do gênero para melhoria desse modelo.

Exemplos importantes de metabólitos obtidos de *Brevibacillus* temos os peptídeos anti-microbianos (AMP), sendo esses classificados pela sua síntese ribossomal ou não, tendo diversos usos como o biocontrole em plantas, preservantes para alimentos em prateleiras (YANG; YOUSEF, 2018). Além dos AMPS podemos citar a probdigiosina com atividade algicida e compostos ainda não elucidados com grande atividade antiproliferativa (ZHANG et al., 2022; ARUMUGAM et al., 2018). Além dos metabólitos, algumas vias bioquímicas dos *Brevibacillus* são interessantes pela capacidade de degradar Ácido Polilático (plástico biodegradável), a síntese de exopolissacarídeos e a biodegradação de polietileno (YU et al., 2022; YILDIZ et al., 2015; HADAD; GERESH; SIVAN, 2005; ALI; ZAKARYA; KHALED, 2022).

A espécie *Brevibacillus Brevis* contém espécies majoritariamente mesofílicas, e sua distinção é baseada em similaridade genômica, sondagem molecular e análises quimiotaxonômicas (RAY; PATEL; AMIN, 2020).

3.5 Estudo MIB's

Em 2012 Bérdy (2012) comentou a respeito do declínio no desenvolvimento de novos fármacos, sendo eles resultantes de falhas humanas devido ao uso irresponsável de medicamentos, falhas científicas devido a limitações técnicas e ambientes econômicos de regulação custosos e estritos que limitam o desenvolvimento de novos medicamentos. Após o fim da era de ouro dos antibióticos, um decréscimo dramático foi observado nos níveis de descoberta de novos fármacos, apesar disso, o desenvolvimento nas áreas de espectrometria de massas, metabolômica, genômica e transcriptômica além do baixo custo para o sequenciamento de um genoma são ferramentas importantes para o direcionamento no desenvolvimento de novos compostos derivados de produtos naturais (KATZ; BALTZ, 2016).

O uso de ferramentas computacionais para a mineração de dados e predição de informações a partir de genomas é uma estratégia promissora por ser eficiente econômica e laborosamente, e podem servir de estratégias guiadoras para o uso de outras ferramentas (ADAMEK et al., 2017). Trivella e Felício (2018) propõe o uso de um tripé para a identificação de novos produtos naturais derivados de bactérias, sendo estes: o uso de mineração genômica, a manipulação de condições de cultivo para eliciação da expressão de genes e a metabolômica baseada em espectrometria de massa. Com essas ferramentas, os autores acreditam que a integração da genômica com técnicas de obtenção e purificação de metabólitos serão as bases para o desenvolvimento de novos produtos farmacológicos pelos próximos anos.

Ramírez-Rendon et al. (2022) ressalta a relevância de bactérias para a descoberta de importantes fármacos e propõe que organismos de fontes não convencionais como cavernas, fontes termais, áreas de alta salinidade, solos áridos, oceanos e mares continuem sendo estudados especialmente com tecnologias como metagenômica e mineração genômica pois podem ter um papel importante no combate de possíveis surtos de doenças como a SARS-COV2 e epidemias causadas por bactérias resistentes.

Em condições laboratoriais, muitos genes relacionados a síntese de compostos bioativos são silenciados, limitando a produção desses produtos, propondo que o uso de eliciadores é necessário para expressão dos genes relacionados a produção desses compostos (RUTLEDGE; CHALLIS, 2015). Felício et al. (2021) propõe uma metodologia de eliciação para expressão, purificação e caracterização desses compostos além de ressaltar que até 45% dos compostos produzidos por microorganismos são metabólitos secundários eliciados.

Através de Tecnologias modernas como a ferramenta ANTI-SMASH (MEDEMA et al., 2011) é possível prever genes putativos e *clusters* gênicos relacionados a produção de metabólitos secundários e de síntese ribossomal. Essa tecnologia de mineração *in silico* permite prever redes metabólicas e possíveis promotores da expressão desses compostos, principalmente por utilizar bancos de dados produzidos a partir de outras ferramentas como BAGEL, NORINE e CLUSEAN (JONG et al., 2010; HEEL et al., 2013; CABOCHE et al., 2007; WEBER et al.,

2009). A incorporação de diversas ferramentas e banco de dados permite uma análise robusta e completa utilizando tecnologias do estado da arte da biologia computacional.

4 METODOLOGIA

4.1 Seleção de amostras

Foram selecionados 2 microorganismos de espécies diferentes do banco de amostras ambientais provenientes do parque estadual Utinga - Belém, PA gentilmente disponibilizadas pelo Centro de Genômica e Biologia de Sistemas. Incluindo uma actinobactéria do gênero *Rhodococcus* (ACT016) e e uma bactéria do filo *Firmicutes*: *Brevibacillus brevis*(ACT094). Essa amostras foram previamente identificadas utilizando sequenciamento do gene de RNA ribossomal 16s utilizando os primers universais 8F: 5'-AGAGTTTGATCATGGCTCAG-3' e 1492R: 5'-CGGTTACCTTGTTACGACTT-3' com o sequenciador ABI Prism 3500 Genetic Analyzer (Applied BioSystems). Posteriormente as espécies foram preditas utilizando homologia baseada no alinhamento contra o banco de dados de RNA ribossomal do NCBI utilizando a ferramenta blast.

4.2 Extração de DNA e sequenciamento

As amostras foram cultivadas em meio Tryptone Soy Broth (TSB) por 48 horas á 28 graus, e seu DNA foi extraído utilizando o kit HiPureA Multi-sample DNA Purification Kit(HI-MEDIA) seguindo as orientações do fabricante. O DNA foi quantificado usando quantificador Qubit(Thermo Fisher) e sua integridade foi avaliada por eletroforese em gel de agarose 1% complementado com brometo de etídeo 0.5%. As bibliotecas foram preparadas utilizando o protocolo do fabricante e sequenciadas no equipamento Ion GeneStudio S5 Plus (Thermo Fisher)

4.3 Análise Genômica

O pipeline Bactopia, filtrou as leituras, montou e anotou o genoma automaticamente. Paralelamente, as amostras foram filtradas manualmente utilizando a ferramenta Trimmomatic utilizando os seguintes parâmetros: "LEADING:3 TRAILING:30 SLIDINGWINDOW:4:15 MIN-LEN:36"e foram realizadas montagens manuais utilizando o software Shovill com os montadores SKESA e SPADES utilizando parâmetros automáticos e kmers específicos (21,33,55,77,99,127). As melhores montagens foram selecionadas após visualização da qualidade no software QUAST. Após isso as melhores montagens foram submetidas ao programa KRAKEN para determinar a pureza das montagens e predição de espécies. Posteriormente os genomas foram montados em um único cromossomo utilizando o software RAGOUT utilizando genomas de referência proximos a espécie predita pelo KRAKEN, a pureza das montagens foi verificada utilizando o software BUSCO.

Finalmente os genomas foram anotados utilizando o software PROKKA, seus genes

de resistência a antibióticos foram preditos utilizando o software armfinder e seus clusters de metabólitos secundários foram preditos utilizando a ferramenta ANTISMASH.

Foram utilizados os softwares fastqc e Artemis para criação de figuras a partir dos dados gerados.

5 RESULTADOS E DISCUSSÃO

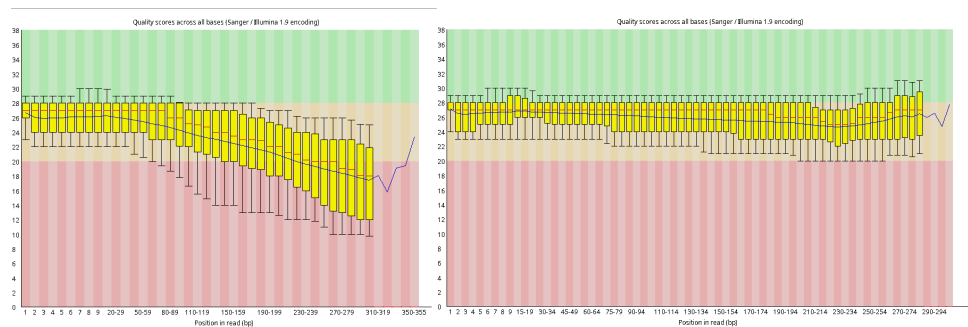
5.1 Sequenciamento e controle de qualidade das leituras

Após o sequenciamento das amostras, foram obtidas 7.8 milhões de leituras de tamanho médio de 223 pares de base para a amostra ACT016 e 7.4 milhões de leituras com tamanho médio de 222 pares de base para a amostra ACT094. Após a filtrar as leituras utilizando a ferramenta Trimmomatic, retivemos 6.2 milhões de leituras com tamanho médio 113 pares de base *perda*21,5% para ACT016 e 6.1 milhões de leituras com tamanho médio 145 pares de base *perda*18,5%.

Baseando-se num tamanho de genoma variável de 3 a 10 milhões de bases para *Rhodococcus*, podemos determinar a cobertura real estimada pela fórmula $C = (L \cdot N)/G$ sendo C a cobertura, L o comprimento médio das reads e G o tamanho do genoma. A partir disso, obtivemos que a cobertura para a amostra ACT016 Após filtrar as leituras está entre 70 e 233,53 vezes. Para a amostra ACT094, consideramos o tamanho do genoma de referência de *Brevibacillus Brevis*(NZ_LR134338) de 6.2 milhões de bases e estimamos a cobertura em aproximadamente 142,66 vezes.

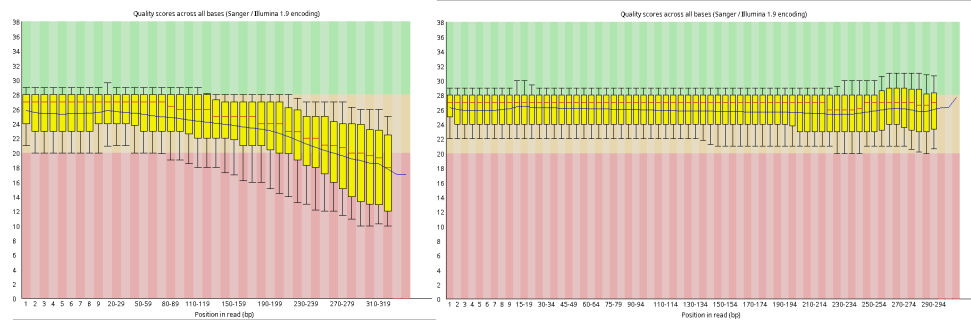
As qualidades médias das sequências pode ser observada a partir dos gráficos a seguir gerados pela ferramenta FASTQC:

Figura 1 – Gráficos representando a qualidade média das leituras da amostra ACT016 na escala PHRED



Fonte: O Autor (2022)

Figura 2 – Gráficos representando a qualidade média das leituras da amostra ACT094 na escala PHRED



Fonte: O Autor (2022)

A partir desses gráficos podemos observar a perda de qualidade no final das leituras, um tipo de limitação técnica comum ao utilizar sequenciadores da plataforma *Illumina*, porém o término em baixa qualidade pode ser removido após a filtração, tendo uma qualidade média ao longo da sequência próximo de PHRED 26 e removendo sequências abaixo de PHRED 20 (que representa probabilidade de erro maior que 1 em 100).

5.2 Montagem das *contigs*

Figura 3 – Report do software QUASt para as montagens da amostra 016

Statistics without reference	016_bactopia_skese_contigs	016_bactopia_spades2_contigs	016_bactopia_spades_contigs	016_manual_more_kmers	016_manual_spades_contigs
# contigs	2893	1073	2859	267	211
# contigs (>= 0 bp)	2893	1073	2859	267	211
# contigs (>= 1000 bp)	2088	976	2069	241	189
# contigs (>= 5000 bp)	154	471	162	185	146
# contigs (>= 10000 bp)	9	195	9	148	119
# contigs (>= 25000 bp)	0	15	0	90	80
# contigs (>= 50000 bp)	0	0	0	31	40
Largest contig	12 514	41 308	12 514	168 716	222 317
Total length	5 931 701	6 487 418	5 931 691	6 431 423	6 438 566
Total length (>= 0 bp)	5 931 701	6 487 418	5 931 691	6 458 455	6 467 051
Total length (>= 1000 bp)	5 340 891	6 418 042	5 350 903	6 413 515	6 424 096
Total length (>= 5000 bp)	1 013 529	5 074 604	1 063 244	6 277 148	6 315 422
Total length (>= 10000 bp)	100 586	3 102 397	100 586	5 997 032	6 110 524
Total length (>= 25000 bp)	0	468 443	0	5 023 485	5 423 825
Total length (>= 50000 bp)	0	0	0	2 903 777	4 003 427
N50	2719	9684	2758	46 660	84 527
N90	1002	2969	1010	14 230	18 884
auN	3216.4	11 552	3258.2	61 550	78 958
L50	702	210	692	38	29
L90	2086	667	2057	130	97
GC (%)	68.06	68.1	68.05	68.09	68.08
Mismatches					
# N's per 100 kbp	0	0	0	0	0
# N's	0	0	0	0	0

Fonte: O Autor (2022)

A melhor montagem para a amostra ACT016 é a montagem *manual_spades* que foi feita utilizando o montador spades junto com a correção do software shovill. Essa montagem foi escolhida por ter o maior valor de L50 (menor quantidade de contigs para atingir 50 % do número de pares de base) e maior *contig* em tamanho absoluto (222 mil pares de base). O conteúdo GC de 60 % dessa montagem está de acordo com o descrito por Yadav et al. (2018) para *Actinomicetos*.

Tabela 1 – Report do software QUASt para as montagens da amostra 016

Estatística	Nome das montagens			
	bactopia_skesa	bactopia_spades	manual_more_kmers	manual_spades
Número de contigs	2893	2859	267	211
Número de contigs (>= 0 bp)	2893	2859	373	314
Número de contigs (>= 1000 bp)	2088	2069	241	189
Número de contigs (>= 5000 bp)	154	162	185	146
Número de contigs (>= 10000 bp)	9	9	148	119
Número de contigs (>= 25000 bp)	0	0	90	80
Número de contigs (>= 50000 bp)	0	0	31	40
Maior contig	12514	12514	168716	222317
Comprimento total	5931701	5931691	6431423	6438566
Comprimento total (>= 0 bp)	5931701	5931691	6458455	6467051
Comprimento total (>= 1000 bp)	5340891	5350903	6413515	6424096
Comprimento total (>= 5000 bp)	1013529	1063244	6277148	6315422
Comprimento total (>= 10000 bp)	100586	100586	5997032	6110524
Comprimento total (>= 25000 bp)	0	0	5023485	5423825
Comprimento total (>= 50000 bp)	0	0	2903777	4003427
N50	2719	2758	46660	84527
N90	1002	1010	14230	18884
auN	3216.4	3258.2	61550	78958
L50	702	692	38	29
L90	2086	2057	130	97
GC (%)	68.06	68.05	68.09	68.08

Fonte: Produzido pelo autor.

Informações obtidas a partir dos resultados do avaliador de montagem quast para a amostra 016

Figura 4 – Report do software QUASt para as montagens da amostra 094

Statistics without reference	094_bactopia_skesa_contigs	094_bactopia_spades2_contigs	094_bactopia_spades_contigs	094_manual_more_kmers	094_manual_spades_contigs
# contigs	1359	90	1356	98	96
# contigs (>= 0 bp)	1359	90	1356	174	145
# contigs (>= 1000 bp)	1184	83	1179	89	86
# contigs (>= 5000 bp)	445	70	442	77	74
# contigs (>= 10000 bp)	134	67	134	72	68
# contigs (>= 25000 bp)	4	54	5	58	54
# contigs (>= 50000 bp)	0	43	0	43	40
Largest contig	47 444	355 470	47 444	426 847	506 270
Total length	6 192 419	6 367 243	6 192 418	6 366 868	6 368 789
Total length (>= 0 bp)	6 192 419	6 367 243	6 192 418	6 385 784	6 381 692
Total length (>= 1000 bp)	6 061 227	6 362 834	6 059 168	6 360 679	6 362 541
Total length (>= 5000 bp)	4 118 533	6 336 758	4 120 357	6 335 651	6 338 367
Total length (>= 10000 bp)	1 956 669	6 317 485	1 976 108	6 300 503	6 297 106
Total length (>= 25000 bp)	132 715	6 107 188	167 349	6 074 220	6 061 343
Total length (>= 50000 bp)	0	5 687 662	0	5 525 199	5 538 539
N50	7064	144 987	7067	140 750	154 727
N90	2188	47 605	2176	41 752	41 705
auN	8683.7	162 005	8819.3	155 188	173 136
L50	272	15	269	16	14
L90	871	44	869	48	45
GC (%)	47.27	47.19	47.27	47.19	47.2
Mismatches					
# N's per 100 kbp	0	0	0	0	0
# N's	0	0	0	0	0

Fonte: O Autor (2022)

De maneira similar, a melhor montagem para a amostra 094 também é a montagem *manual_spades*, possuindo um valor similar a montagem *more_kmers* porém se diferenciando por possuir a maior *contig* com aproximadamente 80 mil pares de base a mais, um fator muito relevante para a posterior montagem do genoma completo. O conteúdo GC também está de acordo com valores comumente encontrados em *Brevibacillus brevis* (NAKAMURA, 1991).

5.3 Predição de espécies e montagem de genoma

Os melhores conjuntos de *contigs* montadas foram submetidas ao software KRAKEN2 para predição de espécie, para a amostra 016 a espécie predita foram *Rhodococcus* com 71,66 % de *contigs* identificadas, mas importunamente 18,45 % das *contigs* não foram identificadas, determinando que a amostra 016 foi predita como *Rhodococcus* não identificado. A partir desse resultado o genoma de referência sugerido para a montagem foi o de código de acesso NZ_CP054690 da cepa *Rhodococcus* sp. W8901.

Para a amostra 094, 85.52 % das *contigs* foram preditas como *Brevibacillus brevis* concordando com os resultados previamente obtidos a partir de sequenciamento de sanger. O genoma NZ_CP030117 da cepa *DZQ7* referência para montagem de genomas dessa espécie.

As montagens utilizando genomas de referência escolhidos, foram avaliadas quanto a presença de genes ortólogos. Na amostra 016 foram encontrados 120 BUSCOs completos e únicos, 2 genes completos e duplicados e 2 genes fragmentados, o valor de genes completos únicos de 98% foi considerada satisfatória para a montagem e de pureza suficiente para a anotação do genoma. Já na amostra 094, 121 BUSCOs completos e únicos foram encontrados, 1 gene completo e duplicado, 1 busco fragmentado e 1 busco faltando, com o percentual de 98,4 % também foi considerada suficiente para prosseguimento da anotação.

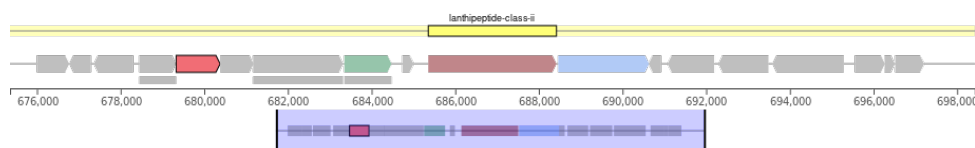
o software PROKKA foi capaz de predizer 5738 CDSs para a amostra 016 e 6082 CDS para a amostra 094, contendo genes de diversas funções celulares.

5.4 Predição de *BGCs* e resistência a antimicrobianos

Para a amostra 016, o software Antismash foi capaz de identificar 16 cluster biosintéticos completos, porém apenas um cluster foi predito com 75% de similaridade para produção de ectoína, um produto biosintético conhecido e produzido por uma ampla gama de bactérias gram positivas e negativas (??). Além disso, através da ferramenta armfinder foi observada a presença de genes codificando uma beta-lactamase de classe A e uma proteína de proteção ribossomal ABC-F. A presença desses genes, pode indicar a resistência desse microorganismo a antibióticos beta-lactâmicos e a macrolídeos.

Já a amostra 094, teve 15 cluster biosintéticos preditos, dentro os quais 4 tiveram sua função identificada por similaridade sendo produtores das seguintes substâncias e identidades percentuais: petrobactina(83%), tyrocidina(81%), gramicidina(91%) e macrobrevina ((100%)).

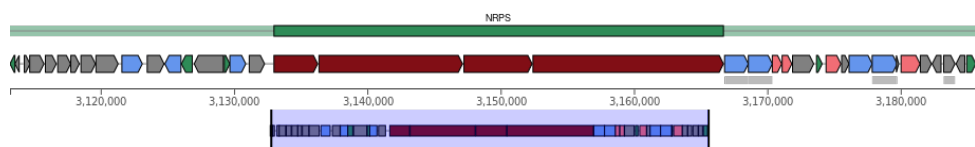
Figura 5 – Cluster biosintético para produção de petrobactina na amostra 094



Fonte: O Autor (2022)

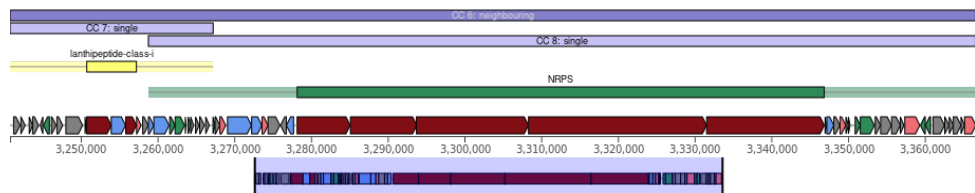
A petrobactina é um sideróforo com protoreatividade capaz de reduzir ferro *III* em ferro *II* (BARBEAU et al., 2002). Interessantemente o cluster presente na amostra 94 capaz de produzir petrobactina, contém uma proteína de resistência tipo A a vancomicina e teicoplanina.

Figura 6 – Cluster biosintético para produção de tirocidina na amostra 094



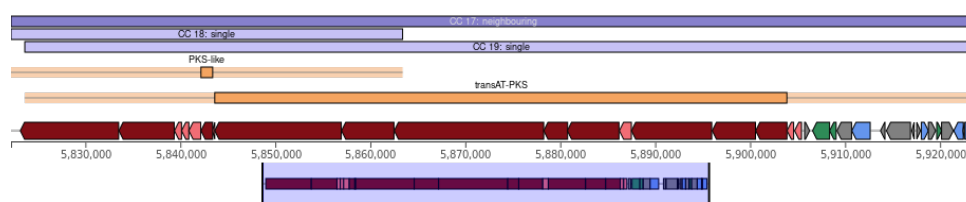
Fonte: O Autor (2022)

Figura 7 – Cluster biosintético para produção de gramicidina na amostra 094



Fonte: O Autor (2022)

A gramicidina e tirocidina são peptídeos lineares de síntese não ribossomal, com mecanismo de ação através do dano na membrana celular de outros organismos, a tirocidina é relatada como um candidato importante a antibiótico de nova geração por não ter induzido resistência nas células afetadas (YANG; YOUSEF, 2018). A presença de clusters capazes de produzir não apenas um mas dois tipos diferentes de peptídeos com atividade antimicrobiana sugere atividade antimicrobiana no isolado 094, o que precisa ser elucidado através de testes *in vitro* em condições que promovam a expressão desses compostos para posterior purificação e descrição de suas estruturas através da técnica de espectrometria de massa.

Figura 8 – Cluster biosintético para produção de macrobrevina na amostra 094

Fonte: O Autor (2022)

O cluster de produção para macrobrevina foi o único encontrado com 100% de similaridade na amostra 094, esse antibiótico com estrutura incomum derivado de uma trans-aciltransferase policetídeo sintase foi inicialmente isolado a partir de *Brevibacillus sp.* associados ao microbioma de plantas (HELFRICH et al., 2018).

A amostra 016 apresenta diversos *cluster* e através de métodos computacionais utilizando bancos de dados extensos não fomos capazes de prever a função desses clusters, o que ressalta a importância de estudos de purificação e descrição das estruturas desses compostos.

Diferentemente, a amostra 094 apresenta 3 clusters identificados como produtores de produtos naturais com atividade antimicrobiana, o que deve ser confirmado através de ensaios elaborados que promovam a expressão desses genes e para sua posterior purificação. Esse microorganismo tem grande potencial para descrição de produtos naturais com atividade antimicrobiana, apesar de serem de classes conhecidas. A elucidação dessas estruturas pode sugerir modificações em antibióticos derivados dessa substância, e também ajudar a complementar informações já existentes a respeito da produção de compostos de interesse biotecnológicos em *Brevibacillus brevis*.

Quanto a presença de genes de resistência a amostra 94 teve a resistência predita para antibióticos como clorofenicol, Kanamicina, Tobramicina, streptomycin, lincossamida, macrolídeos, streptogramina, macrolídeos e beta-lactâmicos.

Tabela 2 – Informações a respeito dos alinhamentos para predição de ARGs

CDS da <i>query</i>	Abrev. do Gene	Cobertura (%)	Identidade (%)	Comprimento
004_01565	catV	100.00	94.52	219
004_04575	ant(4')-Ic	100.00	93.36	256
004_02000	aac(6')-35	100.00	92.74	179
004_02604	ant(6)-Ic	100.00	91.61	286
004_02904	clbC	100.00	90.96	343
004_00361	mphJ	99.68	88.60	307
004_01265	blaBBI	100.00	87.83	304

Fonte: Produzido pelo autor.

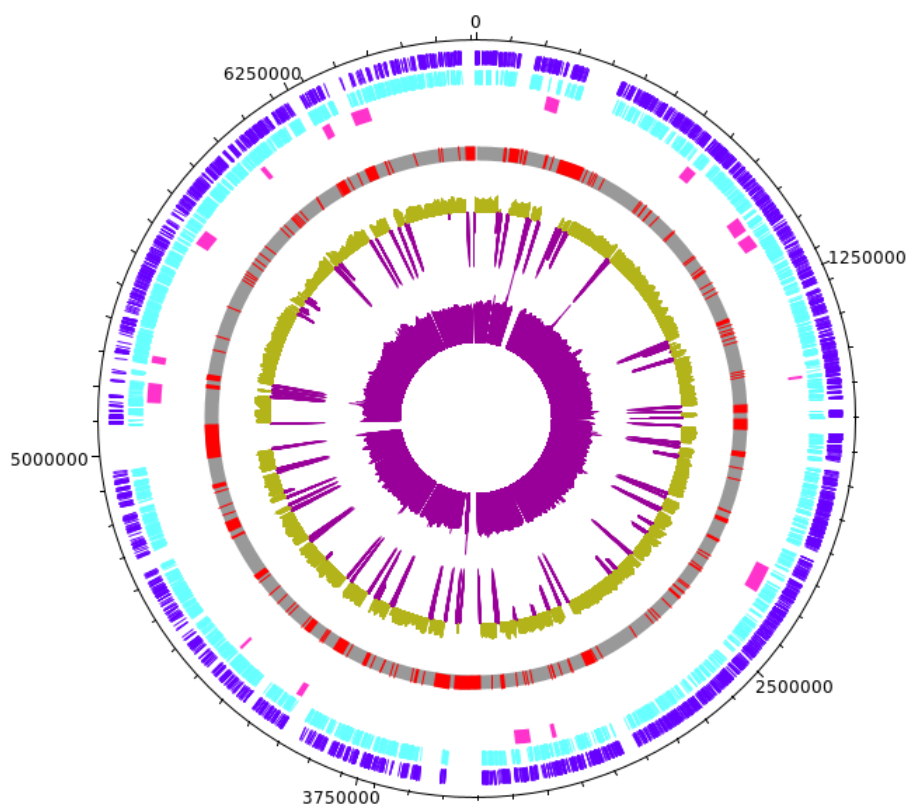
Informações obtidas a partir dos alinhamentos de sequência realizados pela ferramenta armfinder

Os alinhamentos demonstram a presença de vários genes relacionados a resistência,

inclusive a classes inteiras de antibióticos como beta-lactâmicos e macrolídeos, devido ao grande espectro de proteção conferidos por esses genes.

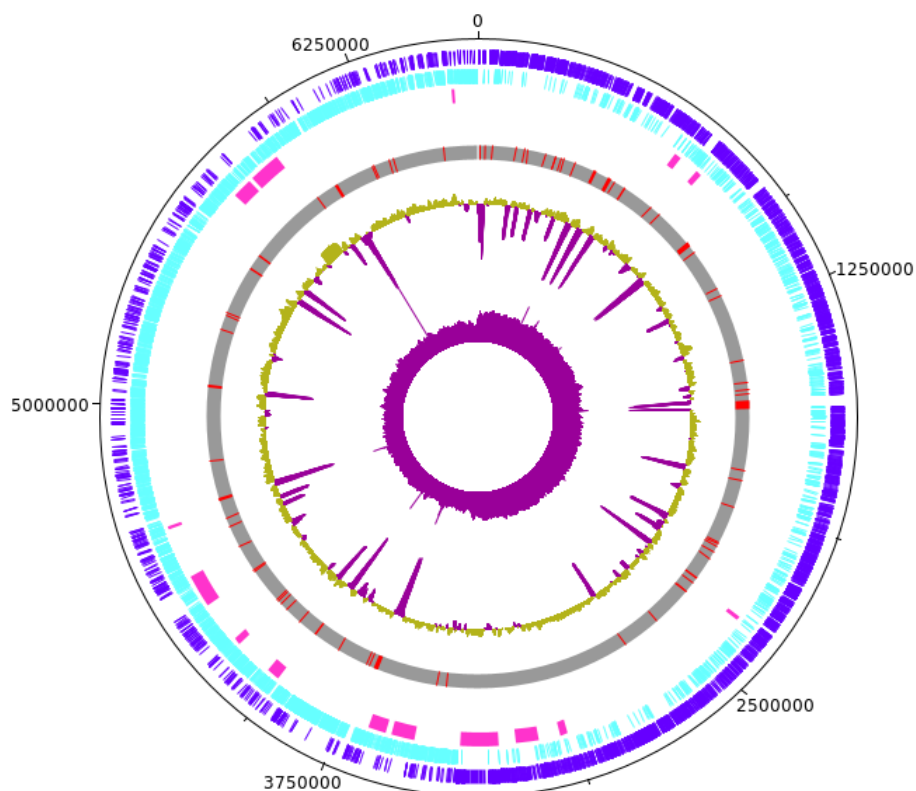
5.5 Estrutura final dos genomas

Figura 9 – Gráficos representando a montagem do cromossomo da amostra 016



Fonte: O Autor (2022)

Figura 10 – Gráficos representando a montagem do cromossomo da amostra 094



Fonte: O Autor (2022)

As figuras ???? representam as montagens completas dos genomas, sendo os círculos do mais exterior para interior: Genes na direção *forward* (roxo), Genes na direção *reverse* (azul claro), *clusters* gênicos (rosa), qualidade da montagem (em cinza a montagem e em vermelho os *gaps* da montagem) e internamente um gráfico representando a variação de conteúdo G+C (amarelo quando maior que 50% e roxo quando menor) e mais internamente um gráfico representando o valor de G+C.

Através dessa visualização, podemos observar a sobreposição de alguns clusters gênicos com áreas de *gaps* na montagem, o que não desqualifica a predição desses clusters mas pode levar a uma baixa identidade ou ausência de componentes nos clusters como genes acessórios importantes para a síntese. O aprimoramento da montagem se faz necessário para o depósito desses genomas em bancos de dados de montagens. Esse pode ter sido um dos motivos para a baixa identificação dos cluster na amostra 016.

6 CONCLUSÃO

Como resultados desse trabalho obtivemos a montagem de 2 novos genomas a fim de depositar no banco de dados de genomas do NCBI. Além de ter descrito o potencial biotecnológico da amostra 094 através de ferramentas *in silico*, abrindo espaço para metodologias confirmatórias para estudo detalhado dos metabólitos desse microorganismo. A cepa da amostra 016 apesar de não tem apresentado clusters identificados ainda deve ser estudada detalhadamente, haja vista que seus metabólitos não puderam ser identificados mesmo ao utilizar os maiores bancos de dados de *clusters* produtores de produtos naturais disponíveis.

O uso das metodologias *in silico* permite a análise detalhada desses organismos, e através desses genomas outras perguntas ainda podem ser respondidas como a presença de ilhas de patogenicidade e genes de virulência além da relação entre esses microorganismos e outros da mesma espécie já descritos anteriormente.

A confirmação da produção de substâncias antimicrobianas será realizada pela equipe do centro de genômica e Biologia de Sistemas para posterior purificação e descrição de sua estrutura.

REFERÊNCIAS

- ABADI, A. T. B. et al. World health organization report: current crisis of antibiotic resistance. **BioNanoScience**, Springer, v. 9, n. 4, p. 778–788, 2019.
- ADAMEK, M. et al. Mining bacterial genomes for secondary metabolite gene clusters. In: **Antibiotics**. [S.l.]: Springer, 2017. p. 23–47.
- ALI, S. A.; ZAKARYA, S.; KHALED, S. Screening and optimisation of the biodegradation potential for low density polyethylene (ldpe) films by fusarium equiseti and brevibacillus parabrevis. **Biosciences Biotechnology Research Asia**, Biosciences Biotechnology Research Asia, v. 19, n. 1, p. 215, 2022.
- ALVAREZ, H. M. et al. Rhodococcus as biofactories for microbial oil production. **Molecules**, MDPI, v. 26, n. 16, p. 4871, 2021.
- ANTHONY, W. E. et al. Development of rhodococcus opacus as a chassis for lignin valorization and bioproduction of high-value compounds. **Biotechnology for biofuels**, Springer, v. 12, n. 1, p. 1–14, 2019.
- ARUMUGAM, T. et al. Isolation, structure elucidation and anticancer activity from brevibacillus brevis egs 9 that combats multi drug resistant actinobacteria. **Microbial pathogenesis**, Elsevier, v. 115, p. 146–153, 2018.
- BARBEAU, K. et al. Petrobactin, a photoreactive siderophore produced by the oil-degrading marine bacterium marinobacter hydrocarbonoclasticus. **Journal of the American Chemical Society**, ACS Publications, v. 124, n. 3, p. 378–379, 2002.
- BÉRDY, J. Thoughts and facts about antibiotics: where we are now and where we are heading. **The Journal of antibiotics**, Nature Publishing Group, v. 65, n. 8, p. 385–395, 2012.
- BERGEIJK, D. A. van et al. Ecology and genomics of actinobacteria: new concepts for natural product discovery. **Nature Reviews Microbiology**, Nature Publishing Group, v. 18, n. 10, p. 546–558, 2020.
- CABOCHE, S. et al. Norine: a database of nonribosomal peptides. **Nucleic acids research**, Oxford University Press, v. 36, n. 1, p. D326–D331, 2007.
- CAPPELLETTI, M. et al. Biotechnology of rhodococcus for the production of valuable compounds. **Applied Microbiology and Biotechnology**, Springer, v. 104, n. 20, p. 8567–8594, 2020.
- CAPPELLETTI, M. et al. Genomics of rhodococcus. In: **Biology of Rhodococcus**. [S.l.]: Springer, 2019. p. 23–60.
- CHATER, K. F. Recent advances in understanding streptomyces. **F1000Research**, Faculty of 1000 Ltd, v. 5, 2016.
- CHATTERJEE, A. et al. Bioconversion of renewable feedstocks by rhodococcus opacus. **Current opinion in biotechnology**, Elsevier, v. 64, p. 10–16, 2020.
- CLAVO, R. F. et al. Evaluation of antimicrobial and antiproliferative activities of actinobacteria isolated from the saline lagoons of northwestern peru. **PloS one**, Public Library of Science San Francisco, CA USA, v. 16, n. 9, p. e0240946, 2021.

CUNDLIFFE, E.; DEMAINE, A. L. Avoidance of suicide in antibiotic-producing microbes. **Journal of Industrial Microbiology and Biotechnology**, Oxford University Press, v. 37, n. 7, p. 643–672, 2010.

DEMAINE, A. L.; SANCHEZ, S. Microbial drug discovery: 80 years of progress. **The Journal of antibiotics**, Nature Publishing Group, v. 62, n. 1, p. 5–16, 2009.

FELÍCIO, R. de et al. Chemical elicitors induce rare bioactive secondary metabolites in deep-sea bacteria under laboratory conditions. **Metabolites**, MDPI, v. 11, n. 2, p. 107, 2021.

FRIERI, M.; KUMAR, K.; BOUTIN, A. Antibiotic resistance. **Journal of infection and public health**, Elsevier, v. 10, n. 4, p. 369–378, 2017.

GENILLOU, O. Actinomycetes: still a source of novel antibiotics. **Natural product reports**, Royal Society of Chemistry, v. 34, n. 10, p. 1203–1232, 2017.

GOKULAN, K.; KHARE, S.; CERNIGLIA, C. Production of secondary metabolites of bacteria. In: BATT, C. A.; TORTORELLO, M. L. (Ed.). **Encyclopedia of Food Microbiology (Second Edition)**. Second edition. Oxford: Academic Press, 2014. p. 561–569. ISBN 978-0-12-384733-1. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780123847300002032>>.

HADAD, D.; GERESH, S.; SIVAN, A. Biodegradation of polyethylene by the thermophilic bacterium *brevibacillus borstelensis*. **Journal of applied microbiology**, Wiley Online Library, v. 98, n. 5, p. 1093–1100, 2005.

HEEL, A. J. V. et al. Bagel3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. **Nucleic acids research**, Oxford University Press, v. 41, n. W1, p. W448–W453, 2013.

HEI, Y. et al. Antimicrobial activity and biosynthetic potential of cultivable actinomycetes associated with lichen symbiosis from qinghai-tibet plateau. **Microbiological Research**, Elsevier, v. 244, p. 126652, 2021.

HELFRICH, E. J. et al. Bipartite interactions, antibiotic production and biosynthetic potential of the arabidopsis leaf microbiome. **Nature Microbiology**, Nature Publishing Group, v. 3, n. 8, p. 909–919, 2018.

HUTCHINGS, M. I.; TRUMAN, A. W.; WILKINSON, B. Antibiotics: past, present and future. **Current Opinion in Microbiology**, v. 51, p. 72–80, 2019. ISSN 1369-5274. Antimicrobials. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1369527419300190>>.

JONG, A. D. et al. Bagel2: mining for bacteriocins in genomic data. **Nucleic Acids Research**, Oxford University Press, v. 38, n. suppl_2, p. W647–W651, 2010.

JOSE, P. A.; MAHARSHI, A.; JHA, B. Actinobacteria in natural products research: Progress and prospects. **Microbiological Research**, Elsevier, v. 246, p. 126708, 2021.

KATZ, L.; BALTZ, R. H. Natural product discovery: past, present, and future. **Journal of Industrial Microbiology and Biotechnology**, Oxford University Press, v. 43, n. 2-3, p. 155–176, 2016.

KRIVORUCHKO, A.; KUYUKINA, M.; IVSHINA, I. Advanced rhodococcus biocatalysts for environmental biotechnologies. **Catalysts**, MDPI, v. 9, n. 3, p. 236, 2019.

- KUYUKINA, M. S.; IVSHINA, I. B. Bioremediation of contaminated environments using rhodococcus. In: **Biology of Rhodococcus**. [S.l.]: Springer, 2019. p. 231–270.
- MADIGAN, M. et al. **Brock biology of microorganisms**. 16. ed. London, England: Pearson Education, 2021.
- MEDEMA, M. H. et al. antismash: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic acids research**, Oxford University Press, v. 39, n. suppl_2, p. W339–W346, 2011.
- MEDEMA, M. H. et al. Minimum information about a biosynthetic gene cluster. **Nature chemical biology**, Nature Publishing Group, v. 11, n. 9, p. 625–631, 2015.
- MEIJ, A. Van der et al. Chemical ecology of antibiotic production by actinomycetes. **FEMS microbiology reviews**, Oxford University Press, v. 41, n. 3, p. 392–416, 2017.
- NAKAMURA, L. *Bacillus brevis* migula 1900 taxonomy: reassociation and base composition of dna. **International Journal of Systematic and Evolutionary Microbiology**, Microbiology Society, v. 41, n. 4, p. 510–515, 1991.
- O'BRIEN, J.; WRIGHT, G. D. An ecological perspective of microbial secondary metabolism. **Current Opinion in Biotechnology**, Elsevier, v. 22, n. 4, p. 552–558, 2011.
- PANDA, A. K. et al. *Brevibacillus* as a biological tool: a short review. **Antonie Van Leeuwenhoek**, Springer, v. 105, n. 4, p. 623–639, 2014.
- RAMÍREZ-RENDON, D. et al. Impact of novel microbial secondary metabolites on the pharma industry. **Applied Microbiology and Biotechnology**, Springer, p. 1–24, 2022.
- RAY, S.; PATEL, N.; AMIN, D. *Brevibacillus*. In: **Beneficial Microbes in Agro-Ecology**. [S.l.]: Elsevier, 2020. p. 149–167.
- RUTLEDGE, P. J.; CHALLIS, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. **Nature reviews microbiology**, Nature Publishing Group, v. 13, n. 8, p. 509–523, 2015.
- SAPKOTA, A. et al. Isolation, characterization, and screening of antimicrobial-producing actinomycetes from soil samples. **International journal of microbiology**, Hindawi, v. 2020, 2020.
- TRIVELLA, D. B.; FELICIO, R. de. The tripod for bacterial natural product discovery: genome mining, silent pathway induction, and mass spectrometry-based molecular networking. **MSystems**, Am Soc Microbiol, v. 3, n. 2, p. e00160–17, 2018.
- WEBER, T. et al. Clusean: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. **Journal of biotechnology**, Elsevier, v. 140, n. 1-2, p. 13–17, 2009.
- YADAV, A. N. et al. Actinobacteria from rhizosphere: molecular diversity, distributions, and potential biotechnological applications. In: **New and future developments in microbial biotechnology and bioengineering**. [S.l.]: Elsevier, 2018. p. 13–41.
- YANG, X.; YOUSEF, A. E. Antimicrobial peptides produced by *brevibacillus* spp.: structure, classification and bioactivity: a mini review. **World journal of microbiology and biotechnology**, Springer, v. 34, n. 4, p. 1–10, 2018.

YAO, D.; ZHANG, K.; WU, J. Available strategies for improved expression of recombinant proteins in *brevibacillus* expression system: a review. **Critical reviews in biotechnology**, Taylor & Francis, v. 40, n. 7, p. 1044–1058, 2020.

YILDIZ, S. Y. et al. Genomic analysis of *brevibacillus thermoruber* 423 reveals its biotechnological and industrial potential. **Applied microbiology and biotechnology**, Springer, v. 99, n. 5, p. 2277–2289, 2015.

YU, J. et al. Comparison of polylactic acid biodegradation ability of *brevibacillus brevis* and *bacillus amyloliquefaciens* and promotion of pla biodegradation by soytone. **Biodegradation**, Springer, p. 1–11, 2022.

ZHANG, J. et al. A glossary for chemical approaches towards unlocking the trove of metabolic treasures in actinomycetes. **Molecules**, MDPI, v. 27, n. 1, p. 142, 2021.

ZHANG, Y. et al. Transcriptome analysis reveals the algicidal mechanism of *brevibacillus laterosporus* against *microcystis aeruginosa* through multiple metabolic pathways. **Toxins**, Multidisciplinary Digital Publishing Institute, v. 14, n. 7, p. 492, 2022.