



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

**DAVI JOSUÉ MARCON**

**ANÁLISE GENÔMICA DE BACTÉRIAS POTENCIALMENTE PRODUTORAS  
DE ANTIMICROBIANOS ISOLADAS DO PARQUE ESTAUDAL 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 POTENCIALMENTE PRODUTORAS  
DE ANTIMICROBIANOS ISOLADAS DO PARQUE ESTAUDAL UTINGA, PARÁ**

Trabalho de Conclusão de Curso apresentado  
para obtenção do grau de Bacharel em Biotecnologia.

Orientador: Prof. Dr. Rafael Azevedo Baraúna

**Belém  
2022**

*Este trabalho é dedicado a todos aqueles que,  
de alguma forma abdicaram de algo e/ou a si mesmos  
pela Ciência.*

# AGRADECIMENTOS

Pela orientação e angario de financiamento: Rafael Baraúna, Paula Schneider, Arthur Silva.

A equipe do Centro de Genômica e Biologia de Sistemas: Beatriz Lobato, Carolina Miranda, Soraya Andrade, Silvanira Barbosa.

Aos desenvolvedores, usuários e contribuintes dos projetos abn $\text{\TeX}$ 2,  $\text{\LaTeX}$ , nextflow, powershell, prokka, antismash, blast, fastq, multiqc, armfinder, ragout, busco, kraken, artemis e demais projetos associados.

Alexandra Elbakyan e ao(s) fundador(es) da Genesis Library.

Aos meus pais e meus irmãos: Fábio, Silvane, Israel, Giulia e Laura Marcon.

Aos meus amigos e colegas: Cayo Uchôa, Tiago Moura, Gustavo Marques, Joaquim Neto, Thiago Cordeiro, Naiana Ribeiro, Valéria Silva, Giovane Pinheiro, Gabriela Campestrini, Wictoria Dias, Aline Castro, Mayanne Farias, Amanda Oliveira, Letícia Lago, José Lucas, Iago Blanco, Isabel Montoril, Gabrielly Andrade, Mayza Miranda, Beatriz Moura, Beatriz Campos.

A minha namorada: Clara Feitosa.

Colaboradores externos a minha formação: Emilyn Conceição, Abhinav Sharma, Marília Lima, Karla Lima, Alex Souza, Robert Petit III.

Aos professores: Luciana Xavier, Agenor Valadares, Rommel Jucá, Moysés Miranda, Adriana Folador, Bruno Duarte, Ricardo de Deus e Alejandro Prado.

Pela estrutura: Universidade Federal do Pará, Centro de Gênômica e Biologia de Sistemas, ENGBIO.

As fontes de financiamento: CAPES, CNPQ, FAPESPA, UFPA.

Nenhum trabalho é feito sozinho.

Obrigado.

*“A consistência é contrária  
à natureza, contrária à vida.  
As únicas pessoas completamente  
consistentes são os mortos.  
(Aldous Huxley - Do What You Will)*

## RESUMO

Tendo em mente a atual crise no mercado farmacológico devido ao surgimento de microrganismos multiresistentes, faz-se necessária a readequação das metodologias de desenvolvimento de fármacos. A mineração genômica permite prever a capacidade de microrganismos produzirem metabólitos sem a necessidade de testes *in vitro*, encurtando os passos até a descoberta de novos fármacos. A partir do sequenciamento de duas cepas bacterianas (*Rhodococcus sp.* e *Brevibacillus brevis*), foi possível montar seu genoma utilizando a ferramenta SPADES, prever os genes nos genomas utilizando a ferramenta PROKKA e prever a produção de Metabólitos secundários usando o ANTI-SMASH. Como principais resultados obtivemos que a cepa de *Rhodococcus sp.*, observamos a presença de 16 clusters ainda sem a função definida. A amostra *Brevibacillus brevis* apresentou 15 clusters sendo 3 (macrobrevidina, tirocidina e gramicidina) com função predita para atividade antimicrobiana. A técnica de mineração genômica, permitiu prospectar informações a respeito da produção de metabólitos, com isso foi possível avaliar o potencial biotecnológico desses organismos com técnicas independentes de cultivo.

**Palavras-chave:** bactérias, potencial biotecnológico, genômica, predição computacional

# ABSTRACT

This is the english abstract. **Keywords:** bacterias, biotechnological potential, genomic, computational prediction.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Gráficos representando a qualidade média das leituras da amostra ACT016 na escala PHRED . . . . .	21
Figura 2 – Gráficos representando a qualidade média das leituras da amostra FIR_094 na escala PHRED . . . . .	22
Figura 3 – Gráficos representando a montagem do cromossomo da amostra 016 . . . . .	24
Figura 4 – Gráficos representando a montagem do cromossomo da amostra 094 . . . . .	25
Figura 5 – <i>Cluster biosintético</i> para produção de petrobactina na amostra 094 . . . . .	26
Figura 6 – <i>Cluster biosintético</i> para produção de tirocidina na amostra 094 . . . . .	26
Figura 7 – <i>Cluster biosintético</i> para produção de gramicidina na amostra 094 . . . . .	26
Figura 8 – <i>Cluster biosintético</i> para produção de macrobrevina na amostra 094 . . . . .	27



## LISTA DE TABELAS

Tabela 1 – Relatório do software QUAST para as montagens da amostra 016 . . . . .	22
Tabela 2 – Relatório do software QUAST para as montagens da amostra 094 . . . . .	23
Tabela 3 – Informações a respeito dos alinhamentos para predição de <i>ARGs</i> . . . . .	28

## LISTA DE ABREVIATURAS E SIGLAS

BGC	<i>Biosyntetic Genes Cluster</i> - Cluster de Genes Biossintéticos
AMP	<i>Antimicrobial Peptides</i> - Peptídeos Antimicrobianos
MIB	Microrganismos de Interesse Biotecnológico
NCBI	<i>National Center of Biotechnology Information</i>
DNA	<i>Deoxyribonucleic acid</i> - Ácido Desoxirribonucleico
RNA	<i>Ribonucleic acid</i> - Ácido ribonucleico
ARG	<i>Antibiotic Resistance Gene</i> - Gene de resistência a antibiótico

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>11</b>
<b>1.1</b>	<b>Contexto . . . . .</b>	<b>11</b>
<b>1.2</b>	<b>Justificativa . . . . .</b>	<b>11</b>
<b>2</b>	<b>OBJETIVOS . . . . .</b>	<b>12</b>
<b>2.1</b>	<b>Objetivo Geral . . . . .</b>	<b>12</b>
<b>2.2</b>	<b>Objetivos Específicos . . . . .</b>	<b>12</b>
<b>3</b>	<b>REFERENCIAL TEÓRICO . . . . .</b>	<b>13</b>
<b>3.1</b>	<b>Metabólitos secundários . . . . .</b>	<b>13</b>
<b>3.2</b>	<b>Resistência a antimicrobianos . . . . .</b>	<b>13</b>
<b>3.3</b>	<b>Mineração genômica . . . . .</b>	<b>14</b>
<b>3.4</b>	<b>Actinomicetos . . . . .</b>	<b>15</b>
<b>3.4.1</b>	<i>Rhodococcus</i> . . . . .	<b>15</b>
<b>3.5</b>	<i>Brevibacillus brevis</i> . . . . .	<b>16</b>
<b>3.6</b>	<b>Estudo MIB's . . . . .</b>	<b>17</b>
<b>4</b>	<b>METODOLOGIA . . . . .</b>	<b>19</b>
<b>4.1</b>	<b>Seleção de amostras . . . . .</b>	<b>19</b>
<b>4.2</b>	<b>Extração de DNA e sequenciamento . . . . .</b>	<b>19</b>
<b>4.3</b>	<b>Análise Genômica . . . . .</b>	<b>19</b>
<b>5</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>21</b>
<b>5.1</b>	<b>Sequenciamento e controle de qualidade das leituras . . . . .</b>	<b>21</b>
<b>5.2</b>	<b>Predição de espécies e montagem de genoma . . . . .</b>	<b>22</b>
<b>5.3</b>	<b>Predição de <i>BGCs</i> e resistência a antimicrobianos . . . . .</b>	<b>25</b>
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>29</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>31</b>
	<b>ANEXO A – RESULTADOS IMPORTANTES DURANTE A MINHA GRADUAÇÃO . . . . .</b>	<b>36</b>
<b>A.1</b>	<b>Artigos em co-autoria . . . . .</b>	<b>36</b>
<b>A.2</b>	<b>Prêmio conferido pelo Comitê internacional de Microbiologia de alimentos e higiene pela 3ª melhor apresentação no 30º Congresso Brasileiro de Microbiologia . . . . .</b>	<b>37</b>

# 1 INTRODUÇÃO

## 1.1 Contexto

A Organização Mundial de Saúde relata uma crise na saúde pública derivada do crescente número de microrganismos 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 silenciosas em microrganismos 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 microrganismos patogênicos. A caracterização genômica e prospecção de genes de interesse desses microrganismos, 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**

Identificar agrupamentos de genes biossintéticos em bactérias de filos com potencial biotecnológico isoladas do Parque Estadual Utinga.

### **2.2 Objetivos Específicos**

1. Sequenciar o genoma de dois isolados de interesse;
2. Montar, anotar e predizer os agrupamentos de genes biossintéticos;
3. Avaliar o potencial biotecnológico dos isolados.

## 3 REFERENCIAL TEÓRICO

### 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 microrganismos, 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, microrganismos 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). Laxminarayan et al. (2013) descrevem extensamente a emergência da crise global no uso dos antibióticos, demarcando o problema não apenas como o uso inadequado pelos pacientes mas um grande conjunto que perpassa pelo enfraquecimento de sistemas públicos de saúde, demora em diagnósticos e o uso de antibióticos como promotores de crescimento na indústria agropecuária.

Para enfrentar esses problemas se faz necessária a readequação da metodologia de desenvolvimento de fármacos, integrando o estado da arte da metabolômica, biologia sintética, genômica e química farmacêutica. O uso de metodologias que integrem as informações gênicas (como o resistoma) de bactérias resistentes com a criação de fármacos (BROWN; WRIGHT, 2016).

De acordo com Cook e Wright (2022), ainda há esperança na luta contra a crise dos antibióticos pois o campos de estudo como a mineração genômica e biologia sintética aliadas de ferramentas poderosas como a inteligência artificial combinadas com a reformulação de técnicas clássicas do desenvolvimento de fármacos permitirão o desenvolvimento de tratamentos eficientes e diversos contra as patologias bacterianas.

### **3.3 Mineração genômica**

Medema, Rond e Moore (2021) descrevem a mineração genômica como uma forma de iluminar a química altamente especializada da vida, uma ferramenta poderosa para o estudo do funcionamento de extensas maquinarias genéticas para a produção de compostos com diversas funcionalidades. Esses estudo servem como unificadores da ecologia e fisiologia além de servir como matéria prima para o desenvolvimento de produtos biotecnológicos.

O estado da arte na mineração genômica tem criado bancos de dados gigantescos com material descrevendo diversos tipos de genes relacionados a produção de metabólitos de interesse, não apenas fornecendo informações ecológicas e filogenéticas, mas também oportunizando o desenvolvimento de novas ferramentas capazes de auxiliar na descoberta de novos genes (CHEVRETTE et al., 2021).

A mineração genômica como metodologia integrativa da bioinformática com outros campos do conhecimento como a química e a fisiologia microbiana, permite que microrganismos com características fenotípicas interessantes sejam extensamente estudados em todos os âmbitos da expressão e produção de metabólitos. Inclusive predizendo propriedades desses metabólitos, sem depender de níveis de expressão, tempo de crescimento ou da purificação. (BAUMAN et al., 2021; BALTZ, 2021)

### 3.4 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.

#### 3.4.1 *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).

### 3.5 *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).

### 3.6 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

Diante da importância biotecnologia dos microrganismos citado anteriormente, para a exploração genômica foram selecionadas duas cepas do parque estadual Utinga - Belém, PA gentilmente disponibilizadas pelo Centro de Genômica e Biologia de Sistemas. Incluindo uma actinobactéria do gênero *Rhodococcus* (ACT\_016) e e uma bactéria do filo *Firmicutes*: *Brevibacillus brevis*(FIR\_094). 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 (III; READ, ), filtrou as leituras, montou e anotou o genoma automaticamente. Paralelamente, as amostras foram filtradas manualmente utilizando a ferramenta Trimmomatic (BOLGER; LOHSE; USADEL, 2014) utilizando os seguintes parâmetros: "LEADING:3 TRAILING:30 SLIDINGWINDOW:4:15 MINLEN:36"e foram realizadas montagens manuais utilizando o software Shovill ( não publicado ) com os montadores SKESA(SOUVOROV; AGARWALA; LIPMAN, 2018) e SPADES(BANKEVICH et al., 2012) utilizando parâmetros automáticos e kmers específicos (21,33,55,77,99,127). As melhores montagens foram selecionadas após visualização da qualidade no software QUAST (GUREVICH et al., 2013).

Após isso as melhores montagens foram submetidas ao programa KRAKEN (WOOD; LU; LANGMEAD, 2019) para determinar a pureza das montagens e predição de espécies. Posteriormente os genomas foram montados em um único cromossomo utilizando o software

RAGOUT(KOLMOGOROV et al., 2014) utilizando genomas de referência próximos a espécie predita pelo KRAKEN, a pureza das montagens foi verificada utilizando o software BUSCO(SIMÃO et al., 2015).

Finalmente os genomas foram anotados utilizando o software PROKKA (SEEMANN, 2014) , seus genes de resistência a antibióticos foram preditos utilizando o software armfinder e seus clusters de metabólitos secundários foram preditos utilizando a ferramenta ANTIS-MASH(BLIN et al., 2021).

Foram utilizados os softwares fastqc(ANDREWS, 2019) e Artemis(CARVER et al., 2012) para criação de figuras a partir dos dados gerados.

## 5 RESULTADOS E DISCUSSÃO

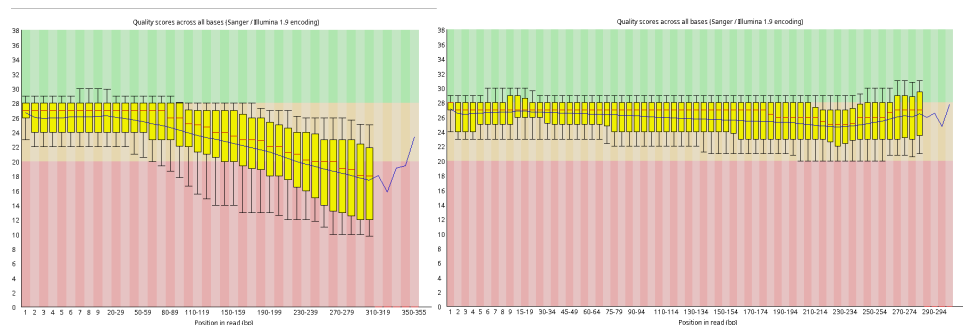
### 5.1 Sequenciamento e controle de qualidade das leituras

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 FIR\_094. 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 FIR\_094, 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.

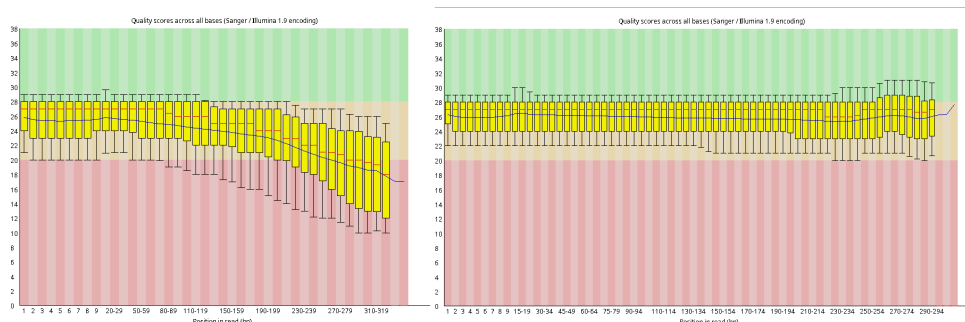
As qualidades médias das sequências pode ser observada a partir dos gráficos a seguir gerados pela ferramenta FASTQC:

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



Box plot representando a distribuição da variação média de qualidade das leituras (eixo Y) em relação a posição da base na leitura (eixo X), a linha azul representa a média. As cores verde, amarelo e vermelho representam respectivamente: qualidade excelente ( $>$ ), boa ( $<27$ ;  $\geq 20$ ) e ruim ( $<20$ ).

**Figura 2 – Gráficos representando a qualidade média das leituras da amostra FIR\_094 na escala PHRED**



Fonte: autoria própria

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 Predição de espécies e montagem de genoma

**Tabela 1 – Relatório do software QUAST para as montagens da amostra 016**

Estatística	Nome das montagens			
	bactopia+skesa	bactopia+spades	manual mais kmers	manual+spadess
Número de contigs	2893	2859	267	211
Número de contigs ( $\geq 0$ bp)	2893	2859	373	314
Número de contigs ( $\geq 1000$ bp)	2088	2069	241	189
Número de contigs ( $\geq 5000$ bp)	154	162	185	146
Número de contigs ( $\geq 10000$ bp)	9	9	148	119
Número de contigs ( $\geq 25000$ bp)	0	0	90	80
Número de contigs ( $\geq 50000$ bp)	0	0	31	40
Maior contig	12514	12514	168716	222317
Comprimento total	5931701	5931691	6431423	6438566
Comprimento total ( $\geq 0$ bp)	5931701	5931691	6458455	6467051
Comprimento total ( $\geq 1000$ bp)	5340891	5350903	6413515	6424096
Comprimento total ( $\geq 5000$ bp)	1013529	1063244	6277148	6315422
Comprimento total ( $\geq 10000$ bp)	100586	100586	5997032	6110524
Comprimento total ( $\geq 25000$ bp)	0	0	5023485	5423825
Comprimento total ( $\geq 50000$ 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: autoria própria

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 2 – Relatório do software QUAST para as montagens da amostra 094**

Estatística	Nome das montagens			
	bactopia+skesa	bactopia+spades	manual mais kmers	manual+spadess
Número de contigs ( $\geq 0$ bp)	1359	1356	174	145
Número de contigs ( $\geq 1000$ bp)	1184	1179	89	86
Número de contigs ( $\geq 5000$ bp)	445	442	77	74
Número de contigs ( $\geq 10000$ bp)	134	134	72	68
Número de contigs ( $\geq 25000$ bp)	4	5	58	54
Número de contigs ( $\geq 50000$ bp)	0	0	43	40
Maior contig	47444	47444	426847	506270
Comprimento total	6192419	6192418	6366868	6368789
Comprimento total ( $\geq 0$ bp)	6192419	6192418	6385784	6381692
Comprimento total ( $\geq 1000$ bp)	6061227	6059168	6360679	6362541
Comprimento total ( $\geq 5000$ bp)	4118533	4120357	6335651	6338367
Comprimento total ( $\geq 10000$ bp)	1956669	1976108	6300503	6297106
Comprimento total ( $\geq 25000$ bp)	132715	167349	6074220	6061343
Comprimento total ( $\geq 50000$ bp)	0	0	5525199	5538539
N50	7064	7067	140750	154727
N90	2188	2176	41752	41705
auN	8683.7	8819.3	155188	173136
L50	272	269	16	14
L90	871	869	48	45
GC (%)	47.27	47.27	47.19	47.2

Fonte: autoria própria

De maneira similar, a melhor montagem para a amostra 094 também é a montagem *manual+spades*, possuindo um valor similar a montagem *manual mais 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).

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

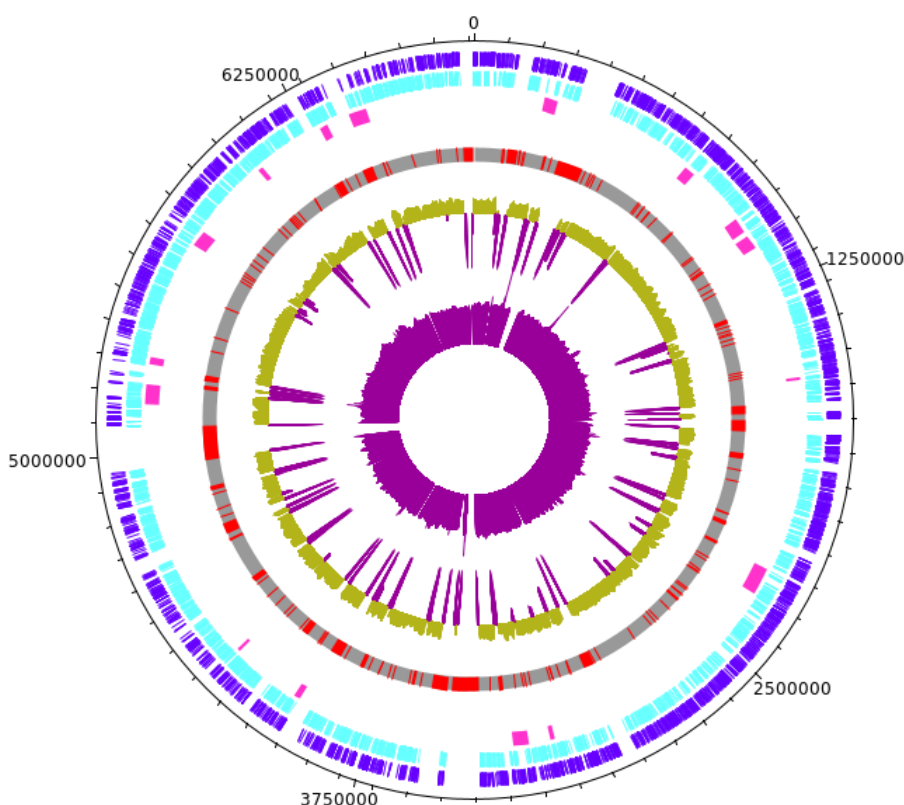


genoma NZ\_CP030117 da cepa *DZQ7* referência para montagem de genomas dessa espécie.

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 prever 5738 CDSs para a amostra 016 e 6082 CDS para a amostra 094, contendo genes de diversas funções celulares.

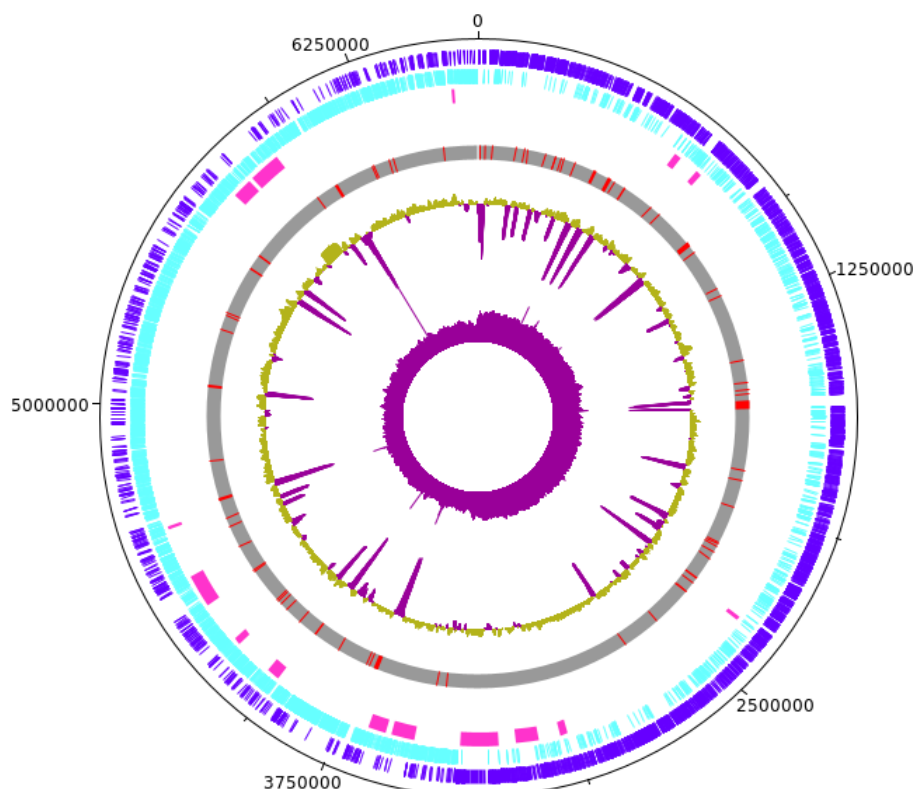
**Figura 3 – Gráficos representando a montagem do cromossomo da amostra 016**



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

Fonte: autoria própria

**Figura 4 – Gráficos representando a montagem do cromossomo da amostra 094**



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

Fonte: autoria própria

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 clusters na amostra 016.

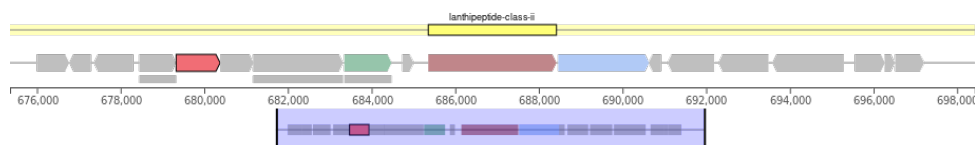
### 5.3 Predição de *BGCs* e resistência a antimicrobianos

Para a amostra 016, o software antismash foi capaz de identificar 16 clusters biossintéticos completos, porém apenas um cluster foi predito com 75% de similaridade para produção de ectoína, um produto biossintético conhecido e produzido por uma ampla gama de bactérias

gram positivas e negativas (STÖVEKEN et al., 2011). Além disso, através da ferramenta arm-finder 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 microrganismo a antibióticos beta-lactâmicos e a macrolídeos.

Já a amostra 094, teve 15 cluster biossintéticos 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%)).

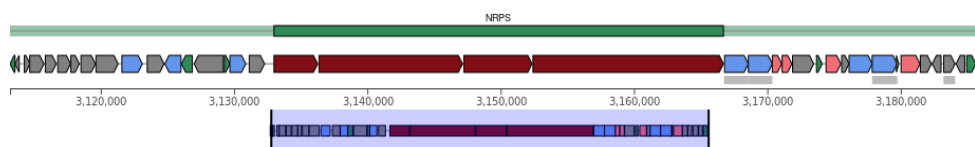
**Figura 5 – Cluster biossintético para produção de petrobactina na amostra 094**



Fonte: autoria própria

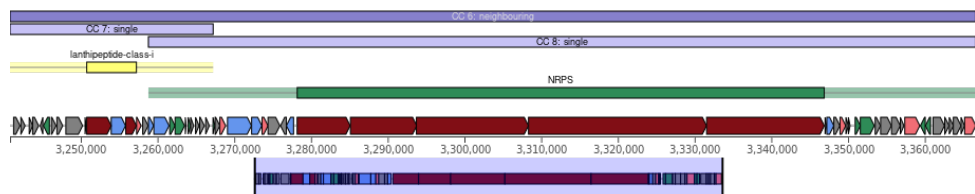
A petrobactina é um sideróforo com protoreatividade capaz de reduzir ferro *III* em ferro *II*(BARBEAU et al., 2002). Interessantemente o cluster presente na amostra 94 capaz de produzir petrobactina, contém uma proteína de resistência tipo A a vancomicina e teicoplanina.

**Figura 6 – Cluster biossintético para produção de tirocidina na amostra 094**



Fonte: autoria própria

**Figura 7 – Cluster biossintético para produção de gramicidina na amostra 094**

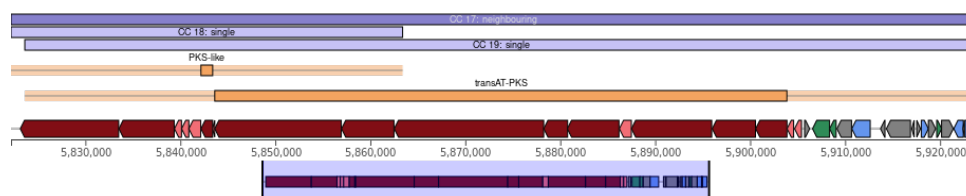


Fonte: autoria própria

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: autoria própria

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 microrganismo 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, streptomicina, lincossamida, macrolídeos, streptogramina, macrolídeos e beta-lactâmicos.

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.

**Tabela 3 – Informações a respeito dos alinhamentos para predição de *ARGs***

<b>CDS da <i>query</i></b>	<b>Abrev. do Gene</b>	<b>Cobertura (%)</b>	<b>Identidade (%)</b>	<b>Comprimento</b>
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:** autoria própria

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

## 6 CONCLUSÃO

A montagem dos genomas foi considerada satisfatória para a sugestão de um *draft*, tendo estatísticas aceitáveis e conteúdo GC próximo ao da literatura para o depósito com número suficiente de CDSs preditas. Porém para a descrição de um genoma completo ainda é necessário o uso de tecnologias de fechamento de *gaps* e aprimoramento das montagens, uma vez que ambas as montagens contêm vários espaçamentos de "NNN" em seus genomas.

Como resultados desse trabalho obtivemos a montagem de 2 novos genomas a fim de depositar no GenBank. 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 microrganismo.

Os *BCGS* encontrados na cepa de *Brevibacillus brevis* (94) correspondem a cluster capazes de produzir macrobrevina, petrobactina, tirocidina e gramicidina. A capacidade de prever esses clusters, permite o uso de eliciadores específicos para esse tipo de cluster conforme a metodologia descrita por Felício et al. (2021).

A técnica de mineração genômica, permitiu prospectar informações a respeito da produção de metabólitos de maneira específica e independente dos padrões de expressão do organismo, o que é de grande interesse uma vez que esse tipo de *cluster* gênico costuma permanecer silenciado em condições de abundância de nutrientes.

Concluimos que a cepa de *Brevibacillus brevis* possui potencial biotecnológico, apesar de se prever como capaz de produzir compostos já conhecidos, seus clusters não são idênticos aos presentes nos bancos de dados e podem conter modificações estruturais interessantes para o desenvolvimento de fármacos baseados nessas modificações, ou o aprimoramento de fármacos já existentes.

A cepa de *Rhodococcus* sp. (16), foi prever por possuir diversos *operons*, mas interessante não foi possível encontrar referências para sua função. Esses dados, apesar de não possuírem impacto biotecnológico direto, são importantes pois através deles, será possível comparar a produção de metabólitos com os genes e através de estudos de expressão heteróloga descrever a função desses clusters e complementar os bancos de dados existentes.

Esse estudo buscou estudar o potencial biotecnológico de dois microrganismos do parque estadual, tendo encontrado resultados importantes em ambas as amostras. Devido a grande quantidade de microrganismos isolados e depositados no banco de bactérias do Centro de Genômica, podemos imaginar a riqueza disponível no genoma desses organismos. O estudo aprofundado do genoma de outros microrganismos pode permitir o estudo de diversos outros *clusters* biossintéticos, e ajudar a compreender a matéria escura gênica (silenciada) desses seres.

O uso das metodologias *in silico* permitiu 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 microrganismos 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.
- ANDREWS, S. **FastQC: A quality control analysis tool for high throughput sequencing data**. [S.l.]: Github, 2019.
- 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.
- BALTZ, R. H. Genome mining for drug discovery: Progress at the front end. **Journal of Industrial Microbiology and Biotechnology**, Oxford University Press, v. 48, n. 9-10, p. kuab044, 2021.
- BANKEVICH, A. et al. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of computational biology**, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 19, n. 5, p. 455–477, 2012.
- 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.
- BAUMAN, K. D. et al. Genome mining methods to discover bioactive natural products. **Natural Product Reports**, Royal Society of Chemistry, v. 38, n. 11, p. 2100–2129, 2021.
- 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.
- BLIN, K. et al. antismash 6.0: improving cluster detection and comparison capabilities. **Nucleic Acids Research**, Oxford University Press, v. 49, n. W1, p. W29–W35, 2021.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for illumina sequence data. **Bioinformatics**, Oxford University Press, v. 30, n. 15, p. 2114–2120, 2014.



BROWN, E. D.; WRIGHT, G. D. Antibacterial drug discovery in the resistance era. **Nature**, Nature Publishing Group, v. 529, n. 7586, p. 336–343, 2016.

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.

CARVER, T. et al. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics**, Oxford University Press, v. 28, n. 4, p. 464–469, 2012.

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.

CHEVRETTE, M. G. et al. The confluence of big data and evolutionary genome mining for the discovery of natural products. **Natural Product Reports**, Royal Society of Chemistry, 2021.

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.

COOK, M. A.; WRIGHT, G. D. The past, present, and future of antibiotics. **Science Translational Medicine**, American Association for the Advancement of Science, v. 14, n. 657, p. eabo7793, 2022.

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>>.

GUREVICH, A. et al. Quast: quality assessment tool for genome assemblies. **Bioinformatics**, Oxford University Press, v. 29, n. 8, p. 1072–1075, 2013.

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>>.

III, R. A. P.; READ, T. D. **Bactopia: a Flexible Pipeline for Complete Analysis of Bacterial Genomes**. **mSystems**. **5** (2020). Disponível em: <<https://github.com/bactopia/bactopia>>.

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.

KOLMOGOROV, M. et al. Ragout — a reference-assisted assembly tool for bacterial genomes. **Bioinformatics**, Oxford University Press, v. 30, n. 12, p. i302–i309, 2014.

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.

LAXMINARAYAN, R. et al. Antibiotic resistance—the need for global solutions. **The Lancet infectious diseases**, Elsevier, v. 13, n. 12, p. 1057–1098, 2013.

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.
- MEDEMA, M. H.; ROND, T. de; MOORE, B. S. Mining genomes to illuminate the specialized chemistry of life. **Nature Reviews Genetics**, Nature Publishing Group, v. 22, n. 9, p. 553–571, 2021.
- 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.
- SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, Oxford University Press, v. 30, n. 14, p. 2068–2069, 2014.
- SIMÃO, F. A. et al. Busco: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, Oxford University Press, v. 31, n. 19, p. 3210–3212, 2015.
- SOUVOROV, A.; AGARWALA, R.; LIPMAN, D. J. Skesa: strategic k-mer extension for scrupulous assemblies. **Genome biology**, Springer, v. 19, n. 1, p. 1–13, 2018.
- STÖVEKEN, N. et al. A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *pseudomonas stutzeri* a1501. **Journal of bacteriology**, Am Soc Microbiol, v. 193, n. 17, p. 4456–4468, 2011.
- 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.

WOOD, D. E.; LU, J.; LANGMEAD, B. Improved metagenomic analysis with kraken 2. **Genome biology**, Springer, v. 20, n. 1, p. 1–13, 2019.

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.

# ANEXO A – RESULTADOS IMPORTANTES DURANTE A MINHA GRADUAÇÃO

## A.1 Artigos em co-autoria

Marine Pollution Bulletin 157 (2020) 111302

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: [www.elsevier.com/locate/marpolbul](http://www.elsevier.com/locate/marpolbul)

Baseline

Occurrence, antibiotic-resistance and virulence of *E. coli* strains isolated from mangrove oysters (*Crassostrea gasar*) farmed in estuaries of Amazonia

Amanda M.S. Oliveira<sup>a,1</sup>, Rafael A. Baraúna<sup>a,b,c,1</sup>, Davi J. Marcon<sup>a</sup>, Letícia A.B. Lago<sup>a</sup>, Artur Silva<sup>a,b</sup>, Joana Lusio<sup>c</sup>, Rafael D.S. Tavares<sup>c,d</sup>, Marta Tacão<sup>c,d</sup>, Isabel Henriques<sup>d,e</sup>, Maria P.C. Schneider<sup>a,b</sup>

<sup>a</sup> Centro de Genética e Biologia de Sistemas, Instituto de Ciências Biológicas, Universidade Federal do Pará, 66075-110 Belém, PA, Brazil  
<sup>b</sup> Laboratório de Engenharia Biológica, Parque de Ciência e Tecnologia Guamã, 66075-750 Belém, PA, Brazil  
<sup>c</sup> Departamento de Biologia, Universidade de Aveiro, 3810-193 Aveiro, Portugal  
<sup>d</sup> CESAM (Centro de Estudos do Ambiente e do Mar), 3810-193 Aveiro, Portugal  
<sup>e</sup> Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade de Coimbra, 3000-456 Coimbra, Portugal

**ARTICLE INFO**

**Keywords:**  
 Oyster  
*Crassostrea gasar*  
*Escherichia coli*  
 Food surveillance  
 Antibiotic resistance  
 Amazonia

**ABSTRACT**

Concentration of bacterial species indicative of fecal contamination in the gut of mangrove oysters (*Crassostrea gasar*) is a major concern for public health and food surveillance. Our work aimed to determine the occurrence, antibiotic-resistance, phylogenetic profile and virulence of *Escherichia coli* strains isolated from *C. gasar* farmed in four estuaries of Amazonia. Santo Antônio de Urindeua was the sampling point with the highest number of *E. coli* cells in oyster samples ( $10^4$  per 100 g of sample). Twenty-four isolates (52.2%) showed resistance to cephalotin and 18 to amoxicillin (39.1%). Eighteen clonal populations were determined by rep-PCR and were mainly affiliated to the pathogenic and commensal phylo-groups B1 and D. The presence of *stx* genes suggests that 10 of these clones belong to the Enterotoxigenic *Escherichia coli* pathotype. Plasmids, mostly of the F incompatibility group, were detected in the majority of the strains. All isolates were susceptible to last-resort antibiotics.

**1. Introduction**

The coastal zone of eastern Amazonia has been used in recent years as an area of intense production of oysters for human consumption. Three oyster species are described in Brazil: *Crassostrea rhizophorae*, *Crassostrea gasar* and *Crassostrea gigas* (Melo et al., 2010). The Pacific oyster *C. gigas* is an exogenous species that was introduced in Brazil and has a significant production in the south of the country. *C. gasar*, the mangrove oyster, grows intensely in the Brazilian north coast and its farming and consumption has increased considerably in recent years (Melo et al., 2010; Varela et al., 2007). *C. gasar* as well as other oyster species are sensitive to environmental variations and have their shell shaped according to the farming site conditions. Due to this constant morphological variation, molecular techniques are commonly employed to determine taxonomic affiliation of these oysters (Boudry et al., 2003; Varela et al., 2007).

Oysters are bivalve mollusks that feed by filtering the water around them and consequently capture microorganisms into their inner tissues. Therefore, the gut microbiome of this animal interacts directly with the microbiome of its aquatic habitat (Pierce et al., 2016). The taxonomic diversity of the microbiome associated to the species *C. virginica*, *C. gigas*, *C. corteziensis*, and *C. sikamea* has been previously described (Fernández et al., 2014; King et al., 2012; Ostrensky et al., 2018; Tróbal et al., 2012), including taxa with potential biotechnological application (Desjac et al., 2014).

Mollusks are commonly used for biomonitoring of environmental pollutants since they are good bioindicators of chemical compounds such as Polycyclic Aromatic Hydrocarbons (PAH's) and Polychlorinated Biphenyl (PCBs) (Aguirre-Rubí et al., 2018), and trace metals such as iron, zinc and copper (Silva et al., 2001). They also serve as bioindicators of fecal contamination, since they act as bioaccumulators of key microorganisms such as *Escherichia coli* (Grevskott et al., 2016; Miotto et al., 2019; Rees et al., 2015). Recently, Grevskott et al. (2016) evaluated the Most Probable Number method specified in the European

\* Corresponding author at: Parque de Ciência e Tecnologia Guamã, Espaço Inovação, Laboratório de Engenharia Biológica, Av. Perimetral, 66075-750 Belém, PA, Brazil.  
 E-mail address: [rbarauna@ufpa.br](mailto:rbarauna@ufpa.br) (R.A. Baraúna).  
<sup>1</sup> These authors contributed equally to this work.

<https://doi.org/10.1016/j.marpolbul.2020.111302>  
 Received 15 July 2019; Received in revised form 19 May 2020; Accepted 19 May 2020  
 Available online 01 June 2020  
 0025-326X/ © 2020 Elsevier Ltd. All rights reserved.

## A.2 Prêmio conferido pelo Comitê internacional de Microbiologia de alimentos e higiene pela 3ª melhor apresentação no 30º Congresso Brasileiro de Microbiologia

INTERNATIONAL COMMITTEE ON FOOD  
MICROBIOLOGY AND HYGIENE (ICFMH)  
INTERNATIONAL UNION OF MICROBIOLOGICAL SOCIETIES

ICFMH Award Receipt

1. Name of conference: 30<sup>th</sup> BRAZILIAN CONGRESS OF MICROBIOLOGY

Location (City, country): MACEIO AL, BRAZIL Date of conference: 2019 Oct 6-9

\*2. Received from ICFMH Executive Board representative: (Name, printed) BERNARDETTE DE M. FRAN  
(Signature) [Signature] the sum of 150 EUROS as prize for  
poster ☒ or oral ☐ presentation.

Full name of awardee in capital letters: DAVI MARCON

3. Signature of awardee: [Signature]

Date: 2019 OCT 09 E-mail address: DAV@MILCONSULTORES.COM.BR

Name of institution: FEDERAL UNIVERSITY OF PARA Country: BRAZIL

4. Permission to upload on ICFMH website (www.icfmh.org):

a. Photograph of winner being presented the award ☒ YES ☐ NO  
Signature [Signature]

b. Abstract and poster ☒ YES ☐ NO  
Signature [Signature]

\*Recipient of this award will comply with tax regulations in his/her country.