



**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS**  
**FACULDADE DE CIÊNCIAS BIOLÓGICAS E AMBIENTAIS**  
**CURSO DE BIOTECNOLOGIA**

**Ryan Fernandes Vieira de Souza**

Anotação funcional de sequências genômicas de polihidroxicanoato em *Priestia megaterium* cepa E1 isolada de solo

**DOURADOS, MS**

**2024**

**Ryan Fernandes Vieira de Souza**

Anotação funcional de sequências genômicas de polihidroxicanoato em *Priestia megaterium* cepa E1 isolada de solo

Trabalho de conclusão de curso apresentado como parte dos requisitos para obtenção do título de Bacharel em Biotecnologia no Curso de Graduação em Biotecnologia pela Universidade Federal da Grande Dourados.

Orientador: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Maricy Raquel Lindebah Bonfá

Coorientador: Prof. Dr. Rodrigo Matheus Pereira.

**DOURADOS, MS**

**2024**

## Dados Internacionais de Catalogação na Publicação (CIP).

S729a Souza, Ryan Fernandes Vieira De  
Anotação funcional de sequências genômicas de polihidroxicanoato em *Priestia megaterium*  
cepa E1 isolada de solo [recurso eletrônico] / Ryan Fernandes Vieira De Souza. -- 2024.  
Arquivo em formato pdf.

Orientador: Maricy Raquel Lindebah Bonfá.

Coorientador: Rodrigo Matheus Pereira.

TCC (Graduação em Biotecnologia)-Universidade Federal da Grande Dourados, 2024.

Disponível no Repositório Institucional da UFGD em:

<https://portal.ufgd.edu.br/setor/biblioteca/repositorio>

1. PHA. 2. genômica. 3. enzimas. I. Bonfá, Maricy Raquel Lindebah. II. Pereira, Rodrigo  
Matheus. III. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a)  
autor(a).

©Direitos reservados. Permitido a reprodução parcial desde que citada a fonte.

Ryan Fernandes Vieira de Souza

**Anotação funcional de sequências genômicas de polihidroxicanoato em *Priestia megaterium* cepa E1 isolada de solo**

Trabalho de conclusão de curso apresentado como parte dos requisitos para obtenção do título de Bacharel em Biotecnologia no Curso de Graduação em Biotecnologia pela Universidade Federal da Grande Dourados.

Orientador: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Maricy Raquel Lindebah Bonfá

Coorientador: Prof. Dr. Rodrigo Matheus Pereira.

Dourados, 03 de maio de 2024.

**Banca Examinadora**

Documento assinado digitalmente  
 **MARICY RAQUEL LINDENBAH BONFA**  
Data: 16/05/2024 09:57:10-0300  
Verifique em <https://validar.iti.gov.br>

---

Prof<sup>ª</sup>. Dr<sup>ª</sup> Maricy Raquel Lindenbah Bonfá - FCBA

Documento assinado digitalmente  
 **SUZANA EIKO SATO GUIMA**  
Data: 28/05/2024 03:20:44-0300  
Verifique em <https://validar.iti.gov.br>

---

Me. Suzana Eiko Sato Guima

Documento assinado digitalmente  
 **RODRIGO MATHEUS PEREIRA**  
Data: 15/05/2024 10:43:53-0300  
Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Rodrigo Matheus Pereira - FCBA

## Agradecimentos

É com gratidão que celebro a jornada trilhada na Universidade Federal da Grande Dourados (UFGD), onde a Faculdade de Ciências Biológicas e Ambientais (FCBA) se tornou o solo fértil para a semeadura do conhecimento, conduzindo-me ao título de Bacharel em Biotecnologia. Sob seus auspícios, encontrei os elementos essenciais e os caminhos para a conclusão deste curso.

Agradeço à minha orientadora e mentora, a professora Dr<sup>a</sup> Maricy Raquel Lindenbah Bonfá, e manifesto profunda admiração. No Laboratório de Metagenômica, desde 2019, mergulhei em um oceano de descobertas, nutrindo uma paixão irrefreável pela ciência e absorvendo os fundamentos práticos do laboratório que só foi possível pela oportunidade concedida pela professora.

Agradeço também ao meu coorientador, o professor Dr<sup>o</sup> Rodrigo Matheus Pereira, em que devo a revelação dos mistérios da Bioinformática e o desvelar de novos horizontes para o desenvolvimento deste presente trabalho.

Expresso minha gratidão aos meus familiares, em especial à minha mãe, Carlione Fernandes Vieira, cujo apoio incansável, motivação incessante e sonhos compartilhados foram o farol que guiou meus passos nesta jornada de conquistas.

Por último, agradeço aos meus amigos, fiéis companheiros de jornada e que compõem a minha grande família Douradense. Suas presenças foram rochas sólidas em meio às correntezas dos anos universitários, sustentando-me com força e coragem para perseverar até a chegada deste momento.

Na arte de aprender, encontramos a poesia da vida, onde cada descoberta é um verso, cada desafio é uma estrofe, e a gratidão é o refrão que ecoa em nosso coração.

## Resumo

Os plásticos são amplamente utilizados devido às suas propriedades vantajosas. No entanto, seu uso extensivo vem causando diversos problemas ambientais devido à poluição e à decomposição lenta, principalmente. Alternativas sustentáveis, como os bioplásticos derivados de fontes renováveis, estão ganhando interesse. Como exemplos disso são os polihidroxicanoatos (PHAs) como alternativas viáveis aos plásticos descartáveis (SUPs), que são produzidos por microrganismos como *Priestia megaterium* (anteriormente conhecido como *Bacillus megaterium*). A genômica adentra esse cenário desempenhando um papel essencial na identificação de genomas de microrganismos com potencial de biossíntese de PHAs para aplicação biotecnológica, especialmente através de programas de anotação funcional. O trabalho presente pretendeu anotar sequências genômicas de PHAs no genoma da bactéria *Priestia megaterium* cepa E1, isolada a partir de amostras de solo coletadas na região de Dourados, MS (2017). Foi realizada a anotação funcional de sequências genômicas da bactéria em questão usando as ferramentas Prokka e eggNOG-mapper. As sequências anotadas foram submetidas a alinhamentos locais nos bancos de dados secundários Swiss-Prot (UniProt) e RefSeq\_protein (NCBI). Os resultados também foram analisados no software InteractiVenn. Em seguida, as sequências foram submetidas a alinhamentos múltiplos globais e análises filogenéticas usando o ClustalX 2.1 e o MEGA 11, respectivamente. Uma análise físico-química foi realizada usando o software ProtParam, e a modelagem 3D das proteínas PhaC sintase foi feita usando o Phyre 2 e Swiss Model. A modelagem das outras proteínas foi limitada pela falta de dados disponíveis na literatura. Os resultados apresentaram 25 sequências de PHA, distribuídas entre diferentes enzimas como PhaA, PhaB e PhaI, entre outras. As árvores filogenéticas revelaram homologias entre as sequências analisadas com aquelas retiradas de banco de dados secundários, e a modelagem 3D das proteínas de PhaC sintase apresentou arquiteturas compatíveis com estudos anteriores. Esses resultados forneceram confiabilidade às anotações de PHAs no genoma da bactéria *Priestia megaterium* cepa E1, o que corrobora com evidências científicas já feitas na espécie.

**Palavras-chaves:** PHA; genômica; enzimas.

## ABSTRACT

Plastics are widely used due to their advantageous properties. However, its extensive use has caused several environmental problems mainly due to pollution and slow decomposition. Sustainable alternatives, such as bioplastics derived from renewable sources, are gaining interest. Examples of this are polyhydroxyalkanoates (PHAs) as viable alternatives to single-use plastics (SUPs), which are produced by microorganisms such as *Priestia megaterium* (formerly known as *Bacillus megaterium*). Genomics enters this scenario by playing an essential role in identifying genomes of microorganisms with potential for biosynthesis of PHAs for biotechnological application, especially through functional annotation programs. The present work intended to annotate genomic sequences of PHAs in the genome of the bacterium *Priestia megaterium* strain E1, isolated from soil samples collected in the region of Dourados, MS (2017). Functional annotation of genomic sequences of the bacterium in question was performed using the Prokka and eggNOG-mapper tools. The annotated sequences were subjected to local alignments in the Swiss-Prot (UniProt) and RefSeq\_protein (NCBI) secondary databases. The results were also analyzed using the InteractiVenn software. Then, the sequences were subjected to global multiple alignments and phylogenetic analyses using ClustalX 2.1 and MEGA 11, respectively. A physicochemical analysis was performed using ProtParam software, and 3D modeling of PhaC synthase proteins was done using Phyre 2 and Swiss Model. Modeling of the other proteins was limited by the lack of data available in the literature. The results showed 25 PHA sequences, distributed among different enzymes such as PhaA, PhaB and PhaI, among others. The phylogenetic trees revealed homologies between the analyzed sequences and those taken from secondary databases, and the 3D modeling of the PhaC synthase proteins presented architectures compatible with previous studies. These results provided reliability to the annotations of PHAs in the genome of the bacterium *Priestia megaterium* strain E1, which corroborates scientific evidence already carried out on the species.

**Keyword:** Polyhydroxyalkanoate; genomics; enzymes.

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>9</b>
<b>2 OBJETIVOS.....</b>	<b>10</b>
<b>2.1 Objetivos gerais.....</b>	<b>10</b>
<b>2.2 Objetivos específicos.....</b>	<b>11</b>
<b>3 REVISÃO LITERÁRIA.....</b>	<b>11</b>
<b>3.1 Biopolímeros.....</b>	<b>11</b>
<b>3.2 Polihidroxicanoatos.....</b>	<b>13</b>
3.2.1 Enzimas.....	18
3.2.2 Vias metabólicas.....	20
<b>3.3 Genômica.....</b>	<b>22</b>
<b>3.4 Bioinformática.....</b>	<b>23</b>
<b>4 MATERIAL E MÉTODO.....</b>	<b>25</b>
<b>4.1 Anotação de sequências.....</b>	<b>26</b>
<b>4.2 Alinhamento local.....</b>	<b>26</b>
<b>4.3 Diagrama de Venn.....</b>	<b>27</b>
<b>4.4 Alinhamento global e Análise filogenética.....</b>	<b>27</b>
<b>4.5 Análise físico química.....</b>	<b>27</b>
<b>4.6 Estrutura 3D.....</b>	<b>27</b>
<b>5 RESULTADOS E DISCUSSÕES.....</b>	<b>28</b>
<b>5.1 Anotação funcional.....</b>	<b>28</b>
<b>5.2 Blast.....</b>	<b>34</b>
<b>5.3 Diagrama de Venn.....</b>	<b>42</b>
<b>5.4 Análise filogenética.....</b>	<b>43</b>
<b>5.5 Análise físico química.....</b>	<b>59</b>
<b>5.6 Modelagem 3D.....</b>	<b>63</b>
<b>6 CONCLUSÕES.....</b>	<b>69</b>
<b>REFERÊNCIAS.....</b>	<b>70</b>

## LISTA DE FIGURAS

**Figura 1.** (European Bioplastics, 2020) Crescimento global de produção de bioplásticos.

**Figura 2.** (Alwuhaib, A. S. et al., 2023) Estrutura química dos polyhydroxialcanoatos.

**Figura 3.** (Acuña, J.M.B; Poblete-Castro, I., 2023) Cluster de PHA e classes e substratos de polimerase de PHA necessários em diversas cepas nativas produtoras de PHA.

**Figura 4.** (Peoples, Sinskey, 1989) Vias metabólicas de produção de PHA.

**Figura 5.** Fluxograma das atividades de bioinformática desenvolvidas no presente trabalho.

**Figura 6.** Representação gráfica da interação das sequências do genoma da bactéria *Priestia megaterium* cepa E1 nos programas Prokka, eggNOG-mapper, UniProt (Swiss-Prot) e Reference Sequence (RefSeq) Protein.

**Figura 7.** Árvore filogenética construída a partir das sete sequências de PhaC sintase presente no genoma da bactéria *Priestia megaterium* cepa E1 alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.

**Figura 8.** Árvore filogenética construída a partir de três sequências de PhaA alinhadas com as sequências mais similares, juntamente com um grupo externo do domínio de Archaea, retiradas do banco de dados RefSeq Protein.

**Figura 9.** Árvore filogenética construída a partir das três sequências de PhaB alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein

**Figura 10.** Árvore filogenética construída a partir das cinco sequências de proteínas fasinas alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.

**Figura 11.** Árvore filogenética construída a partir da única sequência PhaR alinhada, juntamente com dois grupos externos do domínio de Archaea, com a sequência mais similar retirada do banco de dados RefSeq Protein.

**Figura 12.** Árvore filogenética construída a partir das duas sequências de proteínas despolimerases, juntamente com um grupo externo do domínio de Archaea, alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.

**Figura 13.** Árvore filogenética construída a partir das duas sequências de PhaJ da família MaoC desidratase, juntamente com um grupo externo do domínio de Archaea, alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.

**Figura 14.** Visualização da modelagem da proteína Poly-beta-hydroxybutyrate polymerase (PhaC2) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.

**Figura 15.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.

**Figura 16.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.

**Figura 17.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.

**Figura 18.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.

**Figura 19.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase de classe III subunidade PhaC do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no modelo cartoon, à direita.

**Figura 20.** Visualização da modelagem da proteína subunidade PhaE de poli(3-hidroxiálcanoato) polimerase de classe III do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Swiss-model exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no modelo cartoon, à direita.

**Figura 21.** Gráfico de Ramachandran da modelagem da proteína subunidade PhaE de poli(3-hidroxiálcanoato) polimerase de classe III (EAMNKLGA\_22716) realizada pelo Swiss-model.

## LISTA DE TABELAS

**Tabela 1.** As 25 sequências filtradas envolvidas no metabolismo de PHA a partir da anotação feita pela ferramenta Prokka.

**Tabela 2.** Anotação das 25 sequências envolvidas no metabolismo de PHA pela ferramenta eggNOG-mapper.

**Tabela 3.** Alinhamento local das 25 sequências de proteínas realizado no banco de dados Swiss-Prot.

**Tabela 4.** Comparação das descrições feitas pelas ferramentas de anotação Prokka e eggNOG-mapper e pelos alinhamentos no Swiss-Prot.

**Tabela 5.** Alinhamento local das 25 sequências de proteínas realizado no banco de dados Reference Sequence Protein.

**Tabela 6.** Anotações correspondentes entre as ferramentas de anotação funcional, Prokka e eggNOG-mapper, e os bancos de dados secundários, Swiss-prot e RefSeq Protein.

**Tabela 7.** Resultado das características físico químicas obtido no software ProtParam das 25 sequências anotadas.

## LISTA DE ABREVIACÕES

PHAs - Polihidroxicanoatos

SUPs - Plásticos descartáveis

PP - Polipropileno

PE - Polietileno

pH - Potencial hidrogeniônico

PHB - Poli(3-hidroxiobutirato)

scl-PHAs - Polihidroxicanoatos de cadeia curta

mcl-PHA - Polihidroxicanoatos de cadeia média

lcl-PHA - Polihidroxicanoatos de cadeia longa

UV - Ultra Violeta

PLA - Poliacido Láctico

P[3HB-co-4HB] - Poli(3-hidroxiobutirato-co-4-hidroxiobutirato)

UFGD - Universidade Federal da Grande Dourados

LMSEQ - Laboratório Multiusuário Centralizado para Sequenciamento de DNA em Larga Escala e Análise de Expressão Gênica

NCBI - National Center for Biotechnology Information

PM - Peso Molecular

pI -Ponto Isoelétrico

pII - Índice de Instabilidade

IA - Índice Alifático

GRAVY - Grand Average of Hydropathy

CE - Coeficiente de Extinção

## 1 INTRODUÇÃO

Os plásticos são amplamente utilizados devido a suas propriedades vantajosas. No entanto, o seu uso extensivo tem levantado sérias preocupações ambientais, dado que a produção e o descarte inadequado contribuem à poluição e impactos adversos ao meio ambiente (Sakthipriya, N, 2022). Além disso, a decomposição lenta dos plásticos tradicionais, que podem persistir no ambiente por séculos, agrava ainda mais esses problemas (Melchor Martínez, E. M. et al., 2022).

Diante dessas preocupações, cresce o interesse em alternativas sustentáveis, como os bioplásticos derivados de fontes renováveis, como amido e celulose, ou produzidos por microrganismos (Kumar, R. et al., 2023). Os polihidroxialcanoatos (PHAs) destacam-se nesse cenário pela sua capacidade de biodegradabilidade e biocompatibilidade, apresentando-se como alternativas ecologicamente viáveis aos plásticos descartáveis (SUPs), como polietileno e propileno (Wen, Q. et al., 2021). Essa classe de poliésteres, produzida por microrganismos como *Priestia megaterium* (anteriormente *Bacillus megaterium*) (Liu, J. M. et al., 2023), apresenta uma gama versátil de propriedades, tendo diversas aplicações biotecnológicas, desde tecnologia na área da saúde até embalagens e na agricultura (Jaffur, B. N. et al., 2023).

A biossíntese de PHA é mediada por uma série de enzimas, as quais são transcritas por genes que são homólogos entre diferentes organismos. À medida que a tecnologia avança, surgem novas metodologias para a prospecção dessas enzimas e os genes relacionados com as suas sínteses. A genômica, que envolve o estudo do material genético completo de um organismo, destaca-se como uma abordagem bastante abrangente (Yu, S. et al., 2022). Essa área da bioinformática proporciona uma visão ampla e detalhada da composição de um genoma, através de anotações, alinhamentos e análises filogenéticas, permitindo uma identificação mais precisa de enzimas e potenciais alvos biotecnológicos (Wang, R. C., 2023).

Nesse contexto, diversos programas desempenham papéis essenciais, entre eles o Prokka (Seemann, T., 2014; Afgan, E. et al., 2018) e o eggNOG mapper (Cantalapiedra, C. P, et al., 2021), e são amplamente utilizados para executar tarefas específicas para identificar e anotar sequências de interesse e o Mega 11 (Koichiro, T. et al., 2007), na execução de análises filogenéticas.

Dessa forma, a combinação destas diferentes ferramentas juntamente com alinhamentos locais feitos em bancos de dados secundários permite a anotação funcional de sequências genômicas, o que possibilita a identificação de genes e proteínas envolvidas no metabolismo de polihidroxialcanoatos a partir do genoma da bactéria *Priestia megaterium* isolada de solo.

## 2 OBJETIVOS

### 2.1 Objetivos gerais

Realizar a anotação funcional de sequências genômicas de polihidroxicanoato na bactéria *Priestia megaterium* cepa E1 isolada a partir de amostras de solo.

### 2.2 Objetivos específicos

- Identificar e comparar anotações de PHA feitas em diferentes ferramentas;
- Realizar alinhamento local com sequências de PHA em diferentes bancos de dados secundários;
- Analisar a filogenia através do alinhamento múltiplo global de sequências de PHA;
- Examinar as propriedades físico químicas das sequências anotadas e analisar as estruturas de enzimas específicas.

## 3 REVISÃO LITERÁRIA

### 3.1 Biopolímeros

O alerta feita pela ONU, durante a Conferência da ONU sobre Mudanças Climáticas, em 2021, aponta:

“As principais economias mundiais vão produzir, até 2030, mais do dobro da quantidade de carvão, petróleo e gás dos níveis necessários para manter vivas as metas climáticas estabelecidas no Acordo de Paris. Os planos dos governos para extrair combustíveis fósseis até 2030 são incompatíveis com a manutenção das temperaturas globais em níveis seguros”.

Esse comentário é reflexo da exploração escassa de fontes não renováveis pelas principais nações mundiais. Em consequência disso, a mudança climática gerada é superior à capacidade de manutenção de temperatura pelo planeta, o que excede as metas estipuladas pelo Acordo de Paris. Dessa forma, existe uma forte pressão internacional que vem

estimulando a produção de materiais a base de matérias primas de fontes renováveis (Xie, Y. et al., 2023).

Os polímeros de fontes renováveis são vistos como uma alternativa promissora para diminuir a poluição ambiental gerada por polímeros petroquímicos, principalmente os plásticos, devido a sua alta capacidade de biodegradação. Também conhecidos como biopolímeros, eles são uma classe de compostos naturais produzidos por organismos vivos ou por processos químicos de matérias primas renováveis. Sejam eles de origem vegetal, animal ou microbiana (Jiménez, M. M. et al., 2023).

Essas macromoléculas são constituídas por monômeros, pequenas unidades que podem ser de açúcares, nucleotídeos e aminoácidos. Enquanto as fontes fósseis demoram anos para se formarem, no caso dos polímeros à base de petróleo, os biopolímeros são formados em um tempo muito menor. Eles são a base de fontes naturais, as quais estão em constante renovação. Além da característica de biodegradação, a sua produção não gera impacto ambiental, diferente dos polímeros petroquímicos (Xie, Y. et al., 2023).

A capacidade termoplástica dos biopolímeros é uma das principais características que fomenta a sua inserção no mercado. Isso significa que esses materiais podem ser moldados e reprocessados a temperaturas elevadas, o que os torna versáteis para serem aplicados em diversas áreas. Assim, eles podem substituir os plásticos convencionais em muitos produtos, superando impasses como o aumento do preço e o esgotamento do petróleo, além dos impactos ambientais gerados (Khalil, A. H. P. S. et al., 2023).

Os biopolímeros já vêm sendo aplicados em diferentes setores, incluindo embalagens, agricultura, medicina, entre outros. Como exemplo disso, o ácido polilático (PLA), um biopolímero derivado do amido de milho, é usado em embalagens biodegradáveis, copos descartáveis, filmes plásticos e até mesmo em dispositivos médicos. Outro exemplo é o polihidroxialcanoato (PHA), que é produzido principalmente por bactérias e pode ser utilizado na fabricação de produtos como sacolas biodegradáveis e embalagens funcionais (Dana, R. H.; Ebrahimi, F, 2023).

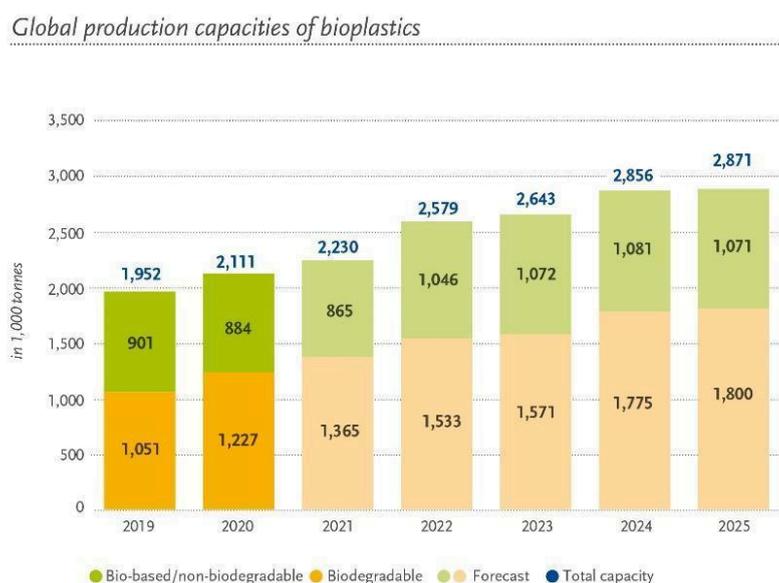
Os bioplásticos são os principais tipos de biopolímeros. Eles apresentam fortes semelhanças com o polipropileno e polietileno e são empregados no segmento de embalagens. Apesar de seu consumo ainda ser baixo, as perspectivas de aumento são promissoras, devido aos aspectos ambientais e econômicos relacionados ao consumo do petróleo pelos países desenvolvidos (Tennakon, P. et al., 2023).

Esses bioplásticos oferecem uma ampla gama de funcionalidades e são otimizados para cada tipo de aplicação. Além disso, já existem muitos plásticos biodegradáveis no

mercado. Dentre aqueles que merecem maior ênfase são os PHAs, PLAs e misturas de amido-polímero (Kim, M. S. et al. 2023).

Em 2020, a capacidade mundial de produção de bioplásticos era de 2,1 milhões de toneladas e há fortes evidências de que até 2025 esse número supere 2,8 milhões de toneladas (Figura 1), de acordo com European bioplastics (Rosenboom, J. G. et al., 2022). Esse crescimento é impulsionado pelos biopolímeros inovadores, como os PLA e, principalmente, os PHAs de bases biológicas. Quanto a este último, espera-se que tenha um aumento na capacidade de produção em quase sete vezes até 2025 (Shaikh, S. et al., 2023).

**Figura 1.** (European Bioplastics, 2020) Crescimento global de produção de bioplásticos 2019 a 2025.



Nesse sentido, os biopolímeros representam uma alternativa promissora e sustentável aos polímeros tradicionais à base de petróleo. No entanto, apesar de suas vantagens, os biopolímeros ainda enfrentam desafios, como o custo de produção comparado aos polímeros petroquímicos (Rahman, M. Z. et al., 2023). Dessa forma, à medida que a conscientização sobre os impactos ambientais dos polímeros não renováveis continua a crescer, é esperado que a demanda por biopolímeros continue a aumentar, diminuindo o preço de produção e impulsionando o crescimento dessa indústria nos próximos anos, validando as expectativas apontadas pela European Bioplastics em 2020.

### 3.2 Polihidroxicanoatos

Os polihidroxicanoatos são os candidatos líderes entre as alternativas potenciais aos plásticos derivados de petróleo. Também conhecidos como PHAs tornaram-se cada vez mais importantes em indústrias tão diversas como tecnologia biomédica, automotiva, embalagens e agricultura devido a versatilidade de suas propriedades (Jaffur, B. N. et al., 2023). Foram descritos como potenciais biomateriais que podem substituir plásticos descartáveis (SUPs) feitos de polipropileno (PP) e polietileno (PE) de diversas áreas da humanidade (Melchor Martínez, E. M. et al., 2022).

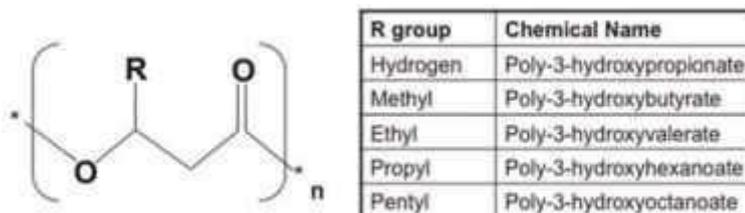
Naturalmente, eles são armazenados como grânulos intracelulares e fornecem reservatórios intermediários para o fornecimento de carbono e energia (Pei, R. et al., 2023). Além disso, são constituídos por monômeros de ácidos hidroxicanoicos, os quais contêm um grupo carboxílico e um grupo hidroxila ligados a mesma cadeia de carbonos (Kumar, et al. 2020).

Esses biopolímeros são sintetizados no momento que a célula estiver sob estresse, como excesso de fonte de carbono, limitação de nitrogênio, pressão osmótica elevada, oscilação de temperatura e pH, além de outros. Dessa forma, são digeridos assim que a célula necessitar (Obruca, S. et al., 2021). Esse processo oferece a capacidade dos microrganismos de sobreviver em ambientes pobres em nutrientes, pois aumenta sua resistência ao estresse local (Jaffur, B.N. et al., 2023).

A primeira vez que os PHAs foram relatados foi em 1926 pelo microbiologista Maurice Lemoigne. Em estudo notou pequenos grânulos produzidos no interior da bactéria *Priestia megaterium* (Alves, A. A. et al. 2023). Esses grânulos são poliésteres alifáticos e foram identificados como PHB (poli(3-hidroxiбутirato)). Os resultados do estudo foram publicados em revistas francesas pouco lidas. Nessa época, os microbiologistas não tinham interesse em lipídeos e desacreditaram que existiam macromoléculas biológicas como polímeros. Assim essa descoberta passou-se um bom tempo despercebida (Rojo, S. G.; Antolínez, D. R, 2023).

O poliéster que Lemoigne isolou e caracterizou foi o poli(3-hidroxiбутirato), o PHB, de monômeros de ácido [R]-3-hidroxiбутírico, contendo um grupo metil ligado a cadeia principal, como indicado na Figura 2 (Alves, A. A. et al. 2023)

**Figura 2.** (Alwuhuib, A. S. et al., 2023) Estrutura química dos polyhydroxicanoatos.



Essa classe de polímeros é referida geralmente como PHAs e todos os membros da classe têm a mesma configuração para o centro quiral do poliéster como mostra a estrutura acima. O que difere é o tipo de grupo R ligado na cadeira principal de carbonos (Rojo, S. G.; Antolínez, D. R, 2023).

Na época, enquanto tentava determinar um problema nas culturas da bactéria *Priestia megaterium*, Lemoigne descobriu a existência do PHB (Pramanik, N., 2023). Em 1923, o cientista publicou as suas descobertas e quatro anos depois, em 1927, descreveu o isolamento de um produto formado no interior de uma célula, descrito como um polímero formado a partir do ácido em questão (Alves, A. A. et al. 2023). Nesse mesmo período, ele chegou a sugerir que o polímero formado dentro da célula ocorria por uma “polimerização por desidratação” (Lemoigne, M., 1926).

Lemoigne e colaboradores relataram seus resultados de polihidroxicanoatos em 27 publicações de 1923 a 1951, e em estudos posteriores descobriram que cepas da bactéria *Priestia megaterium* poderiam conter até 44% do seu peso seco desse polímero (Katagi, V. N. et al., 2023). Apesar das inúmeras publicações, o papel do PHB no metabolismo das bactérias só foi reconhecido no final da década de 1950 e aí que se percebeu a importância dos primeiros estudos feitos por Lemoigne (Kapoor, D. D. et al., 2023).

Quanto à estrutura do PHA, ela é composta por monômeros de hidroxiácidos, como já visto por Lemoigne. Para classificá-la, foi proposto três subconjuntos, dependendo da quantidade de átomos de carbono dos monômeros. O primeiro subconjunto corresponde aos scl-PHAs, que são polímeros de cadeias curtas de 3-5 carbonos. O segundo são os mcl-PHA, de cadeias médias de 6-14 carbonos. Já o terceiro, o menos comum, trata-se dos lcl-PHA, de 15 ou mais carbonos (Liu, J. M. et al., 2021). A enorme diversidade de PHA se deve principalmente à alta aceitação de uma extensa gama de substratos pelas PHA sintases (PhaC) (Grigore, M. E., 2019).

Com base nas composições e tipos de monômeros, muitos homopolímeros e copolímeros de PHA são produzidos para apresentar diversas propriedades de materiais. Por exemplo, scl-PHA é um termoplástico rígido e resistente com alto comportamento cristalino.

Em contraste, mcl-PHA é um polímero elástico e elastomérico semelhante a borracha. Essas diversas propriedades do material são uma das características significativas do PHA que permite a síntese de polímeros modificados para usos preferenciais (Guleria, S. et al., 2022).

Os PHAs mais comuns encontrados em produtos no mercado são os polihidroxibutiratos (PHB), que são polímeros de cadeias curtas (Meereboer, K. W. et al. 2020). No entanto, existe mais de 150 tipos diferentes de PHAs que foram identificados e classificados de acordo com sua unidade monomérica como homopolímeros, contendo um tipo de unidade monomérica, ou heteropolímeros, com mais de um tipo de unidade monomérica (copolímero) (Katagi, V. N. et al., 2023).

A maioria dos microrganismos que acumulam PHA produzem scl- ou mcl-PHA porque a PHA sintase (PhaC) prefere scl- ou mcl-hidroxiacil-CoA como substrato. No entanto, poucas sintases de PHA atuam tanto em scl- quanto em mcl-hidroxiacil-CoAs produzindo scl-mcl-PHA, um copolímero (Gutschmann, et al. 2022).

Como já mencionado anteriormente, os PHAs têm sido descritos como fortes candidatos que podem substituir os plásticos descartáveis (SUPs) feitos de polietileno (PE) e polipropileno (PP). Além do setor de embalagens, eles também podem ser empregados na engenharia de tecidos, transportadores de medicamentos e suplementos nutricionais (Saeed, M. et al., 2022).

A escolha dos PHAs na substituição de SUPs deve-se às suas características de biodegradabilidade, hidrofobicidade e não toxicidade. Além disso, por serem biocompatíveis com processabilidade térmica, permitem ser moldados de diferentes maneiras (Butt, N. et al., 2018). Eles também são insolúveis em água, resistentes aos raios UV e apresentam boa resistência ao ataque hidrolítico. A sua degradação depende principalmente do seu tipo e composição, das condições ambientais e do tipo da despolimerase para degradar os PHAs (Pakalapati, H. et al. 2018).

Por conta de tais características, a procura por esse biopolímero tem aumentado ano após ano desde os primeiros estudos. De acordo com dados recentes, a produção de PHA triplicou de 5,3 para 17,0 milhões de toneladas entre 2013 e 2020. No início desse período, os plásticos biodegradáveis tinham um valor de mercado de 3,02 milhões de dólares em 2018, que deverá atingir 6,73 milhões de dólares até 2025 (Jaffur, N. B. et al., 2023).

Apesar das altas expectativas apontadas por Jaffur e outros estudos, antes que o PHA possa substituir a cadeia de produção de polímeros à base de petróleo, existem gargalos pertinentes. Para aplicação em larga escala, a geração do PHA foi testada a partir de uma

variedade de fontes de carbono e diferentes organismos. No entanto, o custo das fontes de carbonos representa pelo menos metade do custo de produção (Saratale, R. G. et al., 2021).

Os custos elevados de fabricação de PHAs chegam a cerca de US\$ 4.000-15.000/Mt. Enquanto o custo de fabricação de plásticos extraídos do petróleo é de aproximadamente US\$ 1.000-1.500/Mt. Essa diferença é extremamente grande, o que constitui o principal obstáculo que deve ser abordado para viabilizar a comercialização dos PHAs (Raza, Z. A. et al., 2023).

Diversas pesquisas apontam que o preço de produção de PHAs situa-se em torno de 16 dólares por kg, o que representa um custo 18 vezes superior em comparação com os polipropilenos (PP) e polietileno (PE) convencionais. No entanto, é possível reduzir significativamente esse custo para apenas quatro dólares por kg mediante a implementação de estratégias altamente produtivas, como destacado por Adeleye e colaboradores em 2020.

Diante desse cenário, para produzir PHA com boa relação custo-benefício em larga escala, diversos estudos estão sendo realizados para buscar alternativas promissoras. A reutilização de biomassa residual pode ser um excelente substrato de carbono para viabilizar o custo de produção (Raza, Z. A. et al., 2023).

A biomassa residual normalmente inclui resíduos industriais, agrícolas ou domésticos com alto teor de carbono. Produtos residuais das indústrias de alimentos e amido, óleos vegetais, efluentes industriais orgânicos e subprodutos do setor de biodiesel têm sido usados como fontes de carbono para a síntese de PHA (Mahato, R. P. et al., 2023).

Para exemplificar, um estudo realizado por Melo e colaboradores (2023) na síntese de PHB pela cepa *Priestia megaterium* utilizou substratos de fontes renováveis (efluente da indústria) e verificou que há uma redução de custo de aproximadamente US\$ 1,0 por litro de meio de cultura comparado ao meio sintético à base de sacarose.

Desta forma, a escolha adequada de matérias-primas de baixo valor agregado pode ajudar a reduzir os custos de produção e minimizar as rotas de transporte, ao mesmo tempo que minimiza os impactos ambientais negativos (Guleria, S. et al. 2022).

Outro caminho promissor é o da engenharia genética, que atua para melhorar o crescimento de bactérias e aumentar a produtividade dos PHAs. Em estudo publicado por Cal e colaboradores (2021), uma cepa de *Priestia megaterium* foi “engenhariada” para gerar copolímeros de poli(3-hidroxibutirato-co-4-hidroxibutirato) (P[3HB-co-4HB]) (um dos mais industrialmente relevantes). As suas vias metabólicas foram projetadas para produzir P[3HB-co-4HB] com menor fração molar de glicose. Sendo assim, a quantidade de substrato suficiente é utilizada diminui e o custo de produção do copolímero barateia.

### 3.2.1 Enzimas

Dentre as inúmeras enzimas que participam da produção de PHA, a PhaC sintase é considerada a principal enzima desse processo. Esta enzima está presente em todos os microrganismos bioacumuladores de PHA e existem quatro classes principais, diferindo pelas suas preferências de substrato e suas estruturas primárias (Neoh, S. Z et al., 2022).

As PHA sintases de classe I são compostas por uma única subunidade de 60-70kDa e essas enzimas catalisam a polimerização de PHAs de cadeia curtas (scl-PHA) (Tsuge, 2016). Ela é codificada pelo gene *phaC*, encontrado geralmente agrupado no operon *phaCAB*, constituído pelos genes *phaC*, *phaA* e *phaB*, os quais são responsáveis pela síntese desse tipo de PHA, por exemplo P(3HB) (Li, Z. et al., 2017).

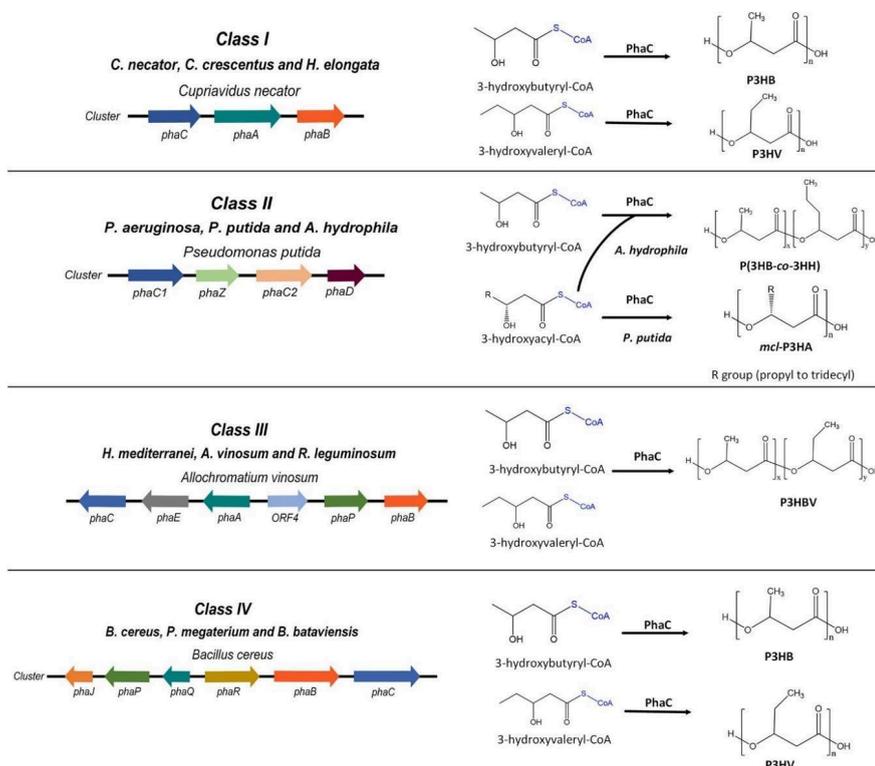
As PHA sintases de classe II são compostas por uma única subunidade de 60-70kDa também e são reguladas por dois genes, *phaC1* e *phaC2* (Zou, H. et al., 2017). Estas enzimas catalisam a polimerização de PHA de cadeias médias (mcl-PHA). Os dois genes codificam para proteínas que regulam a síntese de PHA, polimerizando CoA thioesters de monômeros MCLs. Geralmente eles são encontrados organizados no cluster *phaIFDC2ZC1*, que inclui vários genes como *phaZ*, *phaD*, *phaI* e *phaF*. O gene *phaZ* está presente entre *phaC2* e *phaC1* e codifica a enzima PHA despimerase (Tan, I. K. P, 2020). Em contraste, o gene *phaD* está localizado a jusante desses três genes e regula o número e o tamanho dos grânulos de PHA produzidos dentro da célula (Tarazona, N. A. et al., 2020). Os outros dois genes, *phaI* e *phaF*, estão situados colinearmente com *phaD* e estão associados à regulação das fasinas, que são proteínas responsáveis pela estabilização do grânulo de PHA (Aghaali, Z. et al., 2023). Outra função do gene *phaF* é regular a expressão do gene *phaC1*, regulando assim a síntese de PHA. Em condições ambientais desfavoráveis, a falta de substrato leva à interrupção da síntese de PHA, mas a síntese de proteína associada a grânulos, as fasinas (PhaF), permanece inalterada. Isto resulta na ligação da proteína PhaF que se move livremente à região promotora do operon, suprimindo assim a expressão do gene *phaC1*. No entanto, no advento de condições favoráveis, a síntese de PHA é retomada e a proteína PhaF liga-se ao grânulo de PHA e o *phaC1* é expresso (Mitra, R. et al., 2021).

As PHA sintases de classe III, diferente das duas anteriores, são constituídas por duas subunidades, PhaC e PhaE. Essas subunidades são codificadas pelos genes *phaC* e *phaE* que podem ser encontrados no operon *phaEC*. As duas subunidades contêm aproximadamente 40kDa cada e catalisam a polimerização de HAcOAs de cadeias curtas (SCLs) (MCCOOL, G. J. et al., 2021)

As PHA sintases (PhaCs) de classe IV (presente em *P. megaterium*) incluem heterodímeros como as de classe III, mas a subunidade PhaE é substituída pela PhaR. As duas subunidades, PhaC e PhaR, são reguladas pelos genes *phaC* e *phaR*, respectivamente. A bactéria *P. megaterium* é a espécie representativa desta classe, onde a enzima catalisa a polimerização de SCL-HACoAs (Zou et al., 2017). O operon *phaRBC* das PHA sintases de classe IV regula a síntese de PHA. O gene *phaB* está localizado entre os genes *phaC* e *phaR* (Mitra, R. et al., 2021).

Mohapatra e colaboradores (2021) estudaram em 45 espécies bacterianas o número de genes associados à síntese de PHA. Como resultado, observaram que o número poderia ser superior a 59 genes presentes no organismo das bactérias. Eles afirmam que o número de genes da PHA sintase (PhaC) pode variar entre as bactérias, mas geralmente contém pelo menos duas cópias de cada gene.

**Figura 3.** (Acuña, J. M. B.; Castro, I. P, 2023) Cluster de PHA e classes e substratos de polimerase e PHA necessários em diversas cepas nativas produtoras de PHA.



O polihidroxitirato (PHB) foi estudado extensivamente e foi constatado na maioria das bactérias que é regulado pelo operon *phaCAB* (Fig. 3). No entanto, em algumas espécies incluindo *Acinetobacter sp.*, *Agnes latus*, *P. acidophila* e *R. eutropha*, os genes *phaC*, *phaA* e *phaB* estão em conjunto, mas não necessariamente na mesma ordem no cromossomo. Como

também, os genes *phaA*, *phaB* e *phaC* podem estar separados e não ligados entre si, como acontece em *Paracoccus denitrificans*, *Rhizobium meliloti* e *Zoogloea ramigera* (Umeda, F. et al., 1998).

Paralelo ao que foi dito até então sobre as enzimas envolvidas na síntese de PHA, outro ponto muito importante é a biodegradação. Existem duas enzimas principais responsáveis que são a PHA hidrolase (PhaY) e a PHA despolimerase (PhaZ). Essas enzimas podem agir em condições aeróbicas ou anaeróbicas, resultando em diferentes produtos. Na primeira condição, a degradação do PHA resulta em CO<sub>2</sub> e H<sub>2</sub>O, e na segunda, os produtos são CO<sub>2</sub> e CH<sub>4</sub> (Kumar, M. et al., 2020). Além disso, a degradação do PHA depende de alguns fatores, como a composição química do polímero, o comprimento da cadeia polimérica, a cristalinidade e a complexidade do grânulo (Choi, S.Y. et al., 2020).

### 3.2.2 Vias metabólicas

O primeiro estudo sobre a biossíntese de polihidroxialcanoatos foi feito por Stanier e Wilkinson em 1959. Eles determinaram que os grânulos de PHB são produzidos nas células de *Bacillus sp.* quando uma cultura de bactéria está sob uma certa limitação de certos nutrientes, como o nitrogênio. Essa carência gera uma resposta que sinaliza a ativação de vias metabólicas específicas na célula microbiana (Kapoor, D. D. et al., 2023).

Desde então diversos estudos genômicos e metabólicos permitem a compreensão da biossíntese de PHAs. A sua produção pode ser dividida em duas etapas principais: a geração de hidroxil-acil-CoA e a polimerização de hidroxil-acil-CoA (Muneer, F. et al., 2020). Na primeira etapa, existem três vias metabólicas principais. Sendo estas que são a síntese de acetoacetyl-CoA (via I), a biossíntese de ácidos graxos (via II) e a  $\beta$ -oxidação de ácidos graxos (via III). O acetyl-CoA e o acil-CoA são intermediários comuns em todas as três vias (Choi, S. Y. et al., 2020). Na segunda etapa, o produto da primeira é catalisado pela enzima PHA sintase (PhaC), como esquematizado na Figura 5. A via metabólica utilizada depende da cepa bacteriana e dos substratos disponíveis (Haddadi, M. H. et al., 2019).

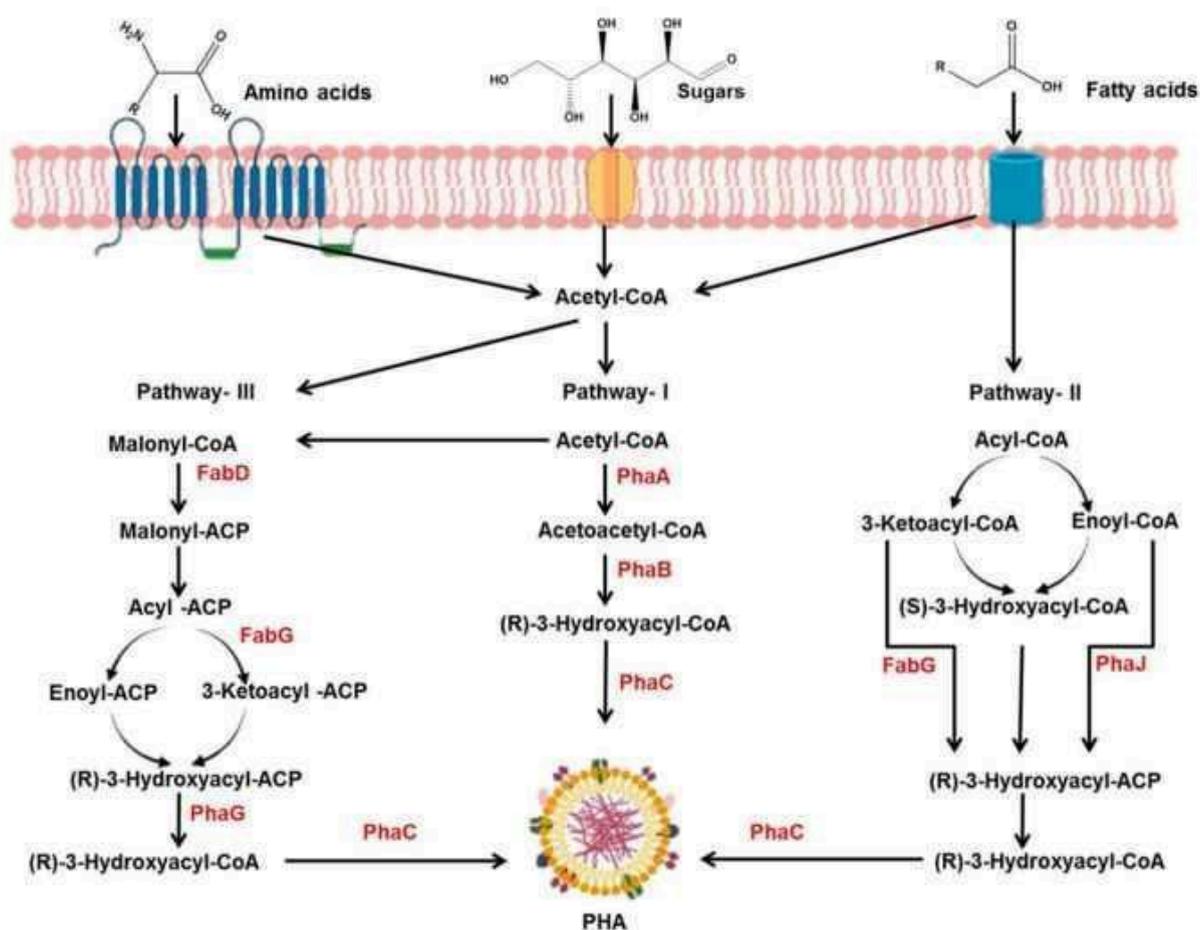
Na via I, três enzimas principais atuam (PhaA, PhaB e PhaC) para produzir o PHA. Nesta via a primeira enzima ativada é a PhaA, também conhecida como  $\beta$ -ketothiolase (ketoacyl-CoA thiolase), que se liga a moléculas de acetyl-CoA como cofatores para produzir acetoacetyl-CoA. A conversão de acetoacetyl-CoA em monômero (R)-3-hydroxybutyryl-CoA é realizada por PhaB, uma acetoacetyl-CoA redutase dependente de NADPH. A última enzima

ativada é a PhaC, uma PHA sintase, que vai catalisar o produto da rota em grânulo de PHA (Madison, L. L.; Huisman, G. W, 1999; Samadhiya, A. et al., 2022).

Na via II é ativada quando existem ácidos graxos disponíveis para o microrganismo durante a síntese de PHA. Após a  $\beta$ -oxidação dos ácidos graxos, o acil-CoA produzido sintetiza monômeros de PHA. Nesta via, enzimas incluindo 3-cetoacil-CoA redutase, epimerase, (R)-enoil-CoA hidratase/enoil-CoA hidratase I, acil-CoA oxidase estão envolvidas. A molécula 3-hidroxiacil-CoA funciona como uma molécula precursora para a síntese de PHA (Costa, C. Z. et al., 2015).

A via III para a síntese de PHA requer duas enzimas principais, 3-hidroxiacil-ACP-CoA transferase (PhaG) e malonil-CoA-ACP transferase (FabD). Estas enzimas fornecem o precursor (3-hidroxiacil-ACP), que é posteriormente convertido em 3-hidroxiacil-CoA. A próxima etapa é catalisada pela PhaC sintase para sintetizar PHA (Zhang, Z. et al., 2020).

**Figura 4.** (Behera, S. et al., 2022) Vias metabólicas de produção de PHA.



No caso da *B. megaterium*, estudos sugerem que a biossíntese de P(3HB), o mesmo composto identificado por Lemoigne, utiliza a via I que começa com duas moléculas de acetil-CoA sendo convertidas em acetoacetil-CoA na presença de uma enzima acetil-CoA acetiltransferase (PhaA). Posteriormente, a acetoacetil-CoA redutase (PhaB) dependente de NADPH ajuda na redução de acetoacetil-CoA a hidroxibutiril-CoA. Este composto é então usado como monômero para polimerizar P(3HB) com ácido poli-hidroxicarboxílico (phaC) sintase (Behera, S. et al., 2022).

### 3.3 Genômica

A genômica é um ramo das ciências ômicas que se concentra no estudo individual do genoma de um só organismo, fornecendo importantes informações sobre os genes e seus mecanismos de expressão. O seu surgimento só foi possível a partir do desenvolvimento de técnicas de sequenciamento que permitiram uma análise detalhada das sequências de DNA, possibilitando o estudo e a busca por genes alvos (Rocha, R. F. B. et al., 2023).

À medida que os estudos genômicos avançam, é possível impulsionar diversas áreas do mercado. Dentre elas está a agricultura, em que os estudos genômicos têm o potencial de impulsionar de maneira mais rápida o desenvolvimento de culturas aprimoradas, desempenhando um papel crucial na segurança alimentar na fase das mudanças climáticas. Isso tem impacto positivo em nações em desenvolvimento, como aquelas do continente africano (Wang, R. C., 2023).

Em trabalho proposto por Nascimento e colaboradores (2020), a genômica foi aplicada para aprimorar o rendimento das colheitas diante das variações de temperatura. Este estudo emprega uma cepa de *Priestia megaterium*, que é uma bactéria reconhecida por promover o crescimento das plantas. Essa bactéria é conhecida por sua versatilidade e resistência a estresses ambientais, demonstrando um potencial significativo para aplicações na agricultura e biotecnologia.

Na esfera da saúde, um outro exemplo é a aplicação da genômica na África, conforme investigado por Karikari e Aleksic (2015). O foco deste estudo reside na obtenção de informações relacionadas a doenças cerebrais, notadamente nas enfermidades neurodegenerativas, como o Parkinson e o Alzheimer. Essa abordagem tem o potencial de aprimorar significativamente a precisão nos diagnósticos e tratamentos dessas condições.

Dessa forma, ao incorporar avanços das tecnologias genômicas, os governos têm a chance de aprimorar significativamente a saúde pública por meio da implementação de novas

tecnologias (Riaz, A. et al., 2021). Além disso, possibilita o desenvolvimento de culturas mais produtivas, dotadas de resistência tanto a estresses abióticos quanto bióticos (Yu, S. et al., 2022).

No Brasil, os primeiros passos na área de genômica podem ser atribuídos a cerca de 1992, quando os pesquisadores Sergio Pena e Andrew Simpson realizaram estudos investigando marcadores de sequência genética no helminto *Schistosoma mansoni*, o causador da doença esquistossomose. Entretanto, foi nos anos 2000 que a genômica no Brasil ganhou destaque, especialmente com o sequenciamento completo do genoma da *Xylella fastidiosa*, um patógeno que afeta os citros. Posteriormente, uma série de estudos emergiu, envolvendo consórcios de pesquisa em diversas espécies, como protozoários, plantas e animais (Salzano, F., 2018).

### **3.4 Bioinformática**

Atualmente, há várias terminologias que buscam descrever o que constitui a biotecnologia. Uma vez que sua aplicação abrange diversas áreas do conhecimento, como engenharia genética, biologia molecular, engenharias, bioquímica, entre outras, sendo, assim, considerada uma disciplina multidisciplinar (Zurron, F., 2016).

Nesse sentido, uma área importante é a bioinformática, que se dedica ao estudo abrangente de um material genético sequenciado. Assim possibilita o estudo de um genoma de uma espécie específica, no caso da genômica, bem como de seus transcritos, proteínas e metabólitos. Utilizando softwares e linguagens de programação específicas, a bioinformática busca elucidar problemas biológicos complexos, buscar por genes alvo e diversas outras tarefas (Carvalho, L. M. et al., 2019).

Dentre esses programas, existe o Prokka (Torsten, S., 2014; Afgan, E. et al, 2018) que realiza a anotação de sequências de DNA genômicos pré-montadas em formato FASTA. Ele depende de ferramentas externas de predição de características para identificar as coordenadas das características genômicas dentro dos contigs, que são sequências maiores montadas a partir de sequências menores. Além disso, o Prokka usa esse método, mas de maneira hierárquica, começando com um banco de dados menor e confiável, passando para bancos de dados de tamanho médio, mas de domínio específico, e, finalmente, para modelos selecionados de famílias de proteínas para aprimorar ainda mais a anotação genômica (Torsten, S., 2014).

Outra ferramenta para anotação funcional é o eggNOG-mapper v2 (Cantalapiedra, C. P. et al., 2021) que realiza anotações funcionais eficientes, refinando as atribuições de ontologia com base em filogenias pré-computadas. Essa abordagem resulta em previsões mais precisas em comparação com métodos baseados apenas em homologia. A atualização para a versão v2 inclui melhorias na cobertura de anotações, desempenho geral e recursos do programa.

O programa Clustal (Larkin, et al., 2007) é o mais antigo dos programas e ainda usados atualmente para realizar alinhamento. Ele fornece diversos pacotes, por exemplo, Clustal W, Clustal X, T-Coffee, MAFFT e MUSCLE. Todos os programas Clustal atuais derivam de Clustal W. O Clustal W e Clustal X eram os programas de alinhamento múltiplo mais utilizados. Eles conseguiram alinhar conjuntos de dados de tamanho médio muito rapidamente e eram fáceis de usar. Com a evolução para a versão 2.0, o Clustal aprimorou sua capacidade de realizar alinhamentos mais rápidos, especialmente para conjuntos de dados consideravelmente grandes, enquanto mantém uma maior precisão no alinhamento (Larkin, et al., 2007).

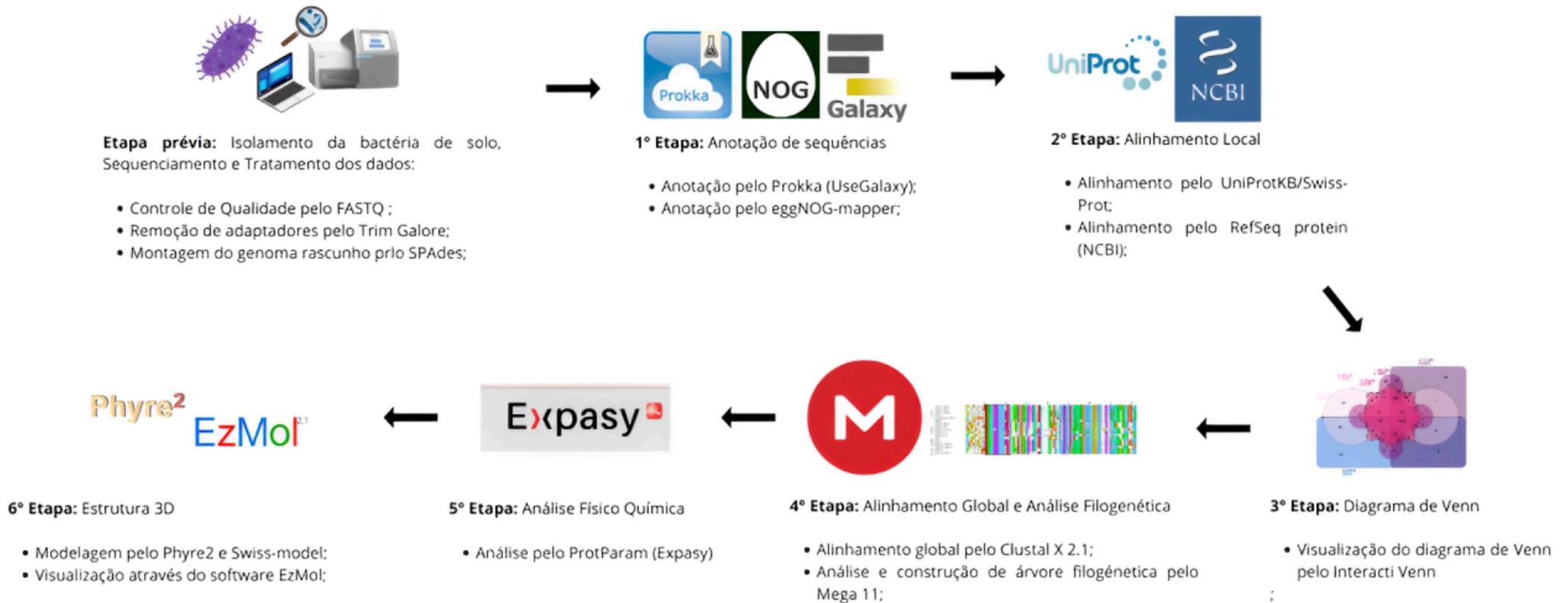
O software Molecular Evolutionary Genetics Analysis (MEGA) (Koichiro, T., et al. 2021) tem a tarefa de realizar a análise filogenética molecular. Desde seu lançamento em 1993, por mais de 25 anos, tem incorporado novos métodos, ferramentas e interfaces. Inicialmente focado em métodos baseados em distância e máxima parcimônia, o MEGA expandiu seu escopo ao longo do tempo, incluindo abordagens inovadoras. O MEGA é agora um conjunto de aplicativos versátil, atendendo aos diversos ambientes de computação utilizados por pesquisadores em evolução molecular e filogenética. A versão 11 do MEGA apresenta adições metodológicas e melhorias técnicas significativas (Koichiro, T., et al. 2021).

#### **4 MATERIAL E MÉTODOS**

Em trabalho prévio, a bactéria *Priestia megaterium* cepa E1 foi isolada e identificada a partir de amostras de solo coletadas na Fazenda Experimental da Universidade Federal da Grande Dourados (UFGD) (Prado, C. C. A. do, 2017). Posteriormente o seu material genético foi extraído e o genoma sequenciado pelo Illumina MiSeq Series no Laboratório Multiusuário Centralizado para Sequenciamento de DNA em Larga Escala e Análise de Expressão Gênica (LMSEQ) do Instituto da Unesp de Jaboticabal-SP seguindo os protocolos do fabricante. Os dados do sequenciamento foram tratados através da plataforma UseGalaxy v. 1.0.1, que incluem o controle de qualidade das sequências brutas de DNA pelo programa FASTQ,

avaliado por meio da medida de qualidade Phred Score para cada base nucleotídica sequenciada, e a montagem do genoma rascunho pela ferramenta SPAdes com parâmetros padrões. O presente estudo se inicia com a anotação funcional dessas sequências genômicas.

**Figura 5.** Fluxograma das atividades de bioinformática desenvolvidas no presente trabalho.



#### 4.1 Anotação de sequências

Utilizando a ferramenta de software Prokka (Prokaryotic genome annotation) v. 1.14.6 com parâmetros padrões, foi feita a predição e anotação funcional das sequências montadas resultantes do sequenciamento da bactéria *Priestia megaterium* cepa E1. A anotação realizada retornou um conjunto de sequências de proteínas contendo informações sobre seus respectivos genes e funções. Estas anotações foram exportadas no formato de um arquivo de planilha Excel (.xlsx). Dentro do arquivo foram feitas várias buscas com termos de “PHA” e “PHB”, as sequências encontradas com anotações correspondentes foram selecionadas (Tabela 1). Para confirmar as anotações feitas pelo Prokka, uma segunda ferramenta foi proposta. As sequências filtradas foram então executadas no eggNOG-mapper v. 2 (Cantalapiedra, C.P. et al., 2021) com parâmetros padrões e aquelas que apresentaram anotações similares em ambas as ferramentas de anotação foram selecionadas (Tabela 2). Além disso, foram rodadas todas as proteínas tidas como hipotéticas pela primeira ferramenta com intuito de encontrar possíveis anotações de PHAs na ferramenta eggNOG-mapper.

#### 4.2 Alinhamento local

As sequências anotadas pelas duas ferramentas anteriores foram submetidas a alinhamentos de similaridade local no banco de dados secundários Swiss-Prot pela plataforma UniProt v. 7.0 (Bairoch, A.; Apweiler R, 2000) e no banco de dados Reference Proteins (RefSeq\_protein) v. 259 por meio da ferramenta Basic Local Alignment Search Tool Protein (blastp) da plataforma do NCBI (National Center for Biotechnology Information) (Pruitt, K.D. et al., 2007). A sequência que apresentou maior identidade de cada busca feita no banco de dados Swiss-Prot foi selecionada e montada a Tabela 3. Além disso, a primeira sequência de cada alinhamento no banco de dados RefSeq\_protein que retornou similaridade com um valor de e-value inferior a  $1 \times 10^{-5}$  e com uma boa identidade foi selecionada e montada a Tabela 5. A sequência mais similar deste último banco de dados foi exportada em formato FASTA

Os genes e EC\_number (número da Enzyme Commission) das sequências de proteínas do genoma da bactéria *Priestia megaterium* cepa E1 anotadas pelas ferramentas Prokka e eggNOG-mapper, juntamente com as sequências de maior identidade selecionadas das buscas feitas no banco de dados Swiss-Prot, foram inseridos e comparados na Tabela 4. Além disso,

foi montada uma tabela consenso (Tabela 6) contendo as correspondências de anotações feitas pelas quatro ferramentas utilizadas até então.

### 4.3 Diagrama de Venn

As sequências do genoma da bactéria *Priestia megaterium* cepa E1 que foram utilizadas como sequências de entrada nos programas de anotação Prokka e eggNOG-mapper e alinhadas nos bancos de dados UniProt (Swiss-Prot) e RefSeq Protein foram inseridas na plataforma de análise Interacti Venn, acessível através do site <<https://www.interactivenn.net>> (Heberle, H. et al., 2015).

### 4.4 Alinhamento global e Análise filogenética

As sequências de proteínas do genoma da bactéria *Priestia megaterium* cepa E1 anotadas e alinhadas foram agrupadas em arquivos separados juntamente com a sequência mais similar exportada do banco RefSeq\_protein. Foram criados sete arquivos contendo somente sequências relacionadas. Tais arquivos foram submetidos a alinhamento múltiplo global utilizando o programa ClustalX v. 2.1 (Larkin, M.A. et al., 2007). Cada alinhamento global gerado foi utilizado para a análise filogenética através da ferramenta MEGA v. 11.0 (Koichiro, T. et al., 2007). As árvores para a pesquisa heurística foram construídas aplicando o método Neighbor-Joining a uma matriz de distâncias emparelhadas estimada usando um modelo de substituição p-distance a 1000 repetições de bootstraps. Vale ressaltar que todas as árvores foram construídas com deleção parcial das lacunas (GAPs) nos alinhamentos múltiplos globais, com exceção da árvore das proteínas fasinas (PhaI) que foi utilizado a deleção completa de GAPs.

### 4.5 Análise físico química

A análise foi feita através de software computacional, o ProtParam v. 8.0, que possibilita a predição de diversas características físicas e químicas de proteínas. As sequências de proteínas da bactéria *Priestia megaterium* cepa E1 anotadas e alinhadas foram submetidas a análise e são exibidas na Tabela 7. A ferramenta pode ser acessada através do site <<https://web.expasy.org/protparam/>> (Gasteiger, 2005).

## 4.6 Estrutura 3D

As sequências das proteínas de PhaC sintase foram modeladas usando o software Phyre v. 2.0 (Kelley, L. A et al., 2015), empregando o modelo oculto de Markov e o Swiss-model v. 2024.2 (Schwede, T. et al., 2003). Após a conclusão de cada análise, um arquivo de saída no formato PDB (Protein Data Bank) foi gerado. Em seguida, esse arquivo foi inserido no software EzMol v. 2.1 para a visualização da proteína em modelo de nuvem eletrônica e cartoon (Reynolds, C. R., 2018). Vale ressaltar que não foi realizada a modelagem das outras diferentes proteínas anotadas anteriormente devido a uma carência em artigos científicos já publicados que contenha a predição de suas estruturas 3D para elucidar a investigação.

## 5 RESULTADOS E DISCUSSÕES

### 5.1 Anotação funcional

A Tabela 1 mostra as sequências filtradas com anotações de polihidroxicanoatos (PHAs), dentre as 23145 sequências anotadas do genoma *Priestia megaterium* cepa E1 que o Prokka (Torsten, S., 2014; Afgan, E. et al., 2018) retornou. Das 25 sequências filtradas, foram sete proteínas PhaC sintases, cinco Acetyl-CoA transferases (PhaA), três Acetoacetyl-CoA reductases (PhaB) e duas Enoyl-CoA hydratase (PhaJ). Outras 8 proteínas foram tidas como hipotéticas, que são sequências previstas pela análise de sequências genômicas, mas cujas funções biológicas específicas não foram determinadas experimentalmente. Estas proteínas foram anotadas sendo PHAs pelas ferramentas posteriores, por isso foram adicionadas à Tabela 1 para evidenciar que não foram anotadas por uma das ferramentas utilizadas no presente estudo. De modo geral, os tamanhos das 25 sequências variam de 246 a 1683 aminoácidos. Além disso, algumas apresentaram as classificações enzimáticas: 2.3.1.- (PHAs sintases), 2.3.1.9 (Acetyl-CoA transferases), 1.1.1.36 (Acetoacetyl-CoA reductases) e 4.2.1.119 (Enoyl-CoA hydratase). Estas sequências pertencem aos grupos ortólogos de proteínas COG3243 (PhaC), COG0183 (PhaA) e COG2030 (PhaJ), ou seja, proteínas que são homólogas entre diferentes organismos e desempenham funções similares (Galperin, M. Y. et al., 2019). Com exceção das proteínas Acetoacetyl-CoA reductases que não foram fornecidas a ortologia.

A Tabela 2 mostra a segunda ferramenta utilizada para confirmar as anotações feitas pelo Prokka. O eggNOG mapper forneceu anotações das oito sequências que tinham sido apresentadas como hipotéticas anteriormente. Sendo elas, uma proteína reguladora (PhaR), cinco proteínas fasinas (PhaI) e duas PHB despolimerases (PhaZ). Além disso, as 17 sequências inicialmente anotadas pelo Prokka foram confirmadas pela segunda ferramenta.

Pela anotação feita pelo eggNOG-mapper (Pearson, W. R. et al., 2013), as sequências de proteínas do genoma da bactéria *Priestia megaterium* cepa E1 apresentaram boas pontuações de e-value e score, que indicam respectivamente o número de ocorrências (alinhamentos) esperadas com pontuação igual ou melhor do que obtido usando uma base de dados de mesmo tamanho e comprimento da sequência *query* similar ao utilizado e a qualidade da correspondência encontrada (Pearson, W. R. et al., 2013). Quanto menor o e-value, maiores são as chances de que a ocorrência não foi encontrada ao acaso. As sequências EAMNKLGA\_07463, EAMNKLGA\_07465 e EAMNKLGA\_11085 apresentaram as melhores pontuações com 0.0 de e-value e score acima de 885.0. Por outro lado, a sequência EAMNKLGA\_13677 e EAMNKLGA\_13678 foram as que apresentaram o pior valor de e-value e score ( $7.74e-22$  e 94.4;  $7.32e-33$  e 124.0, respectivamente), possivelmente sejam sequências resultante de um sequenciamento incompleto uma vez que foram as sequências de menor tamanho em pares de base (246 e 261 aminoácidos, respectivamente).

As cinco sequências de Acetyl-CoA transferases (PhaA) apresentaram a mesma classificação enzimática (E.C\_number) em ambas ferramentas de anotação funcional. Pelo eggNOG-mapper, as proteínas Acetoacetyl-CoA reductases (PhaB) apresentaram a classificação 1.1.1.36, as quais não tinham sido classificadas pelo Prokka. Além disso, o eggNOG-mapper fornece a classificação das duas proteínas despolimerases (EAMNKLGA\_07464 e EAMNKLGA\_11085), 3.1.1.76 e 3.1.1.75, respectivamente. As quais tinham sido hipotéticas pelo Prokka. As cinco proteínas fasinas (PhaI) e a proteína reguladora (PhaR) não apresentaram classificação enzimática (EC\_number) em nenhuma das duas ferramentas de anotação funcional.

**Tabela 1.** As 25 sequências filtradas envolvidas ao metabolismo de PHA a partir da anotação feita pela ferramenta Prokka.

Prokka					
Locus_tag	Length_pb	Descrição	Genes	E.C_number	COG
EAMNKLGA_05793	1185	Acetyl-CoA transferase	<i>phaA_1</i>	2.3.1.9	-
EAMNKLGA_05917	1086	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_1</i>	2.3.1.-	COG3243
EAMNKLGA_07463	1683	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_2</i>	2.3.1.-	COG3243
EAMNKLGA_07465	1680	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_3</i>	2.3.1.-	COG3243
EAMNKLGA_10818	1179	Acetyl-CoA transferase	<i>phaA_2</i>	2.3.1.9	COG0183
EAMNKLGA_13058	1644	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_4</i>	2.3.1.-	COG3243
EAMNKLGA_13061	747	Acetoacetyl-CoA reductase	<i>phaB_1</i>	1.1.1.36	-
EAMNKLGA_13677	246	Acetyl-CoA transferase	<i>phaA_3</i>	2.3.1.9	COG0183
EAMNKLGA_13678	261	Acetyl-CoA transferase	<i>phaA_4</i>	2.3.1.9	COG0183
EAMNKLGA_16837	741	Acetoacetyl-CoA reductase	<i>phaB_2</i>	1.1.1.36	-
EAMNKLGA_17817	471	(R)-specific enoyl-CoA hydratase	<i>phaJ_1</i>	4.2.1.119	COG2030
EAMNKLGA_17957	741	Acetoacetyl-CoA reductase	<i>phaB_3</i>	1.1.1.36	-
EAMNKLGA_20212	408	(R)-specific enoyl-CoA hydratase	<i>phaJ_2</i>	4.2.1.119	COG2030
EAMNKLGA_20844	1182	Acetyl-CoA transferase	<i>phaA_5</i>	2.3.1.9	COG0183
EAMNKLGA_22690	1152	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_5</i>	2.3.1.-	COG3243
EAMNKLGA_22716	1065	Poly(3-hydroxyalkanoate) polymerase subunit PhaE	<i>phaE</i>	-	-
EAMNKLGA_22717	1068	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>phaC_6</i>	2.3.1.-	COG3243
EAMNKLGA_17958	552	hypothetical protein	-	-	-
EAMNKLGA_05951	456	hypothetical protein	-	-	-
EAMNKLGA_07460	411	hypothetical protein	-	-	-
EAMNKLGA_13384	348	hypothetical protein	-	-	-
EAMNKLGA_16952	492	hypothetical protein	-	-	-
EAMNKLGA_07461	972	hypothetical protein	-	-	-
EAMNKLGA_07464	1005	hypothetical protein	-	-	-
EAMNKLGA_11085	885	hypothetical protein	-	-	-

Locus\_tag - Localização da sequência no genoma, Length\_pb – Tamanho da sequência, E.C\_number – Classificação enzimática, COG – grupos ortólogos.

**Tabela 2.** Anotação das 25 sequências envolvidas no metabolismo de PHA pela ferramenta eggNOG-mapper.

eggNOG-mapper					
Query	Descrição	Genes	E-value	Score	E.C_number
EAMNKLGA_05793	Belongs to the thiolase family	<i>phbA</i>	1.77e-280	767.0	2.3.1.9
EAMNKLGA_05917	alpha/beta hydrolase fold	<i>phaC</i>	4.58e-248	682.0	-
EAMNKLGA_07463	Poly-beta-hydroxybutyrate polymerase (PhaC) N-terminus	<i>phaC2</i>	0.0	1100.0	-
EAMNKLGA_07465	poly(R)-hydroxyalkanoic acid synthase	<i>phaC1</i>	0.0	1117.0	-
EAMNKLGA_10818	Belongs to the thiolase family	<i>atoB</i>	2.07e-262	721.0	2.3.1.9
EAMNKLGA_13058	Poly-beta-hydroxybutyrate polymerase	<i>phaC</i>	7e-268	749.0	-
EAMNKLGA_13061	Acetoacetyl-CoA reductase	<i>phaB</i>	2.4e-141	403.0	1.1.1.36
EAMNKLGA_13677	Belongs to the thiolase family	<i>phbA</i>	7.74e-22	94.4	2.3.1.9
EAMNKLGA_13678	Belongs to the thiolase family	<i>phaA</i>	7.32e-33	124.0	2.3.1.9
EAMNKLGA_16837	Catalyzes from in polyhydroxyalkanoate synthesis	<i>phbB</i>	5.41e-171	478.0	1.1.1.36
EAMNKLGA_17817	N-terminal half of MaoC dehydratase	<i>phaJ</i>	6.1e-101	293.0	-
EAMNKLGA_17957	Belongs to the short-chain dehydrogenases reductases (SDR) family	<i>phbB</i>	1.39e-173	484.0	1.1.1.36
EAMNKLGA_20212	Acyl dehydratase	-	3.6e-92	269.0	-
EAMNKLGA_20844	Coenzyme A transferase	<i>atoB</i>	5.6e-272	745.0	2.3.1.9
EAMNKLGA_22690	Poly-beta-hydroxybutyrate polymerase (PhaC) N-terminus	<i>phbC</i>	8.12e-285	777.0	-
EAMNKLGA_22716	Poly(R)-hydroxyalkanoic acid synthase subunit (PHA_synth_III_E)	-	1.68e-219	609.0	-
EAMNKLGA_22717	acid) synthase	<i>phbC</i>	2.4e-256	702.0	-
EAMNKLGA_17958	synthesis repressor, PhaR, PHB_acc,PHB_acc_N	<i>phaR</i>	1.2e-106	310.0	-
EAMNKLGA_05951	Poly(hydroxyalkanoate) granule associated protein (phasin)	<i>phaI</i>	9.11e-96	280.0	-
EAMNKLGA_07460	Poly(hydroxyalkanoate) granule associated protein (phasin)	<i>phaI</i>	1.16e-84	250.0	-
EAMNKLGA_13384	Phasin protein	-	4.05e-35	124.0	-
EAMNKLGA_16952	Poly(hydroxyalkanoate) granule associated protein (phasin)	<i>phaI</i>	4.76e-106	306.0	-
EAMNKLGA_07461	Granule-associated protein	<i>phaF</i>	1.52e-71	230.0	-
EAMNKLGA_07464	poly(3-hydroxyalkanoate) depolymerase	<i>phaZ</i>	5.52e-240	659.0	3.1.1.76
EAMNKLGA_11085	de-polymerase	<i>phaZ</i>	0.0	885.0	3.1.1.75

Query - Sequências de entrada, E-value - Pontuação de correspondência ao acaso, Score – Grau de semelhança entre duas sequências.

O eggNOG mapper fornece mais informações sobre os genes que codificam as proteínas presentes no genoma da bactéria *Priestia megaterium* cepa E1, o que ajuda a compreender a disposição destas regiões codificantes do genoma. As sequências de PhaC, EAMNKLGA\_05917 e EAMNKLGA\_22690, codificadas pelos genes *phaC*, aparecem em locus distantes e isolados de outras quaisquer sequências relacionadas com polihidroxicanoato. Enquanto a PhaC (EAMNKLGA\_13058) aparece em locus próximos com os genes *phaA* (EAMNKLGA\_13060) e *phaB* (EAMNKLGA\_13061) - possivelmente estão inseridos no mesmo operon (*phaABC*), o qual já é muito estudado. Este operon é responsável por codificar as principais enzimas da rota metabólica I de síntese de PHA (Zhou, W. et al., 2023).

Outras duas sequências de PhaC são codificadas pelos genes *phaC1* (EAMNKLGA\_07465) e *phaC2* (EAMNKLGA\_07463), os quais regulam a expressão de PhaC sintase de classe II. Esta classe é composta por somente uma subunidade e atua polimerizando coa tioesters em monômeros MCLs (PHAs de cadeias médias) (Song, L. et al., 2023). Geralmente esses dois genes se encontram em *cluster* com outros genes (*phaIFDC2ZC1*), entretanto no genoma em questão, os genes *phaC1* e *phaC2* aparecem em locus consecutivos com os genes *phaI* (EAMNKLGA\_07460), *phaF* (EAMNKLGA\_07461) e *phaZ* (EAMNKLGA\_07464) – o gene *phaD* não foi constatado, neste locus está presente um regulador transcricional TetR/AcrR (EAMNKLGA\_07462), o qual não apresenta uma relação direta com PHA e não foi inserido ao presente trabalho (Deng, W. et al., 2013).

Outras duas sequências de PhaC sintase restantes, são codificadas pelos genes *phaE* (EAMNKLGA\_22716) e *phaC* (EAMNKLGA\_22717), que apareceram em locus pareados, possivelmente se trata do operon *phaEC*. O primeiro gene codifica a subunidade PhaE, enquanto o segundo codifica a subunidade PhaC. Estas duas subunidades atuam na formação da PhaC sintase de classe III, que catalisam a polimerização de HAcCoAs em PHAs de cadeias curtas (SCLs) (Abdallah, B. M. et al., 2023).

Vale ressaltar que, embora o gênero *Bacillus* seja o principal representante das PhaC sintases de classe IV, na cepa E1 da espécie *Bacillus megaterium*, que foi renomeada como *Priestia megaterium* (Liu, J. M et al., 2023), não foi encontrada a região codificadora da subunidade PhaR, a qual é essencial para formar a PhaC sintase desta classe (Kumar, R. et al., 2023). No entanto, foi encontrada uma proteína PhaR, codificada pelo gene *phaR*, que atua reprimindo a síntese de PHA, geralmente associada à superfície do grânulo de PHA (Aghaali, Z. et al., 2023).

A conversão da molécula de acetyl-CoA em 3-hidroxi-butiril-CoA é mediada pelas enzimas PhaA e PhaB, uma das principais rotas de síntese de PHA. As Acetyl-CoA transferases (PhaA) se ligam às moléculas de acetyl-CoA como cofatores para produzir acetoacetyl-CoA. Esta última molécula é substrato para a Acetoacetyl-CoA redutase (PhaB), que converte acetoacetyl-CoA em monômero (R)-3-hydroxybutyryl-CoA, o qual é catalisado por uma PhaC sintase (Samadhiya, A. et al. 2022).

Além disso, o interesse biotecnológico nesse tipo de biopolímero (polihidroxialcanoato), não se limita somente a sua produção a partir de microrganismos. A biodegradação é outra característica importante, visto que essa capacidade é crucial em contextos ambientais, como o acúmulo de resíduos plásticos por vários anos. Geralmente, os microrganismos bioacumuladores de PHA carregam tanto genes relacionados no processo de biossíntese, como também genes relacionados à biodegradação (Kumar, R. et al., 2023).

No genoma em questão, foram anotadas duas proteínas despolimerases (EAMNKLGA\_07464 e EAMNKLGA\_11085) codificadas pelos genes *phaZ*. Embora o Prokka não tenha fornecido a anotação delas, o eggNOG-mapper forneceu anotação dela, com classificação enzimática 3.1.1.76 e 3.1.1.75, respectivamente.

## 5.2 BLAST

A Tabela 3 exibe as 25 sequências de maior identidade retornadas pelo BLAST no banco de dados UniProtKB/Swiss-Prot, que contém entradas anotadas manualmente. As sequências apresentaram correspondência com as anotações feitas pelas ferramentas anteriores, com variações nas nomenclaturas empregadas aos genes codificantes. No entanto, as proteínas expressas foram as mesmas. De modo geral, as sequências obtiveram valores de identidades superiores a 60%, com exceção de seis sequências com identidade inferior a 50%.

A Tabela 4 traz um comparativo das anotações dos genes codificantes e das classificações de E.C number das 25 sequências anotadas pelo Prokka e eggNOG-mapper, e alinhadas pelo banco de dados UniProtKB/Swiss-Prot. Em todas as ferramentas apresentaram variação na nomenclatura dos genes codificantes. Além disso, dentre as 25 sequências, 17 apresentam os mesmos E.C number em pelo menos duas ferramentas. A sequência EAMNKLGA\_22716 anotada como sendo a subunidade PhaE não apresentou E.C number em nenhuma das ferramentas, a sua função biológica ainda permanece desconhecida, como sugere Neoh, Z. S. e colaboradores (2022). No entanto deve estar presente, obrigatoriamente, na composição das PhaC sintases de classe III para que esta seja funcional. Outras sete

sequências não apresentaram o E.C number em pelo menos duas ferramentas. Algumas sequências de proteínas não foram revisadas pelo banco de dados Swiss-Prot e foram fornecidas pelo banco de dados TrEBML (Translation of the EMBL database). Este consiste de entradas traduzidas de sequências de nucleotídeos do banco de dados EMBL, que não estão no banco de dados Swiss-Prot (Junker, V. et al., 2000). O Swiss-Prot é um banco de dados de sequências de proteínas com curadoria, o que significa que contém somente sequências revisadas e estudadas em bancadas por cientistas (Bairoch, A.; Apweiler R, 2000).

**Tabela 3.** Alinhamento local das 25 sequências de proteínas realizado no banco de dados Swiss-Prot.

Query	UniProt (Swiss-Prot)				
	Taman ho	Genes	Proteína	Organismo	Identidade
EAMNKLGA_05793	393	<i>phaA</i>	Acetyl-CoA acetyltransferase	<i>Rhizobium meliloti (strain 1021)</i>	78.6%
EAMNKLGA_05917	492	<i>phaC</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>Haloferax mediterranei</i>	27.7%
EAMNKLGA_07463	560	<i>phaC2</i>	Poly(3-hydroxyalkanoate) polymerase 2	<i>Pseudomonas oleovorans</i>	70.7%
EAMNKLGA_07465	559	<i>phaC1</i>	Poly(3-hydroxyalkanoate) polymerase 1	<i>Pseudomonas oleovorans</i>	82.1%
EAMNKLGA_10818	393	<i>atoB</i>	Acetyl-CoA acetyltransferase	<i>Pseudomonas aeruginosa</i>	83.7%
EAMNKLGA_13058	583	<i>phaC</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>Azorhizobium caulinodans</i>	39.5%
EAMNKLGA_13061	246	<i>phaB</i>	Acetoacetyl-CoA reductase	<i>Cupriavidus necator</i>	56.6%
EAMNKLGA_13677	393	<i>atoB</i>	Acetyl-CoA acetyltransferase	<i>Pseudomonas aeruginosa</i>	60%
EAMNKLGA_13678	393	<i>atoB</i>	Acetyl-CoA acetyltransferase	<i>Pseudomonas aeruginosa</i>	69.8%
EAMNKLGA_16837	246	<i>phaB</i>	Acetoacetyl-CoA reductase	<i>Cupriavidus necator</i>	64.2%
EAMNKLGA_17817	144	<i>phaJ</i>	(R)-specific enoyl-CoA hydratase	<i>Rhodospirillum rubrum</i>	45%
EAMNKLGA_17957	246	<i>phaB</i>	Acetoacetyl-CoA reductase	<i>Allochromatium vinosum</i>	53.7%
EAMNKLGA_20212	2,086	<i>FAS1</i>	Fatty acid synthase subunit beta	<i>Yarrowia lipolytica</i>	33.8%
EAMNKLGA_20844	393	<i>atoB</i>	Acetyl-CoA acetyltransferase	<i>Pseudomonas aeruginosa</i>	78.6%
EAMNKLGA_22690	492	<i>phaC</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>Haloferax mediterranei</i>	29.4%
EAMNKLGA_22716	357	<i>phaE</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaE	<i>Allochromatium vinosum</i>	26.3%
EAMNKLGA_22717	355	<i>phaC</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	<i>Stenotrophomonas maltophilia</i>	99.2%
EAMNKLGA_17958	175	<i>Smlt3326</i>	Poly-hydroxy-butyrate synthesis protein	<i>Stenotrophomonas maltophilia</i>	95.6%
EAMNKLGA_05951	151	<i>BDGL_00100</i>	Poly(Hydroxyalkanoate) granule associated protein	<i>Acinetobacter calcoaceticus</i>	93.4%
EAMNKLGA_07460	136	<i>UB43_16185</i>	Poly(3-hydroxyalkanoate) granule-associated protein PhaI	<i>Pseudomonas sp. 21</i>	94.1%
EAMNKLGA_13384	115	<i>Smlt0600</i>	Phasin domain-containing protein	<i>Stenotrophomonas maltophilia</i>	91.3%
EAMNKLGA_16952	163	<i>phaF</i>	Poly(Hydroxyalkanoate) granule-associated protein	<i>Stenotrophomonas maltophilia</i>	96.9%
EAMNKLGA_07461	352	<i>algP</i>	Transcriptional regulatory protein AlgP	<i>Pseudomonas aeruginosa</i>	52.6%
EAMNKLGA_07464	291	<i>UB43_16165</i>	Poly(3-hydroxyalkanoate) depolymerase	<i>Pseudomonas sp. 21</i>	98.6%
EAMNKLGA_11085	435	<i>Smlt1486</i>	PHB depolymerase	<i>Stenotrophomonas maltophilia</i>	99.1%

**Tabela 4.** Comparação das descrições feitas pelas ferramentas de anotação Prokka e eggNOG-mapper e pelos alinhamentos no Swiss-Prot.

Query	Prokka		eggNOG-mapper		Uniprot (Swiss-Prot)	
	Genes	E.C number	Genes	E.C number	Genes	E. C number
EAMNKLGA_05793	<i>phaA_1</i>	2.3.1.9	<i>phbA</i>	2.3.1.9	<i>phaA</i>	2.3.1.9
EAMNKLGA_05917	<i>phaC_1</i>	2.3.1.-	<i>phaC</i>	-	<i>phaC</i>	2.3.1.-
EAMNKLGA_07463	<i>phaC_2</i>	2.3.1.-	<i>phaC2</i>	-	<i>phaC2</i>	2.3.1.-
EAMNKLGA_07465	<i>phaC_3</i>	2.3.1.-	<i>phaC1</i>	-	<i>phaC1</i>	2.3.1.-
EAMNKLGA_10818	<i>phaA_2</i>	2.3.1.9	<i>atoB</i>	2.3.1.9	<i>atoB</i>	2.3.1.9
EAMNKLGA_13058	<i>phaC_4</i>	2.3.1.-	<i>phaC</i>	-	<i>phaC</i>	2.3.1.-
EAMNKLGA_13061	<i>phaB_1</i>	1.1.1.36	<i>phaB</i>	1.1.1.36	<i>phaB</i>	1.1.1.36
EAMNKLGA_13677	<i>phaA_3</i>	2.3.1.9	<i>phbA</i>	2.3.1.9	<i>atoB</i>	2.3.1.9
EAMNKLGA_13678	<i>phaA_4</i>	2.3.1.9	<i>phaA</i>	2.3.1.9	<i>atoB</i>	2.3.1.9
EAMNKLGA_16837	<i>phaB_2</i>	1.1.1.36	<i>phbB</i>	1.1.1.36	<i>phaB</i>	1.1.1.36
EAMNKLGA_17817	<i>phaJ_1</i>	4.2.1.119	<i>phaJ</i>	-	<i>phaJ</i>	4.2.1.119
EAMNKLGA_17957	<i>phaB_3</i>	1.1.1.36	<i>phbB</i>	1.1.1.36	<i>phaB</i>	1.1.1.36
EAMNKLGA_20212	<i>phaJ_2</i>	4.2.1.119	-	-	<i>FAS1</i>	4.2.1.119
EAMNKLGA_20844	<i>phaA_5</i>	2.3.1.9	<i>atoB</i>	2.3.1.9	<i>atoB</i>	2.3.1.9
EAMNKLGA_22690	<i>phaC_5</i>	2.3.1.-	<i>phbC</i>	-	<i>phaC</i>	2.3.1.-
EAMNKLGA_22716	<i>phaE</i>	-	-	-	<i>phaE</i>	-*
EAMNKLGA_22717	<i>phaC_6</i>	2.3.1.-	<i>phbC</i>	-	<i>phaC</i>	2.3.1.-
EAMNKLGA_17958	-	-	<i>phaR</i>	-	<i>Smlt3326</i>	-*
EAMNKLGA_05951	-	-	<i>phaI</i>	-	<i>BDGL_001005</i>	-*
EAMNKLGA_07460	-	-	<i>phaI</i>	-	<i>UB43_16185</i>	-*
EAMNKLGA_13384	-	-	-	-	<i>Smlt0600</i>	-*
EAMNKLGA_16952	-	-	<i>phaI</i>	-	<i>phaF</i>	-*
EAMNKLGA_07461	-	-	<i>phaF</i>	-	<i>algP</i>	-*
EAMNKLGA_07464	-	-	<i>phaZ</i>	3.1.1.76	<i>phaZ</i>	3.1.1.-
EAMNKLGA_11085	-	-	<i>phaZ</i>	3.1.1.75	<i>UB43_16165</i>	-*

\* - Sequência não revisada pelo SwissProt fornecida pelo TrEMBL.

A Tabela 5 apresenta o BLAST das 25 sequências no banco de dados Reference Sequence Protein do NCBI (National Center for Biotechnology Information), o qual é secundário e todas sequências depositadas passam por curadoria (Pruitt, K. D. et al., 2007). De modo geral, todas as sequências apresentaram boas identidades com as mais similares do banco de dados e a porcentagem de positivos superior a 83%. Além disso, as proteínas que retornaram melhores similaridades confirmam as anotações e o alinhamento feito no banco de dados Swiss-Prot. Embora o banco de dados RefSeq Protein não tenha fornecido o nome de todas as proteínas, fornece pelo menos a família as quais as proteínas estão inseridas. Como é o caso das PhaC sintase que fazem parte da família Alpha/beta fold hydrolase, que se trata de um conjunto de enzimas com essa conformação estrutural (Alaux, E. et al., 2022). Como também a superfamília MaoC family dehydratase que é composta por proteínas como a PhaJ (Enoyl-CoA hydratase), a qual atua na rota metabólica  $\beta$ -oxidação fornecendo monômeros para a biossíntese de PHA (Wan, J. H., 2023). Em relação às proteínas fasinas (PhaI) também não foram fornecidos os nomes dos genes, foram descritas como pertencentes à família de proteínas fasinas (Choi, S. Y., 2020).

**Tabela 5.** Alinhamento local das 25 sequências de proteínas realizado no banco de dados Reference Sequence Protein.

Query	RefSeq Protein					ID_sequence
	Score	Positives	E-value	Identidade	Proteínas	
EAMNKLGA_05793	2053	100%	0.0	99%	Acetyl-CoA C-acetyltransferase	WP_094514208.1
EAMNKLGA_05917	1873	98%	0.0	98%	Alpha/beta fold hydrolase	WP_198071066.1
EAMNKLGA_07463	2940	98%	0.0	97%	Class II poly(R)-hydroxyalkanoic acid synthase	WP_138525372.1
EAMNKLGA_07465	2880	97%	0.0	94%	Class II poly(R)-hydroxyalkanoic acid synthase	WP_236172704.1
EAMNKLGA_10818	1973	96%	0.0	100%	Acetyl-CoA C-acetyltransferase	WP_138524486.1
EAMNKLGA_13058	2309	78%	0.0	85%	Alpha/beta fold hydrolase	WP_035224716.1
EAMNKLGA_13061	1188	94%	3e-164	88%	Acetoacetyl-CoA reductase	WP_035224625.1
EAMNKLGA_13677	267	86%	3e-26	80%	Acetyl-CoA C-acetyltransferase	WP_058490479.1
EAMNKLGA_13678	305	83%	5e-33	74%	Acetyl-CoA C-acyltransferase, partial	WP_139205583.1
EAMNKLGA_16837	1290	100%	8e-180	100%	Acetoacetyl-CoA reductase	WP_005414435.1
EAMNKLGA_17817	780	99%	4e-105	96%	MaoC family dehydratase	WP_138525083.1
EAMNKLGA_17957	1279	98%	4e-178	99%	Acetoacetyl-CoA reductase	WP_005410389.1
EAMNKLGA_20212	626	92%	3e-82	85%	MaoC family dehydratase	WP_238474373.1
EAMNKLGA_20844	1993	99%	0.0	98%	Acetyl-CoA C-acetyltransferase	WP_169938770.1
EAMNKLGA_22690	2031	99%	0.0	99%	Alpha/beta fold hydrolase	WP_198072391.1
EAMNKLGA_22716	1636	95%	0.0	91%	Class III poly(R)-hydroxyalkanoic subunit PhaE	WP_150454071.1
EAMNKLGA_22717	1836	99%	0.0	99%	Class III poly(R)-hydroxyalkanoic subunit PhaC	WP_005410089.1
EAMNKLGA_17958	900	97%	1e-122	97%	Polyhydroxyalkanoate synthesis repressor PhaR	WP_032956189.1
EAMNKLGA_05951	763	100%	1e-102	100%	Phasin family protein	WP_004701909.1
EAMNKLGA_07460	674	99%	2e-89	98%	Phasin family protein	WP_138525375.1
EAMNKLGA_13384	556	98%	3e-72	97%	Phasin family protein	WP_010484090.1
EAMNKLGA_16952	824	99%	2e-111	98%	Phasin family protein	WP_032961659.1
EAMNKLGA_07461	733	98%	9e-93	97%	Phasin family protein	WP_236201306.1
EAMNKLGA_07464	1711	98%	0.0	97%	Poly(3-hydroxyalkanoate) depolymerase	WP_226045854.1
EAMNKLGA_11085	2315	99%	0.0	99%	Polyhydroxyalkanoate depolymerase	WP_010486357.1

**Tabela 6.** Anotações correspondentes entre as ferramentas de anotação funcional, Prokka e eggNOG-mapper, e os bancos de dados secundários, Swiss-prot e RefSeq Protein.

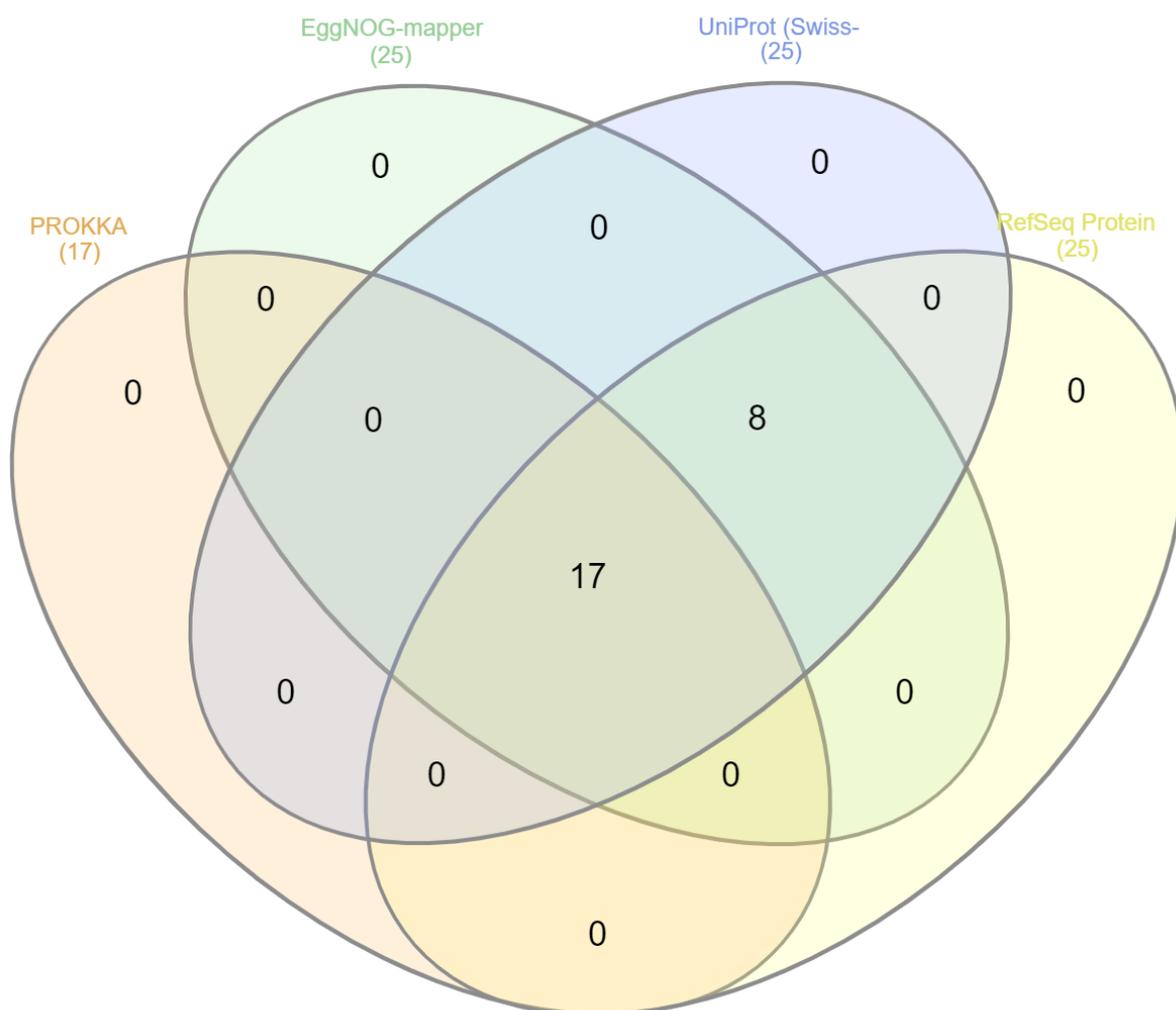
Genes	Cópias	Proteínas	Prokka	eggNOG-mapper	Swiss-prot	RefSeq Protein
<i>phaA</i>	5	Acetyl-CoA transferase	•	•	•	•
<i>phaB</i>	3	Acetoacetyl-CoA reductase	•	•	•	•
<i>phaC</i>	7	Poly(3-hydroxyalkanoate) polymerase	•	•	•	•
<i>phaJ</i>	2	Enoyl-CoA hydratase	-#	•	•	•
<i>phaI</i>	5	Proteínas fasinias	•	•	•	•
<i>phaR</i>	1	Proteína reguladora	-#	•	•	•
<i>phaZ</i>	2	Poly(3-hydroxyalkanoate) depolymerase	-#	•	•	•

- - Sequência que obteve anotação correspondente entre todas as ferramentas; -# - Sequência que não foi anotada pela ferramenta.

### 5.3 Diagrama de Venn

O programa InteractiveVenn (Heberle, H. et al.,2015) permite visualizar a representação gráfica que ilustra as relações entre diferentes conjuntos de itens através do diagrama de Venn. Nesse sentido, o programa possibilita observar as 25 sequências do genoma da bactéria *Priestia megaterium* cepa E1 que apresentaram resultados em comum entre os programas e bancos de dados analisados. A Figura 6 exibe o diagrama de Venn representando as sequências de entrada para as quatro ferramentas utilizadas no processo de anotação.

**Figura 6.** Representação gráfica da interação das sequências do genoma da bactéria *Priestia megaterium* cepa E1 nos programas Prokka, eggNOG-mapper, UniProt (Swiss-Prot) e Reference Sequence (RefSeq) Protein.



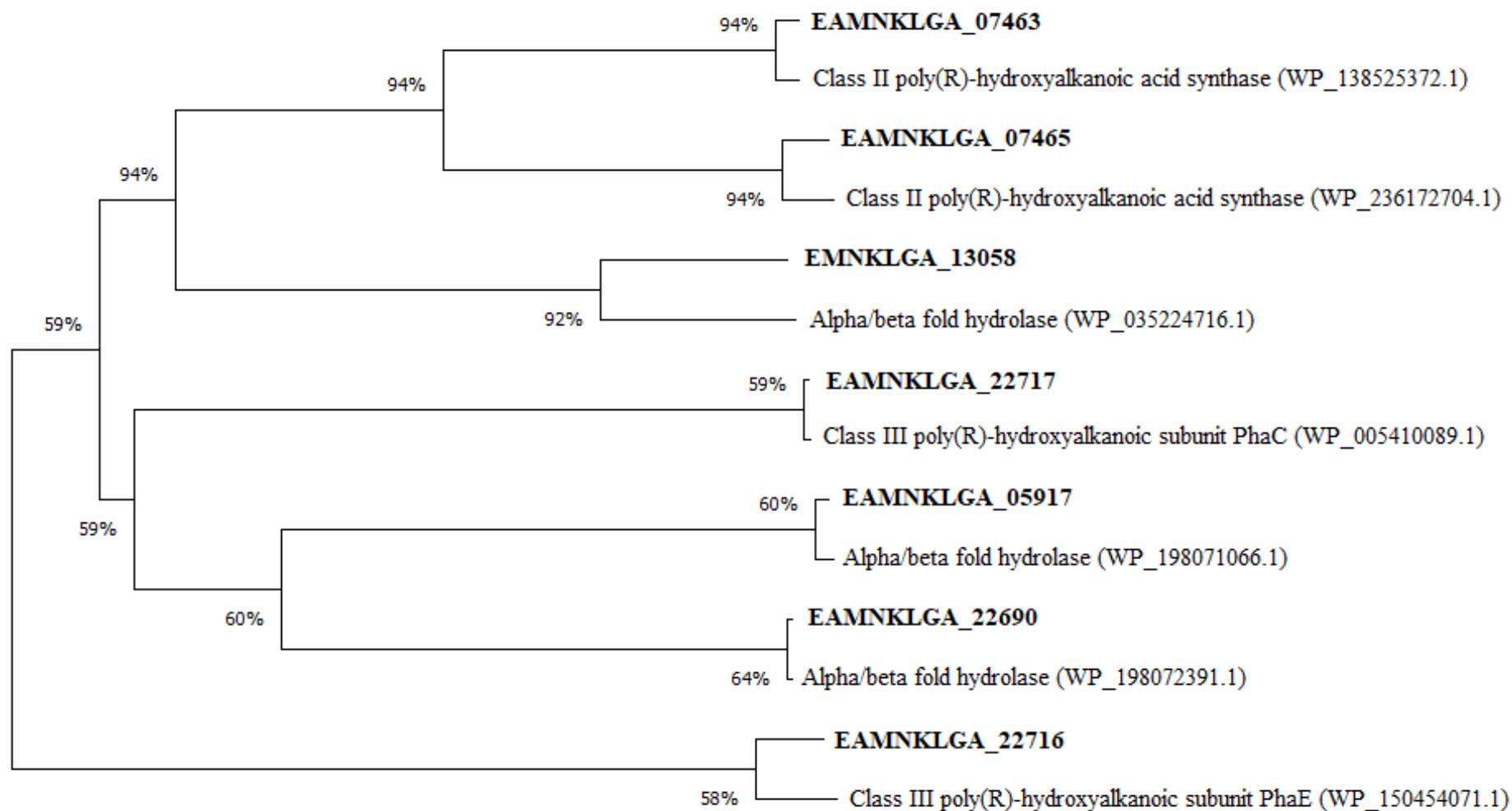
Das 25 sequências analisadas pelas ferramentas, somente 17 sequências apresentaram resultados em comuns entre as quatro ferramentas de análise. Outras oito sequências foram exclusivas somente das ferramentas eggNOG-mapper, Swiss-Prot (UniProt) e RefSeq protein (NCBI). O Prokka não apresentou anotações destas oito sequências e foram identificadas como sequências hipotéticas, que são sequências previstas pela análise de sequências genômicas, mas foram depositadas no banco de dados do programa cujas funções biológicas não foram determinadas experimentalmente. Essas oito sequências foram anotadas pelas outras três ferramentas sendo uma proteína reguladora (PhaR), cinco proteínas fasinases (PhaI) e duas proteínas despolimerases (PhaZ).

#### **5.4 Análise filogenética**

A Figura 7 mostra a árvore construída a partir das sete sequências de PhaC sintases agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu três grupos principais. O primeiro grupo contendo seis sequências, sendo dois locus alinhados com duas sequências de PhaC sintase de classe II, outro um locus alinhado com uma sequência de PhaC sintase, com bootstrap variando de 92% à 94% indicando homologia. O segundo grupo contendo seis sequências, sendo dois locus alinhados com duas sequências de PhaC sintase, outro um locus alinhado com uma sequência de PhaC sintase de classe III, com bootstrap variando de 59% a 64%, apoiado no primeiro grupo com 94% de bootstrap. O terceiro grupo contendo 2 sequências, 1 locus alinhado com 1 sequência de PhaE sintase de classe III com bootstrap de 58% e apoiado nos outros grupos com 59% de bootstrap.

Vale ressaltar que algumas sequências retiradas do banco de dados Reference Sequence Protein (NCBI) foram descritas como “alpha/beta fold hydrolase”. O banco agrupou estas sequências como pertencentes a família de proteínas com dobra de alfa/beta hidrolase, a qual faz parte as PhaC sintases que, de modo geral, têm essa conformação estrutural (Neoh, Z. S. et al., 2022).

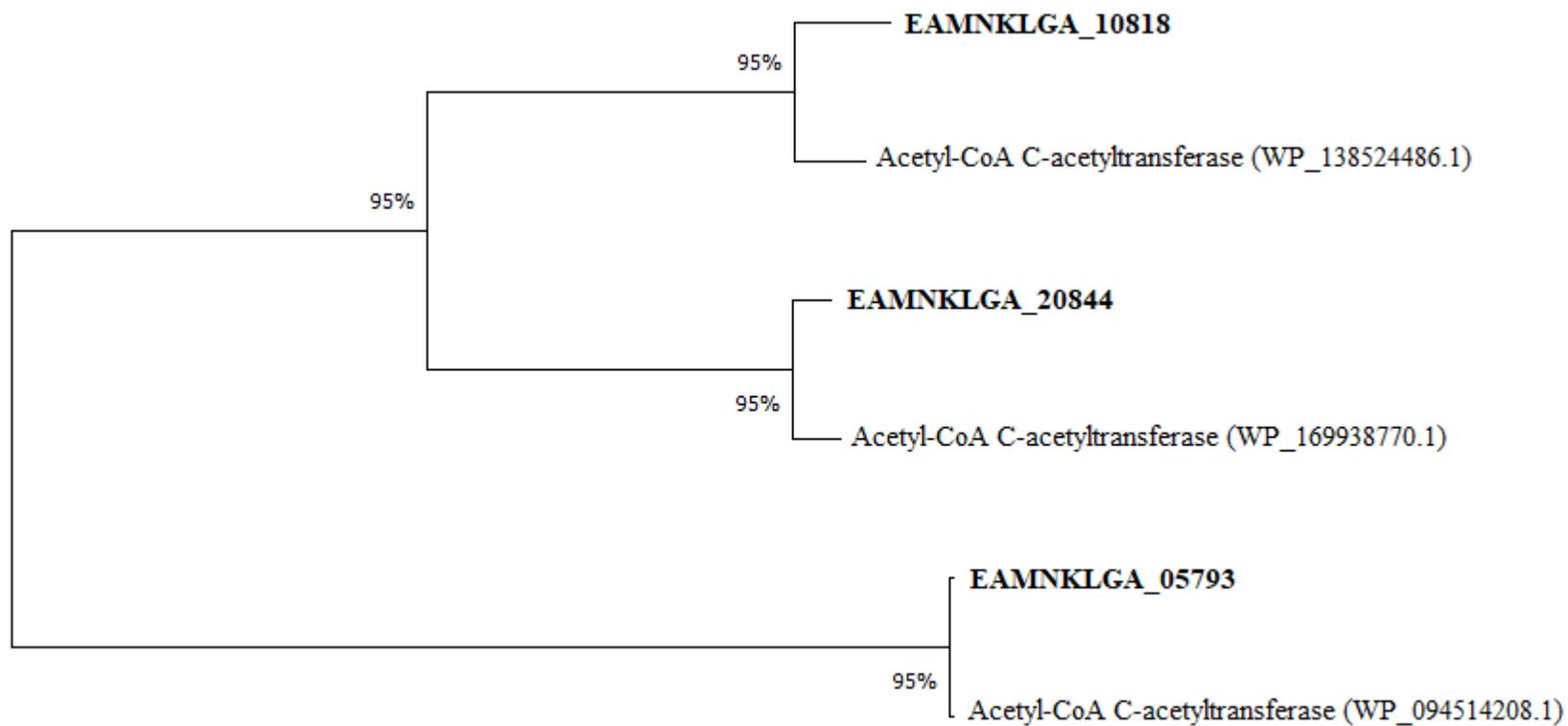
**Figura 7.** Árvore filogenética construída a partir das sete sequências de PhaC sintase presente no genoma da bactéria *Priestia megaterium* cepa E1 alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.



Ao examinar a árvore obtida das PhaC sintases, nota-se a presença de três grupos distintos. O primeiro grupo exibe sequências de PhaC sintase de classe II, enquanto o segundo grupo contém PhaC de classe III. O terceiro grupo, por sua vez, inclui sequências de PhaE sintase de classe III. Essa diversidade sugere uma relação evolutiva entre sequências presentes no genoma da bactéria *Priestia megaterium* cepa E1 com PhaC sintase de pelo menos duas classes. No caso da PhaC de classe II, ela é composta por somente uma subunidade e regulada por duas regiões gênicas (*phaC1* e *phaC2*). Diferente da PhaC de classe III que é composta por duas subunidades (PhaC e PhaE); uma unidade catalítica codificada pelo gene *phaC* e, outra unidade, codificada pelo gene *phaE*, cuja função permanece desconhecida (Neoh, S. Z. et, al, 2022).

A Figura 8 mostra a árvore filogenética construída a partir de três sequências de Acetyl-Coa transferase (PhaA) agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu dois grupos principais. O primeiro grupo contendo quatro sequências, sendo dois locus alinhados com duas sequências de PhaA com bootstrap variando de 95% indicando homologia. O segundo grupo contendo duas sequências, sendo um locus alinhado com uma sequência de PhaA com bootstrap 95% e apoiado no grupo anterior com 95% de bootstrap.

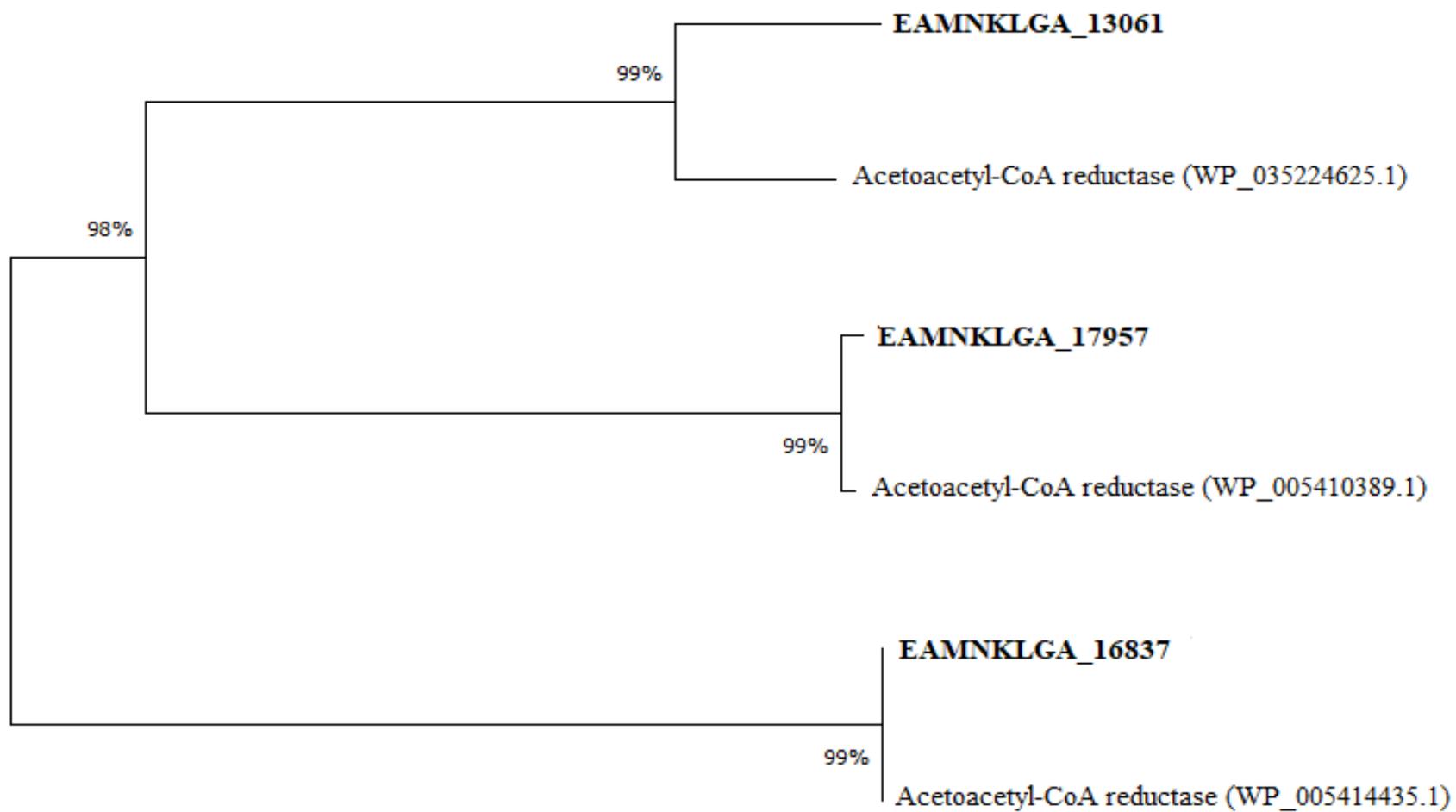
**Figura 8.** Árvore filogenética construída a partir de três sequências de PhaA alinhadas com as sequências mais similares, juntamente com um grupo externo do domínio de Archaea, retiradas do banco de dados RefSeq Protein.



No contexto das Acetyl-CoA Transferases (PhaA), utilizando o programa Mega 11, nos parâmetros escolhidos não foi possível construir uma árvore com todas as cinco sequências anotadas anteriormente, duas sequências não foram incluídas. As sequências EAMNKLGA\_13677 e EAMNKLGA\_13678 foram as que apresentaram os piores valores de e-value e score, como visto na Tabela 5 acima. Devido às suas baixas similaridades com as sequências retiradas do Reference Sequence Protein, não ocorreu um alinhamento global adequado entre todas as sequências de PhaA para que o programa Mega 11 conseguisse estabelecer uma relação mínima evolutiva entre elas. Possivelmente, essas duas sequências sejam sequências parciais, por isso os seus tamanhos reduzidos quando comparadas com as outras sequências anotadas. A árvore construída (Figura 8), revelou sequências de PhaA com possibilidade de serem altamente homólogas, baseado nos valores de bootstrap. Nesse caso, o genoma da bactéria *Priestia megaterium* cepa E1 carrega sequências de PhaA herdadas de um ancestral em comum com as sequências retiradas do banco de dados.

A Figura 9 mostra a árvore filogenética construída a partir das três sequências de Acetyl-Coa reductase (PhaB) agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu dois grupos. O primeiro grupo contendo quatro sequências, sendo dois locus alinhados com duas sequências de PhaB com bootstrap de 99% indicando homologia. O segundo grupo contendo duas sequências, sendo um locus alinhado com uma sequência de PhaB com bootstrap de 99% e apoiado no grupo anterior com 98% de bootstrap.

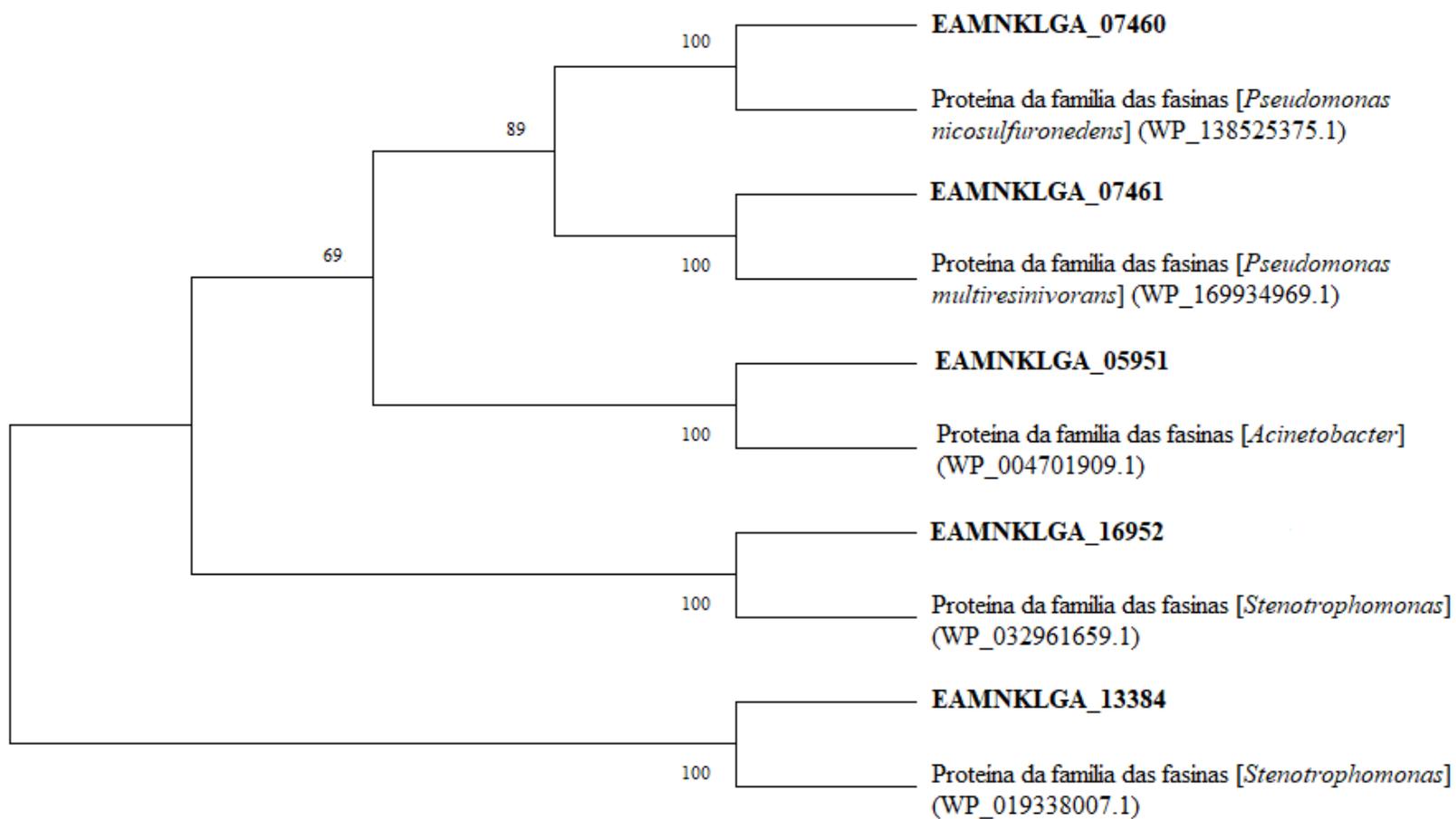
**Figura 9.** Árvore filogenética construída a partir das três sequências de PhaB alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.



A análise das Acetyl-CoA reductases (PhaB) revela dois grupos bem definidos. O alto valor de bootstrap de 99% nas sequências de PhaB sugere uma conservação evolutiva significativa, indicando uma herança genética compartilhada por diferentes organismos. Assim como as proteínas PhaA, as PhaB são cruciais em umas das rotas metabólicas de síntese de PHA, juntamente com a PhaC. Nesta rota, a PhaA e a PhaB sintetizam os substratos necessários para a ação da PhaC, que converte Acetil CoA em P3HB. Esse *cluster* de três enzimas (*phaACB*) é um dos mais comuns entre os organismos bioacumuladores de PHA já estudados, conforme destacado por Zhou e colaboradores (2023).

A Figura 10 mostra a árvore filogenética construída a partir das cinco sequências de proteínas da família de fasínas (PhaI) agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu três grupos. O primeiro grupo contendo seis sequências, sendo três loci alinhados com três sequências de PhaI com bootstrap de 100% e com um enraizamento de 89% de bootstrap. O segundo grupo contendo duas sequências, sendo um locus alinhado com uma sequência de PhaI com bootstrap de 100% e apoiado no grupo anterior com 69% de bootstrap. O terceiro grupo contendo duas sequências, sendo um locus alinhado com uma sequência de PhaI com bootstrap de 100% e apoiado nos grupos anteriores com 100% de bootstrap.

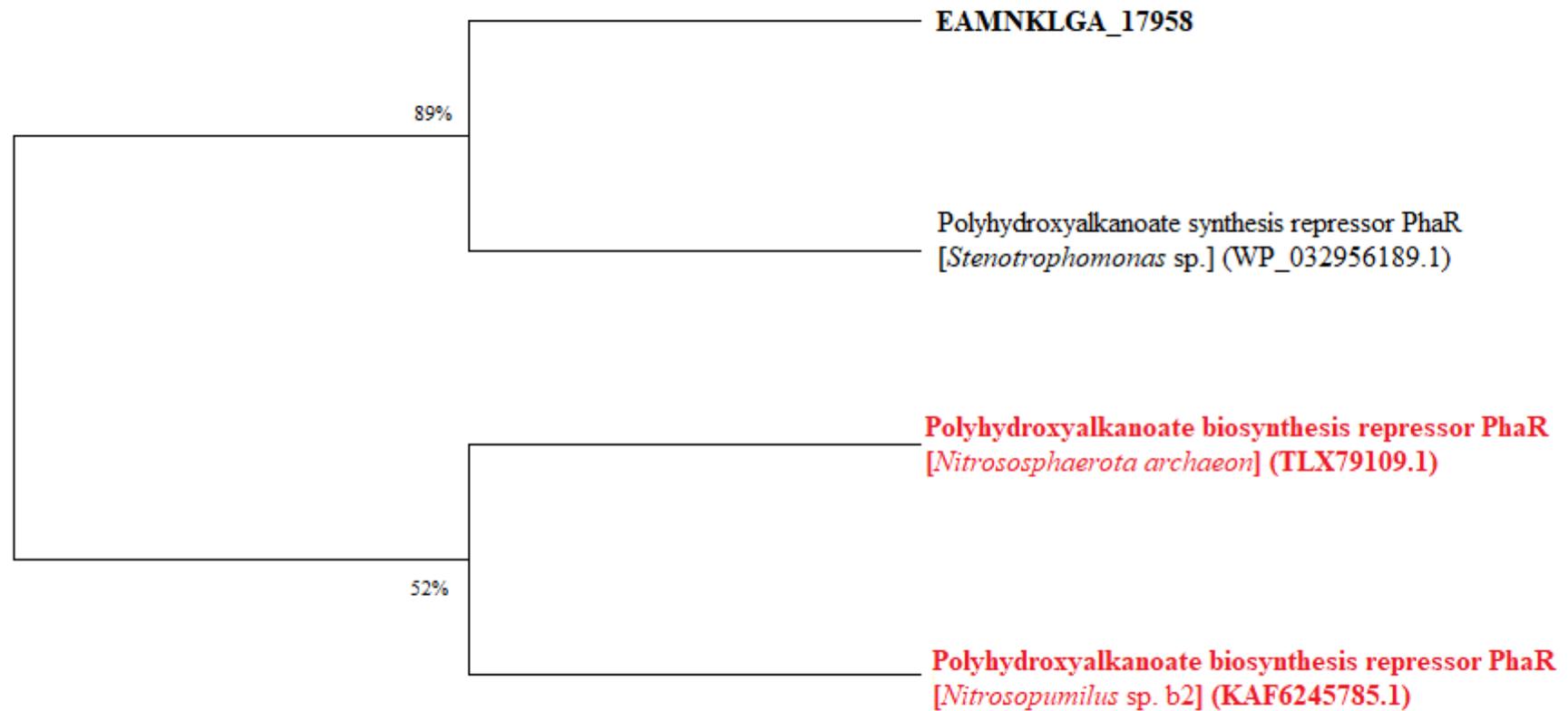
**Figura 10.** Árvore filogenética construída a partir das cinco sequências de proteínas fasinias alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.



A árvore filogenética das proteínas da família de fasínas (PhaI) mostram três grupos distintos. Diferente das outras árvores, essa em particular foi a única que foi utilizada a deleção completa das lacunas do alinhamento múltiplo global e também foi a única exibida no presente trabalho a sua árvore consenso. Dessa forma, o software Mega11 considera somente as regiões alinhadas e permite que a análise tenha sequências com o mesmo comprimento e dados completos, o que explica os agrupamentos terem obtido 100% de bootstrap. Essas enzimas são muito importantes na síntese de PHA e desempenham diferentes funções como a estabilização do polímero, regulação do tamanho e do número de grânulos (Vicente, D. et al., 2023). Em estudo, Aghaali, Z. e colaboradores (2023) sugere que existe uma relação direta positiva entre a expressão de proteínas fasinas durante síntese de PHAs – à medida que o PHA é sintetizado, aumenta a expressão de proteínas fasinas, em que aumenta o número de grânulos e, enquanto isso, amplia o tamanho dos grânulos ao se ligarem à superfície.

A Figura 11 mostra a árvore filogenética construída a partir da única sequência da proteína reguladora de PHA (PhaR) agrupada com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu um único grupo. O grupo contém quatro sequências, sendo um locus alinhado com uma sequência de PhaR e outras duas sequências de PhaR alinhadas do domínio de Archaeas (grupos externos), KAF6245785.1 e TLX79109.1, com bootstrap variando de 89% a 52% indicando homologia.

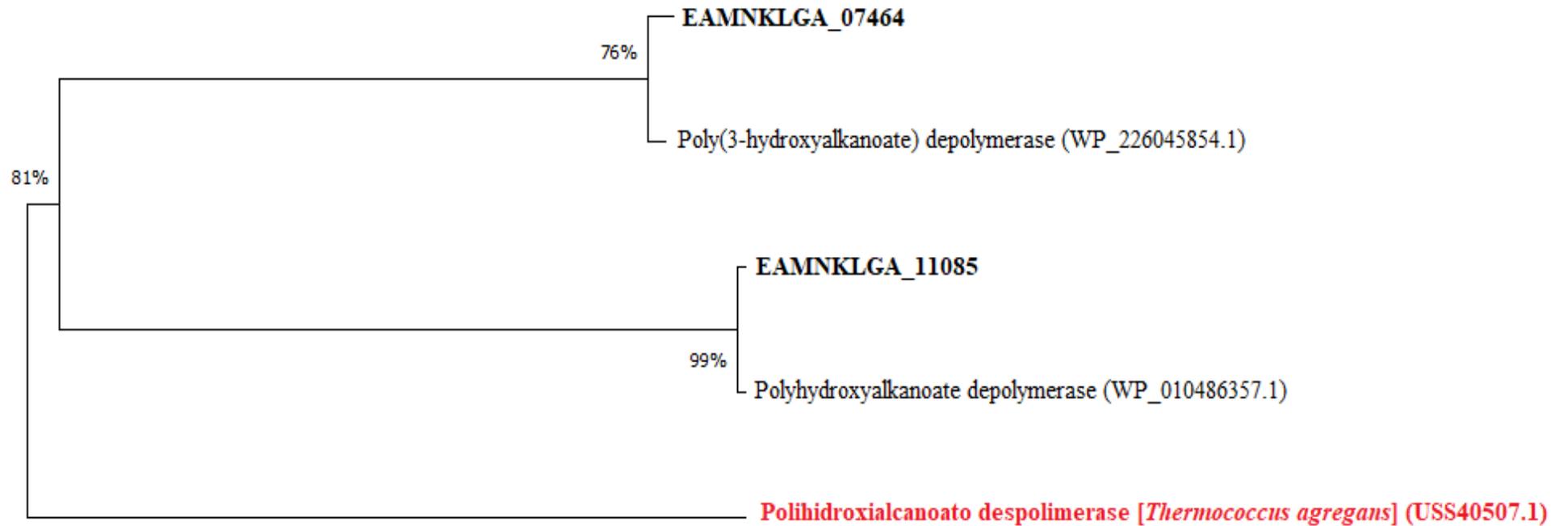
**Figura 11.** Árvore filogenética construída a partir da única sequência PhaR alinhada, juntamente com dois grupos externos do domínio de Archaea, com a sequência mais similar retirada do banco de dados RefSeq Protein.



Como a proteína PhaR foi a única sequência relacionada com PHA encontrada em uma única cópia no genoma da bactéria *Priestia megaterium* cepa E1, foi necessário adição de grupos externos. Para que o programa MEGA11 consiga estabelecer uma relação filogenética, são necessários no mínimo quatro táxons. Os grupos externos são adicionados nas análises filogenéticas para melhor compreensão da relação evolutiva de sequências homólogas entre diferentes organismos (Steenwyk, J.; King, N, 2023). A árvore filogenética da proteína reguladora de PHA (PhaR) revela um único grupo, mas compartilhando forte homologia entre o locus do genoma em questão com outras sequências de PhaR. Esta proteína age reprimindo a transcrição de proteínas fasinás, se ligando a região promotora e impedindo que sejam transcritas quando as bactérias não acumulam PHA. Assim, que as bactérias começam a síntese de PHAs, a PhaR sai e as fasinás são expressas (Aghaali, Z. et al., 2023).

A Figura 12 mostra a árvore filogenética construída a partir das duas sequências de PHA despolimerases agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu dois grupos. O primeiro grupo contendo quatro sequências, sendo dois locus alinhados com 2 sequências de PHA despolimerases com bootstrap variando de 76% a 99% indicando homologia. O segundo grupo (grupo externo), contém uma sequência de PHA despolimerase do domínio de Archaeas apoiado no outro grupo com 81% de bootstrap.

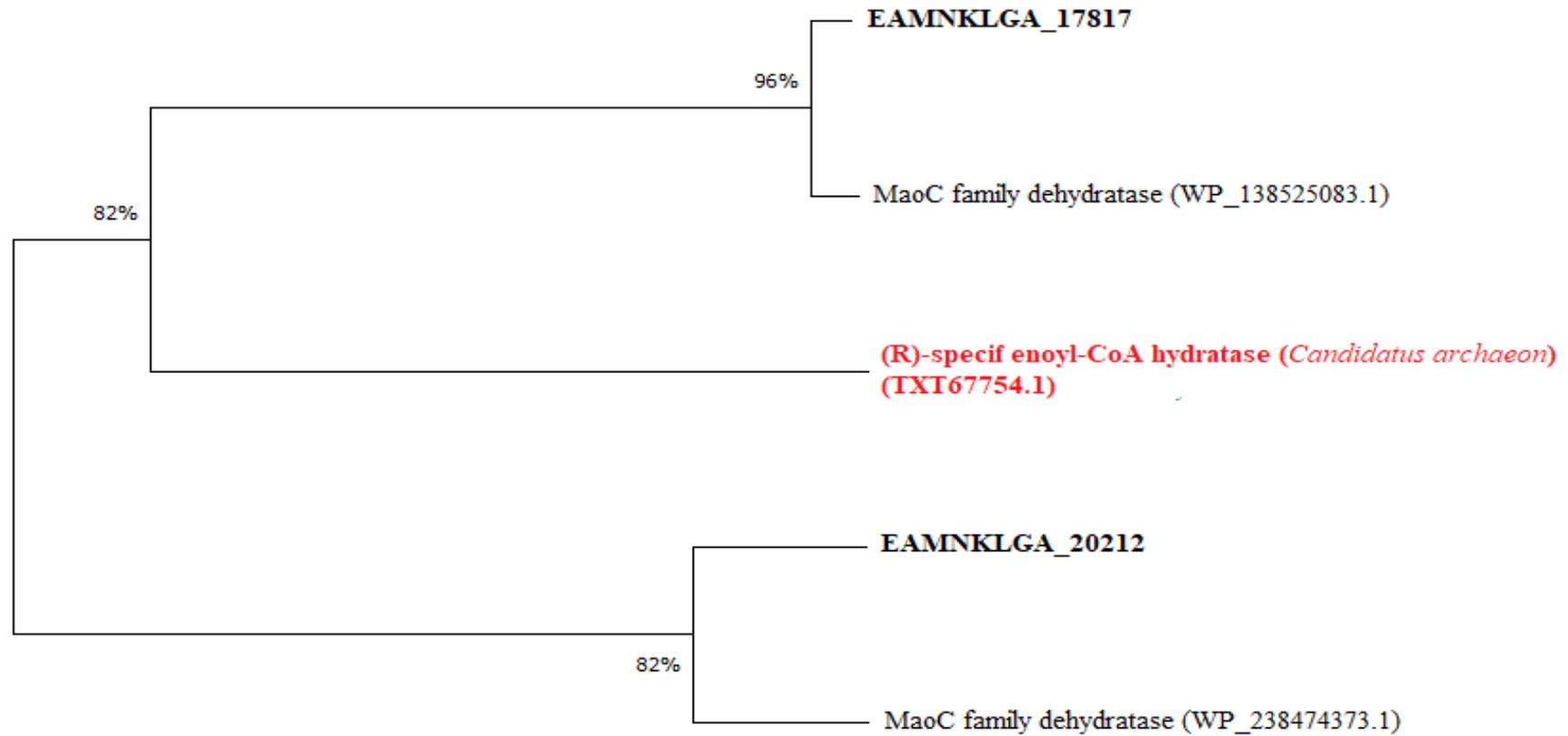
**Figura 12.** Árvore filogenética construída a partir das duas sequências de proteínas despolimerases, juntamente com um grupo externo do domínio de Archaea, alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.



A árvore filogenética das PHA despolimerases destaca a presença de dois grupos distintos com possibilidade de serem alta homólogos, baseado nos valores de bootstrap, e apoiado a um grupo externo de uma sequência do domínio de Achaea, o que ressalta também a diversidade evolutiva dessa enzima. Como o PHA é produzido pelos microrganismos como reserva de fonte de carbono e de energia no momento que a célula precisar, as PHA despolimerases são expressas para degradar os grânulos presentes no citoplasma da célula. Essa capacidade fornece aos polímeros de PHA a característica de biodegradabilidade, a qual é de suma importância no contexto ambiental – uma alternativa sustentável à persistência do plástico convencional. Desse modo, quando essas proteínas são traduzidas, diminui concomitante o acúmulo de PHAs no interior do citoplasma dos microrganismos (Zhou, W. et al., 2023).

A Figura 13 mostra a árvore filogenética construída a partir das duas sequências de (R)-specific enoyl-CoA hydratase (PhaJ) da família de MaoC desidratase agrupadas com as sequências mais similares retiradas do banco de dados RefSeq Protein (NCBI). Ela exibiu dois grupos. O primeiro grupo contendo três sequências, sendo um locus alinhado com uma sequência de PhaJ com bootstrap de 96% e sendo apoiado por um grupo externo do domínio de Archaeas (TXT67754.1) com 96% de bootstrap. O segundo grupo contendo duas sequências, sendo um locus alinhado com uma sequência de PhaJ com bootstrap de 82% e apoiado no grupo anterior com 82% de bootstrap.

**Figura 13.** Árvore filogenética construída a partir das duas sequências de PhaJ da família MaoC desidratase, juntamente com um grupo externo do domínio de Archaea, alinhadas com as sequências mais similares retiradas do banco de dados RefSeq Protein.



Já se tratando da árvore filogenética das (R)-specific enoyl-CoA hydratase (PhaJ), ela destaca a presença de dois grupos distintos com alta indicação de homologia. Assim como as PHA despolimerases, a inclusão de uma sequência do domínio Archaea no grupo externo ressalta a diversidade evolutiva dessas enzimas. A PhaJ é a principal enzima que atua na rota de  $\beta$ -oxidação que fornece intermediários que são essenciais para o crescimento celular e a síntese de PHA. Ela catalisa Enoyl-CoA em (R)-3-hydroxyacyl-ACP, o qual é substrato para PhaC sintase formar produtos de PHA (Giorgio, M. et al., 2020).

### 5.5 Análise físico química

A Tabela 7 mostra as características físico-químicas das 25 sequências de proteínas. Os tamanhos das sequências variam de 81 a 560 aminoácidos com um peso molecular (PM) variando de 8556.79 a 62966.70 kDa.

De acordo com Neoh, Z. S. e colaboradores (2022), as PhaC sintase de classe I têm o peso molecular de 63 a 73 kDa. Dentre as sete PhaC sintase anotadas no genoma da bactéria *Priestia megaterium* cepa E1, nenhuma apresentou PM acima de 62kDa. Já as PhaC de classe II, abrangem enzimas com o PM em torno de 62kDa, o que corresponde com os PM das sequências EAMNKLGA\_07463 (62966,70 kDa) e EAMNKLGA\_07465 (62312,08 kDa) presentes no genoma da bactéria *Priestia megaterium* cepa E1, codificadas pelos genes *phaC1* e *phaC2*, respectivamente, responsáveis pela expressão dessa classe de enzima (Song, L. et al., 2023). Já as PhaC sintases de classe III, são compostas por duas subunidades: uma subunidade maior PhaC, de PM 39 a 53 kDa, e outra subunidade menor PhaE, de 20 a 43 kDa. As sequências de PhaC de classe III presentes no genoma da bactéria *Priestia megaterium* cepa E1, a sequência EAMNKLGA\_22716, anotada como sendo uma subunidade PhaE, e a sequência EAMNKLGA\_22717, sendo uma subunidade PhaC, têm os PM dentro da faixa de peso apontada por Neoh – a primeira, 38841,58 kDa, e a segunda, 39928,87 kDa. As outras duas sequências de PhaC sintase, EAMNKLGA\_05917 e EAMNKLGA\_13058, apresentaram o peso molecular de 40944,23 kDa e 61420,96 kDa, respectivamente. A partir das ferramentas de anotação, não foi especificado a qual classe de PhaC sintase estas duas pertencem, somente que são polimerases de polihidroxicanoato (PHA), pertencentes a família de alfa/beta hidrolase, sendo esta uma estrutura geral de PhaC sintase semelhante à das lipases, com base em sua

semelhança e modelos de encadeamento, conforme aponta Neoh, Z. S. e colaboradores (2022).

O valor de pH em que todas as cargas negativas e positivas de uma proteína se anulam é conhecido como ponto isoelétrico (pI) (Guimarães, M. B., 2021). O ProtParam forneceu o pI teórico das proteínas que variaram de 4.61 a 9.58. Dentre as 25 sequências, 17 apresentaram o ponto isoelétrico em pH ácido. Outras oito sequências tiveram o valor de pH acima ou próxima de sete (Santos, F. R. F. et al., 2021).

O índice de instabilidade (pII) é uma medida teórica para determinar a estabilidade de uma proteína em tubo de ensaio (Guimarães, M. B., 2021). A ferramenta determina um valor de corte sendo 40, então qualquer pII acima de 40 representa uma provável instabilidade e abaixo uma provável estabilidade no comportamento das proteínas *in vitro*. Os valores obtidos de pII para oito sequências foram superiores a 40, sendo consideradas instáveis. As outras 17 sequências foram consideradas estáveis os seus comportamentos *in vitro*, com pII abaixo de 40 (Santos, F. R. F. et al., 2021).

O índice alifático (IA) de uma proteína é uma medida que indica a quantidade de grupos alifáticos em sua estrutura (Guimarães, M. B., 2021). Os grupos alifáticos são aqueles compostos principalmente por átomos de carbono e hidrogênio, como as cadeias laterais de aminoácidos não polares. As proteínas tiveram o valor do IA variando de 59.45 a 113.49, o que significa que, em porcentagem (%), o volume da proteína é composto por essas cadeias laterais alifáticas. Ikai (1980) descobriu que há uma correlação entre o IA de uma proteína e sua termoestabilidade, ou seja, a capacidade da proteína de resistir a altas temperaturas sem perder sua estrutura ou função. Portanto, quanto maior o valor do IA, maior a probabilidade de a proteína ser termoestável, pois as cadeias alifáticas contribuem para a estabilidade da estrutura da proteína em condições de calor (Ramya, P. et al., 2023).

O Grand Average of Hydropathy (GRAVY) reflete a interação da proteína com a água, variando de -2 a +2. Valores positivos indicam maior hidrofobicidade, enquanto valores negativos indicam maior afinidade pela água (Guimarães, M. B., 2021). Das 25 proteínas em questão, sete apresentaram carácter hidrofóbico (valores positivos). As outras 18 sequências apresentaram valores inferiores a 0,0 (valores negativos), indicando uma afinidade pela água (Bouqellah, N. A., 2023).

O coeficiente de extinção (CE) fornece informações sobre a capacidade de absorção de luz de uma proteína em um comprimento de onda específico ( $\lambda$ ), sendo diretamente influenciado pela quantidade de resíduos aromáticos presentes (Guimarães,

M. B., 2021). Os valores obtidos para o coeficiente de extinção das 25 sequências variam de 2980 a 122840 M<sup>-1</sup> cm<sup>-1</sup> a 280 nm em condições aquáticas (SANTOS, F. R. F. et al., 2021).

**Tabela 7.** Resultado das características físico químicas obtido no software ProtParam das 25 sequências anotadas.

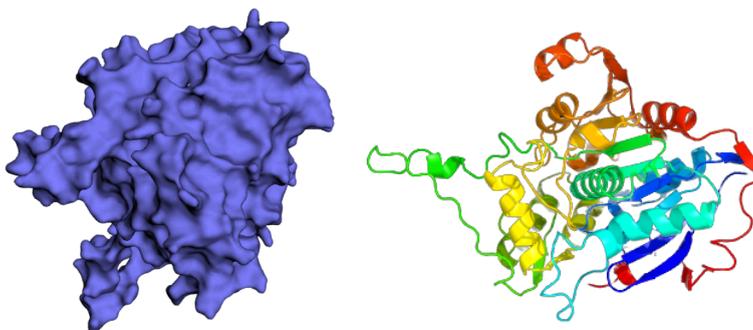
Query	Parâmetros						
	pI	II	IA	GRAVY	CE	PM	Tamanho
EAMNKLGA_05793	5.46	24.08	91.04	0.073	20970	40745,53	394
EAMNKLGA_05917	9.58	38.89	99.42	-0.145	53860	40944.23	361
EAMNKLGA_07463	8.86	47.33	86.43	-0.338	109320	62966.70	560
EAMNKLGA_07465	6.96	36.70	86.78	-0.277	101870	62312.08	559
EAMNKLGA_10818	6.33	33.02	98.55	0.111	22460	40454.38	392
EAMNKLGA_13058	5.87	41.18	86.03	-0.246	122840	61420.98	547
EAMNKLGA_13061	8.95	23.01	88.95	-0.007	23950	26199.96	248
EAMNKLGA_13677	4.61	34.03	111.98	0.246	2980	8556.79	81
EAMNKLGA_13678	4.55	19.67	113.49	0.066	8480	9212.52	86
EAMNKLGA_16837	7.73	36.75	85.73	0.036	36440	25886.20	246
EAMNKLGA_17817	5.20	26.71	102.50	-0.064	4470	16818.33	156
EAMNKLGA_17957	6.31	24.96	84.47	-0.139	36440	26496.92	246
EAMNKLGA_20212	5.26	25.25	95.48	0.047	8480	14517.57	135
EAMNKLGA_20844	6.04	35.42	101.70	0.110	15470	40839.90	393
EAMNKLGA_22690	9.56	33.77	98.72	-0.045	61880	43136.17	383
EAMNKLGA_22716	9.39	59.62	63.45	-0.604	69440	38841.58	354
EAMNKLGA_22717	5.05	30.80	100.79	-0.123	49850	39928.87	355
EAMNKLGA_17958	5.09	43.68	67.16	-0.679	18450	20391.83	183
EAMNKLGA_05951	6.34	39.34	91.66	-0.728	6990	16932.22	151
EAMNKLGA_07460	9.45	28.29	92.65	-0.342	11460	14849.94	136
EAMNKLGA_13384	5.26	32.67	73.39	-0.377	17990	12231.53	115
EAMNKLGA_16952	5.06	37.74	59.45	-0.974	23490	17989.57	163
EAMNKLGA_07461	10.41	40.43	66.25	-0.341	19480	31522.87	323
EAMNKLGA_07464	9.54	49.56	89.79	-0.080	46410	36738.65	334
EAMNKLGA_11085	6.20	32.22	85.20	-0.272	70360	49055.70	435

pI – Ponto isoelétrico, II – Índice de instabilidade, IA – Índica alifático, GRAVY - Grand average of hydropathy, CE – Coeficiente de extinção, PM – Peso molecular, Tamanho – quantidade de aminoácido

## 5.6 Modelagem 3D

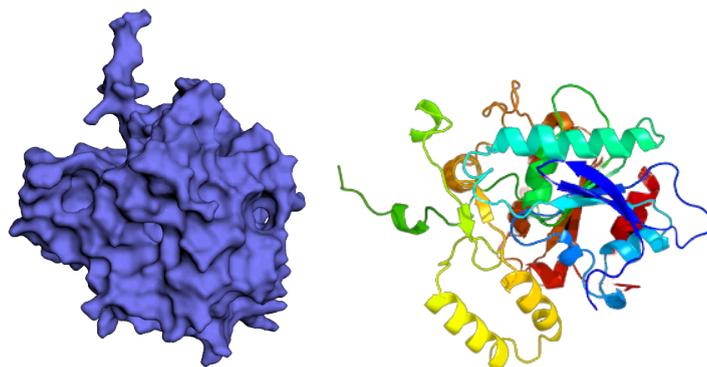
A modelagem da sequência Poly-beta-hydroxybutyrate polymerase (PhaC2) (EAMNKLGA\_07463) apresentou uma confiança de 100% e uma cobertura de 67% de acordo com o programa Phyre2. A Figura 14 está representando a visualização da modelagem da proteína PhaC sintase de classe II codificada pelo gene *phaC2* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 14.** Visualização da modelagem da proteína Poly-beta-hydroxybutyrate polymerase (PhaC2) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.



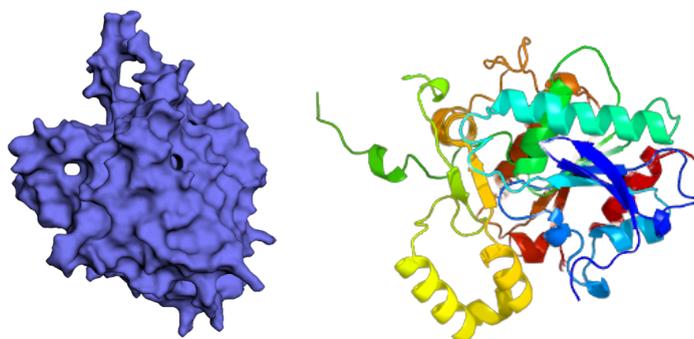
A Figura 15 está representando a modelagem da sequência Poly(3-hydroxyalkanoate) polymerase (PhaC) (EAMNKLGA\_07465), a qual apresentou uma confiança de 100% e uma cobertura de 87% de acordo com o programa Phyre2. Na Figura está representando a visualização da modelagem da proteína PhaC sintase codificada pelo gene *phaC* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 15.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.



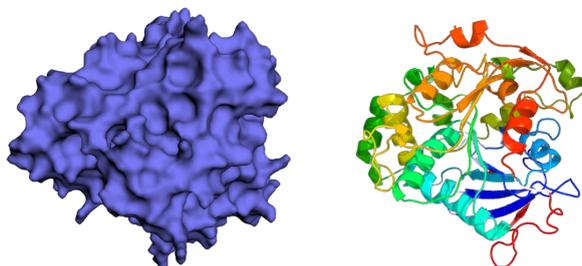
A Figura 16 está representando a modelagem da sequência Poly-beta-hydroxybutyrate polymerase (PhaC) (EAMNKLGA\_05917), a qual apresentou uma confiança de 100% e uma cobertura de 67% de acordo com o programa Phyre2. Na Figura está representando a visualização da modelagem da proteína PhaC codificada pelo gene *phaC* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 16.** isualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.



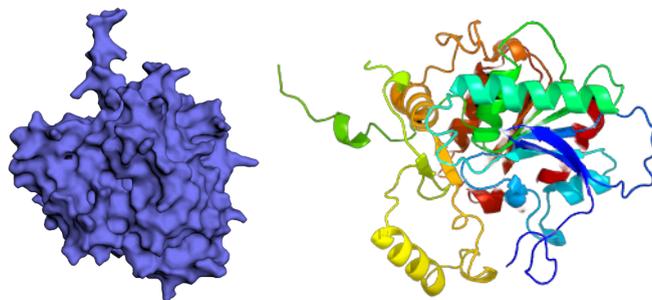
A Figura 17 está representando a modelagem da sequência Poly-beta-hydroxybutyrate polymerase (PhaC) (EAMNKLGA\_13058), a qual apresentou uma confiança de 100% e uma cobertura de 67% de acordo com o programa Phyre2. Na Figura está representando a visualização da modelagem da proteína PhaC codificada pelo gene *phaC* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 17.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.



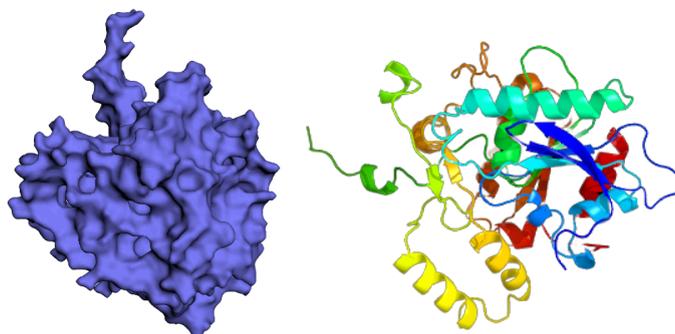
A Figura 18 está representando a modelagem da sequência Poly-beta-hydroxybutyrate polymerase (PhaC) (EAMNKLGA\_22690), a qual apresentou uma confiança de 100% e uma cobertura de 67% de acordo com o programa Phyre2. Na Figura está representando a visualização da modelagem da proteína PhaC codificada pelo gene *phaC* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 18.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase (PhaC) do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no model cartoon, à direita.



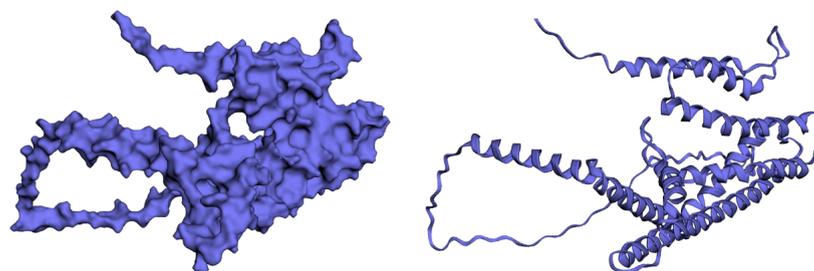
A Figura 19 está representando a modelagem da sequência Poly(3-hydroxyalkanoate) polymerase subunidade PhaC de classe III (EAMNKLGA\_22717), a qual apresentou uma confiança de 100% e uma cobertura de 87% de acordo com o programa Phyre2. Na Figura está representando a visualização da modelagem da proteína PhaC codificada pelo gene *phaC* presente no genoma da bactéria *Priestia megaterium* cepa E1.

**Figura 19.** Visualização da modelagem da proteína Poly(3-hydroxyalkanoate) polymerase de classe III subunidade PhaC do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Phyre2 exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no modelo cartoon, à direita.

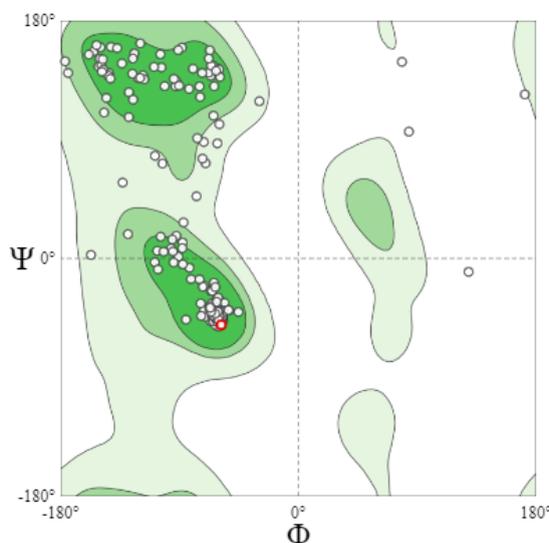


O programa Phyre2 não conseguiu modelar a sequência de PhaC sintase subunidade PhaE (EAMNKLGA\_22716) a um nível de cobertura e confiança significativo. Então, foi utilizado o Swiss-model (Schwede, T. et al., 2003) para modelar esta sequência. A Figura 20 está representando a modelagem da sequência Subunidade PhaE de poli(3-hidroxiálcanoato) polimerase de classe III, a qual apresentou uma identidade de 71,14% e uma cobertura de quase 98% de acordo com o Swiss-model. Além disso, o programa fornece outros parâmetros avaliativos, como o gráfico de Ramachandran (Figura 21) que apresentou 91.84% de favorecimento à modelagem. Na Figura 20 abaixo está representando a visualização da modelagem da proteína que é codificada pelo gene *phaE* presente no genoma da bactéria *Priestia megaterium* cepa E1

**Figura 20.** Visualização da modelagem da proteína subunidade PhaE de poli(3-hidroxiálcanoato) polimerase de classe III do genoma da bactéria *Priestia megaterium* cepa E1 realizada pelo programa Swiss-model exibida no programa EzMol no modelo de nuvem de elétrons, à esquerda, e no modelo cartoon, à direita.



**Figura 21.** Gráfico de Ramachandran da modelagem da proteína subunidade PhaE de poli(3-hidroxialcanoato) polimerase de classe III (EAMNKLGA\_22716) realizada pelo Swiss-model.



A estrutura geral das PhaC sintases, em geral, é de uma dobra típica  $\alpha/\beta$ -hidrolase (alfa/beta hidrolase). Essa arquitetura é semelhante à das esterases, com base em suas semelhanças e modelos de encadeamento (Jendrossek, D. 2009). É composta por dois complexos, um domínio C-terminal e outro domínio N-terminal. As estruturas cristalográficas do domínio catalítico C-terminal foram resolvidas com sucesso para a PHA sintase (PhaC) de *C. necator* cepa H16 e *Chromobacterium sp.* cepa USM2 (Wittenborn, E. C. et al., 2016; Chek, M. F, 2017). O domínio catalítico C-terminal contém um subdomínio, o qual é um canal de entrada do substrato, sítio ativo e túnel de saída do produto. Até o momento, a função exata do domínio N-terminal não foi determinada e a estrutura dessa região não pode ser visualizada com sucesso por meio da cristalização (Wittenborn, E. C. et al., 2016).

A partir das modelagens das PhaC sintase pelo Phyre2 e visualizadas pelo EzMol, no modelo cartoon, foi possível observar a presença das estruturas alfa-hélice e folha-beta pregueadas em todas as sequências. As estruturas dessas proteínas se assemelham com a investigação das estruturas de PhaC sintases feita por Neoh, Z. S. e colaboradores (2022). Quanto a modelagem da proteína subunidade PhaE (EAMNKLGA\_22716) realizada pelo Swiss-model, ela apresentou semelhança com as estruturas das outras PhaC sintase. Além disso, o gráfico de Ramachandran apresentou cerca de 91,84% dos resíduos de aminoácidos dentro da zona de favorecimento,

enquanto resíduos fora dessa zona pode indicar problemas na estrutura, como erros de modelagem ou desdobramento incorreto (Schwede, T. et al., 2003).

## 6 CONCLUSÕES

Foram anotadas 25 sequências de polihidroxialcanoato (PHA) no genoma da espécie *Priestia megaterium* cepa E1 isolada de solo. Estas sequências foram anotadas como cinco referentes a PhaA (Acetyl-CoA transferase), sete PhaC (Poly(3-hydroxyalkanoate) polymerase), três PhaB (Acetoacetyl-CoA reductase), duas PhaJ (Enoyl-CoA hydratase), cinco PhaI (Proteínas fasinas), uma PhaR (Proteína reguladora) e duas PhaZ (Poly(3-hydroxyalkanoate) depolymerase).

As anotações automáticas realizadas pelos programas PROOKA e eggNOG-mapper demonstraram similaridades entre as anotações feitas das 25 sequências nos dois programas. Os alinhamentos realizados nos bancos de dados Swiss-Prot (UniProt) e RefSeq Protein com as sequências do genoma da bactéria *Priestia megaterium* cepa E1 apresentaram valores bons de e-value e fortes identidades, com exceção de duas sequências (EAMNKLGA\_13677 e EAMNKLGA\_13678) anotadas como sendo PhaAs, que tiveram os piores valores de score, e-value e identidade - possivelmente sejam sequências parciais.

De modo geral, as árvores apresentadas demonstraram altos valores de bootstrap indicando homologia entre as sequências do genoma da bactéria *Priestia megaterium* cepa E1 com aquelas depositadas em banco de dados secundário (RefSeq protein). Vale ressaltar que não foi possível construir uma árvore filogenética contendo todas as cinco sequências de PhaAs, devido estas duas sequências supracitadas (EAMNKLGA\_13677 e EAMNKLGA\_13678) não apresentarem agrupamentos com as outras sequências.

Através da análise físico química pelo programa ProtParam foi possível constatar que as sete sequências anotadas como sendo PhaC sintases apresentaram pesos moleculares semelhantes com estudos já feitos destas enzimas.

A modelagem 3D das sete sequências de proteínas de PhaC sintases apresentaram arquiteturas com estruturas semelhantes a estudos com as estruturas cristalográficas já resolvidas dessas enzimas. Vale ressaltar que a sequência de PhaE sintase (EAMNKLGA\_22716) não foi modelada programa Phyre2 por não apresentar nível de cobertura e confiança significativo. Ela foi então modelada pelo programa Swiss Model que apresentou uma modelagem confiável, segundo o gráfico de

Ramachandran. Todos esses resultados e indícios forneceram mais confiabilidade ao processo de anotação de sequências genômicas de PHAs feita.

Por fim, vale ressaltar, que o presente estudo identificou uma diversidade de genes e proteínas que estão relacionados com o metabolismo de PHA na bactéria *Priestia megaterium* cepa E1 isolada de solo, corroborando com evidências científicas já feitas na espécie desde 1926. Estas informações são relevantes e podem ser utilizadas para avaliar experimentalmente o potencial de produção de PHA no isolado, bem como complementar a compreensão sobre as vias metabólicas de PHA em bactérias. Essas descobertas também podem ser aplicadas em outras espécies, ampliando assim o conhecimento de possíveis bioacumuladores de PHA.

## REFERÊNCIAS

Abdallah, B. M.; Sayadi, S.; Karray, F. Overproduction of poly(3-hydroxybutyrate) by the extremely halophilic archaeon *Haloarcula* sp. CEJ40-10 using starchy substrate. *Biomass Conv. Bioref.* 2023. <https://doi.org/10.1007/s13399-023-03909-3>.

Acuña, J. M. B.; Castro, I. P. Rational engineering of natural polyhydroxyalkanoates producing microorganisms for improved synthesis and recovery. *Microbial Biotechnology*. 16. 10.1111/1751-7915.14109. 2022. <https://doi.org/10.1111/1751-7915.14109>.

Adeleye, A. T.; Odoh, C. K.; Enudi, O. C.; Banjoko, O. O.; Odediran, E. T.; Louis, H. Sustainable synthesis and applications of polyhydroxyalkanoates (PHAs) from biomass, *Process Biochemistry*, Volume 96, Pages 174-193, ISSN 1359-5113. 2020. <https://doi.org/10.1016/j.procbio.2020.05.032>.

Afgan, E.; Baker, D.; Batut, B.; Beek, M.; Bouvier, D.; Cech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B. A.; Guerler, A.; Hillman-Jackson, J.; Hiltemann, S.; Jalili, V.; Rasche, H.; Soranzo, N.; Goecks, J.; Taylor, J.; Nekrutenko, A.; Blankenberg, D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46(W1):W537-W544. 2018. <https://doi.10.1093/nar/gky379>.

Aghaali, Z.; Naghavi, M. R. Biotechnological Approaches for Enhancing Polyhydroxyalkanoates (PHAs) Production: Current and Future Perspectives. *Curr Microbiol* 80, 345. 2023. <https://doi.org/10.1007/s00284-023-03452-4>.

Alaux, E.; Couvreur, M.; Marie, B.; Bounouba, M.; Hernandez-Raquet, G. Biosynthesis of medium chain length polyhydroxyalkanoates (mcl-PHA) by activated sludge: Impact of phosphorus limitation on pha production and microbial diversity. **SSRN Electronic Journal**. 2022. <https://ssrn.com/abstract=4028808>.

Alves, A. A.; Siqueira, E. C.; Barros, M. P. S.; Polyhydroxyalkanoates: a review of microbial production and technology application. **International J. Environment. Science. Technology**. 20, 3409–3420. 2023. <https://doi.org/10.1007/s13762-022-04213-9>.

Alwuhaib, A. S.; Zinkevich, V.; Kartvelishvili, T.; Asatiani, N.; Sapojnikova, N. Bioplastics against Microplastics: Screening of Environmental Bacteria for Bioplastics Production Chapters. **Advances and Challenges in Microplastics**. 2023. <https://ideas.repec.org/h/ito/pchaps/290559.html>.

Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A. A.; Dvorkin, M.; Kulikov, A. S.; Lesin, V. M.; Nikolenko, S. I.; Pham, S.; Prjibelski, A. D.; Pyshkin, A. V.; Sirotkin, A. V.; Vyahhi, N.; Tesler, G.; Alekseyev, M. A.; Pevzner, P. A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **J Comput Biol**. 19(5):455-77. 2012. <https://doi.org/10.1089/cmb.2012.0021>.

Bairoch, A.; Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL. **Nucleic Acids Res**. 1;28(1):45-8. 2000. <https://doi:10.1093/nar/28.1.45>

Behera, S.; Priyadarshane, M.; Surajit, M. D.. Polyhydroxyalkanoates, the bioplastics of microbial origin: Properties, biochemical synthesis, and their applications. **Chemosphere**, 294, 133723. 2022. <https://doi.org/10.1016/j.chemosphere.2022.133723>.

Bouqellah, N. A.; Farag, P. F. In Silico Evaluation, Phylogenetic Analysis, and Structural Modeling of the Class II Hydrophobin Family from Different Fungal Phytopathogens. **Microorganisms**. 11, 2632.2023. <https://doi.org/10.3390/microorganisms11112632>.

Cal, A. J.; Kibblewhite, R. E.; Sikkema, W. D.; Torres, L. F.; Hart-Cooper, W. M.; Orts, W. J.; Lee, C. C. Production of polyhydroxyalkanoate copolymers containing 4-hydroxybutyrate in engineered *Bacillus megaterium*. **Int. J. Biol. Macromol**. 168, 86–92. 2021. <http://doi.org/10.1016/j.ijbiomac.2020.12.015>.

Cantalapiedra, C. P.; Hernández-Plaza, A.; Letunic, I.; Bork, P.; Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain

Prediction at the Metagenomic Scale. **Mol Biol Evol.** 9;38(12):5825-5829. 2021  
<https://doi.org/10.1093/molbev/msab293>.

Carvalho, L. M; Borelli, G; Assis, M. A; Ferraz, S. M. F; Fiamenghi, M. B; José, J; Mofatto, L. S; Nagamatsu, S. T; Persinoti, G. F; Silva, N. V; Vasconcelos, A. A; Pereira, G. A. G, Carazzole, M. F. Bioinformatics applied to biotechnology: A review towards bioenergy research. **Biomass and Bioenergy**, v. 123, p. 195-224, 2019.

Choi, S. Y.; Cho, I. J.; Lee, Y.; Kim, Y. J.; Kim, K. J.; Lee, S. Y. Microbial Polyhydroxyalkanoates and Nonnatural Polyesters. **Adv. Mater.** 2020, 32, e1907138.  
<http://doi.org/10.1007/s10924-020-01772-1>.

Cock, P. J.; Fields, C. J.; Goto, N.; Heuer, M. L.; Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. **Nucleic Acids Res.**38(6):1767-71. 2010. <http://doi.org/10.1093/nar/gkp1137>.

Costa, C. Z.; Albuquerque, M. C. C.; Brum, M. C.; Castro, A. M. DEGRADAÇÃO MICROBIOIÓGICA E ENZIMÁTICA DE POLÍMEROS: UMA REVISÃO. **Quím Nova.** 38(2):259–67. 2015. <https://doi.org/10.5935/0100-4042.20140293>.

Dana, R. H.; Ebrahimi, F. Synthesis, properties, and applications of polylactic acid-based polymers. **Polym. Eng. Ciência.** , 63 ( 1 ), 22. 2023.  
<https://doi.org/10.1002/pen.26193>.

Deng, W.; Li, C.; Xie, J. The underling mechanism of bacterial TetR/AcrR family transcriptional repressors, Cellular Signalling, Volume 25, Issue 7, Pages 1608-1613, ISSN 0898-6568. 2013. <https://doi.org/10.1016/j.cellsig.2013.04.003>.

European Bioplastics. **europaebioplastics**. Global bioplastic production capacities continue to grow despite low oil prices. 2020. Disponível em:  
<<https://www.european-bioplastics.org/market-data-update-2020/>> Acesso em: 30 dez. 2023.

Fortunato, P. E.; Vittoria, V.; Carmine, B.; Pietro, T.; Brendon, N.; Pooja, B.; Donatella, P. Enhanced production of biobased, biodegradable, Poly(3-hydroxybutyrate) using an unexplored marine bacterium *Pseudohalocynthiibacter aestuariivivens*, isolated from highly polluted coastal environment, **Bioresource Technology**, Volume 368, 2023.  
<https://doi.org/10.1016/j.biortech.2022.128287>.

Galperin, M. Y.; Kristensen, D. M.; Makarova, K. S.; Wolf, Y. I.; Koonin, E. V. Microbial genome analysis: the COG approach. **Brief Bioinform.** 19;20(4):1063-1070. 2019. <https://doi:10.1093/bib/bbx117>

Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): **The Proteomics Protocols Handbook**, Humana Press, pp. 571-607. 2005.

Giorgio, M.; Dario, P.; Jarillo, M. Recovery of polyhydroxyalkanoates (PHAs) from wastewater : a review». **Bioresource Technology**, Vol. 297, art. 122478. 2020. <https://doi.10.1016/j.biortech.2019.122478>.

Grigore, M. E.; Grigorescu, R. M.; Iancu, L.; Ion, R.M.; Zaharia, C.; Andrei, C. R. Synthesis methods, properties and biomedical applications of polyhydroxyalkanoates: a review. **Journal of Biomaterials Science**. Polymer Edition, 30:9, 695-712. 2019. <https://doi.org/10.1080/09205063.2019.1605866>.

Guimarães, M. B. Bioprospecção in sílico de genes codificadores de Laspariginase em amostras de solo de Dourados-MS. 2021. 52 f. TCC (Graduação) - Curso de Biotecnologia, FCBA, Universidade Federal da Grande Dourados, Dourados, 2021. Disponível em: <https://repositorio.ufgd.edu.br/jspui/handle/prefix/4959>. Acesso em: 20 nov. 2023.

Guleria, S.; Singh, H.; Sharma, V.; Bhardwaj, N.; Arya, K. S.; Puri, S. ; Khatri, M. Polyhydroxyalkanoates production from domestic waste feedstock: A sustainable approach towards bio-economy. **Journal of Cleaner Production**. Volume 340, 130661, ISSN 0959-6526. 2022. <https://doi.org/10.1016/j.jclepro.2022.130661>.

Gutschmann, B.; Huang, B.; Santolin, L.; Thiele, I.; Neubauer, P.; Riedel, S. L. Native feedstock options for the polyhydroxyalkanoate industry in Europe: A review, **Microbiological Research**, Volume 264, 127177, ISSN 0944-5013, 2022. <https://doi.org/10.1016/j.micres.2022.127177>.

Haddadi, M. H.; Asadolahi, R.; Negahdari, B. The bioextraction of bioplastics with a focus on polyhydroxybutyrate: a review. **International J. Environment. Science. Technology**. 16, 3935–3948. 2019. <https://doi.org/10.1007/s13762-019-02352-0>.

Heberle, H.; Meirelles, G. V.; da Silva, F. R.; Telles, G. P.; Minghim, R. InteractiVenn: a webbased tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics** 16:169. 2015.

Herrera, R. E. M. et al. A comprehensive view of *Bacillus cereus* as a polyhydroxyalkanoate (PHA) producer: A promising alternative to Petroplastics. **Process Biochemistry**. 129. 281-292. 2023. <https://www.elsevier.com/locate/procbio> .

Jaffur, B. N.; Kumar, G; Pratima; Seeram, J; Bhatia, S. C. Current advances and emerging trends in sustainable polyhydroxyalkanoate modification from organic waste streams for material applications, International. **Journal of Biological Macromolecules**, Volume 253 .2023. <https://doi.org/10.1016/j.ijbiomac.2023.126781> .

Junker, V.; Contrino, S.; Fleischmann, W.; Hermjakob, H.; Lang, F.; Magrane, M.; Martin, M. J.; Mitalaritonna, N.; O'Donovan, C.; Apweiler, R. The role SWISS-PROT and TrEMBL play in the genome research environment. **J Biotechnol**, 31;78(3):221-34. 2000. [https://doi.org/10.1016/s0168-1656\(00\)00198-x](https://doi.org/10.1016/s0168-1656(00)00198-x)

Kapoor, D.D.; Yadav, S.; Gupta, R. K. Comprehensive study of microbial bioplastic: current status and future prospects for sustainable development. **Support for Environmental Development**. 2023. <https://doi.org/10.1007/s10668-023-03620-3>

Karikari, T.K.; Aleksic, J. Neurogenomics: An opportunity to integrate neuroscience, genomics and bioinformatics research in Africa. **Applied & Translational Genomics**. Volume 5, Pages 3-10, ISSN 2212-0661. 2015. <https://doi.org/10.1016/j.atg.2015.06.004>.

Katagi, V. N.; Bhat, S. G.; Paduvari, R.; Kodavooru, D.; Somashekara, D. Waste to value-added products: an innovative approach forsustainable production of microbial biopolymer (PHA) - emphasis on inexpensive carbon feedstock. **Environmental Technology Reviews**. 2162-2515. 2023. <https://doi.org/10.1080/21622515.2023.2250066>.

Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. **Nat Protoc**. 10(6):845-58. 2015. <https://doi.org/10.1038/nprot.2015.053>.

Khalil, A. H. P. S.; Jha, K.; Yahya, E. B.; Panchal, S.; Patel, N.; Garai, A.; Kumari, S.; Jameel, M. Insights into the potential of biopolymeric aerogels as advanced soil fertilizer delivery systems. **Gels**, 9, 666. 2023. <https://doi.org/10.3390/gels9080666>

Kim, M. S.; Chang, H.; Zheng, L.; Yan, Q.; Brian, F. P.; John, K.; Nelson, K.; Majumder, E. L. W.; Huber, G. W. A review of biodegradable plastics: chemistry, applications, properties and future research needs. **Chemical Reviews**. 123, (16). 2023. <https://doi.org/10.1021/acs.chemrev.2c00876>.

Koichiro, T.; Glen, S.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Versão 11, **Molecular Biology and Evolution**, Volume 38, Edição 7, julho de 2021, páginas 3022–3027. <https://doi.org/10.1093/molbev/msab120>.

Koonin, E. V. "Orthologs, Paralogs, and Evolutionary Genomics", *Annu. Rev. Genet.*, 39, 309-338. 2005.

Krueger, F. Trim Galore. In **GitHub repository**. 2021. Acesso em: <<https://github.com/FelixKrueger/TrimGalore.com>> Acessado em: 17 Apr 2024.

Kumar, M.; Rathour, R.; Singh, R.; Sun, Y.; Pandey, A.; Gnansounou, E.; Lin, Y. L.; Tsang, D. C. W.; Thakur, I. S. Bacterial polyhydroxyalkanoates: Opportunities, challenges, and prospects, **Journal of Cleaner Production**, Volume 263, 121500, ISSN 0959-6526. 2020. <https://doi.org/10.1016/j.jclepro.2020.121500>

Kumar, R.; Dongyi, L.; Liwen, L.; Manu, M. K.; Jun, Z.; Rajeshwar, D. T.; Jonathan, W.C. Genome-centric polyhydroxyalkanoate reconciliation reveals nutrient enriched growth dependent biosynthesis in *Bacillus cereus* IBA1, **Bioresource Technology**, Volume 382, 129210, ISSN 0960-8524. 2023. <https://doi.org/10.1016/j.biortech.2023.129210>.

Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W e Clustal X versão 2.0. **Bioinformatics**, 23, 2947-2948. 2007.

Lemoigne, M. Dehydration and polymerization products of  $\beta$ -hydroxybutyric acid. *Touro.Soc. Chemical. Biol.* 8, 770–782. 1926.

Li, Z.; Loh, X. J. Recent advances in the use of polyhydroxyalkanoate-based nanovehicles as therapeutic delivery carriers. **WIREs Nanomed Nanobiotechnol**, 9: e1429. 2017. <https://doi.org/10.1002/wnan.1429>

Liu, J. M.; Liang, Y. T.; Wang, S. S.; Jin, N.; Sun, J.; Lu, C.; Sun, Y. F.; Li, S. Y.; Fan, B.; Wang, F. Z. Antimicrobial activity and comparative metabolomic analysis of *Priestia megaterium* strains derived from potato and dendrobium. *Sci Rep.* 31;13(1):5272. 2023. <https://doi.org/10.1038/s41598-023-32337->.

Madison, L. L.; Huisman, G. W. Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. **Microbiol.** 63 ,21–53.1999.  
<https://doi.org/10.1128/MMBR.63.1.21-53>.

Mahato, R. P.; Kumar, S.; Singh, P. Production of polyhydroxyalkanoates from renewable resources: a review on prospects, challenges and applications. **Arch Microbiol.** 205, 172. 2023. <https://doi.org/10.1007/s00203-023-03499-8>

Martins, K.; Anindya, M. A New Wave of Industrialization of PHA Biopolyesters. **Bioengineering.** 9, 74. 2022. <https://doi.org/10.3390/bioengineering9020074>.

Meereboer, K. W.; Misra, M.; Mohanty, A. K. Review of recent advances in the biodegradability of polyhydroxyalkanoate (PHA) bioplastics and their composites. **Green Chemistry,** 2, 5519. 2020. <https://doi.org/10.1039/D0GC01647K>.

Melchor Martínez, E. M.; Macias-Garbett, R.; Malacara, A.; Iqbal, H; Sosa-Hernández, J.; Parra, R.Environmental impact of emerging contaminants from battery waste: A mini review. **Case Studies in Chemical and Environmental Engineering.** 2021.  
<https://dx.doi.org/10.1016/j.cscee.2021.100104>.

Melo, R. N.; Souza, H. G.; Steffens, J. Recent updates on microbial production and recovery of polyhydroxyalkanoates. 3 **Biotechnology** 13, 204. 2023.  
<https://doi.org/10.1007/s13205-023-03633-9>.

Michaeli, W. et al. Tecnologia dos Plásticos. **Editora Edgard Blücher** Ltda., São Paulo. Introdução e Lição 1, p. 1 a 13. 1995.

Mitra, R.; Xiang, H.; Han, J. Current advances toward the production of polyhydroxyalkanoates containing 4-hydroxybutyrate for biomedical applications. **Molecules.** 26, 7244. 2021. <https://doi.org/10.3390/molecules26237244>.

Mohapatra, S.; Vishwakarma, K.; Joshi, N.C.; Maity, S.; Kumar, R.; Ramchander, M.; Pattnaik, S.; Samantaray, D. P. A review on PHAs: the future biopolymer. In **Environmental and Agricultural Microbiology.** 2021.  
<https://doi.org/10.1002/9781119525899.ch5>.

Jendrossek, D. Polyhydroxyalkanoate granules are complex subcellular organelles (carbonosomes). **J. Bacteriol.** 191 (10), 3195-3202. 2009.  
<https://doi.org/10.1128/jb.01723-08>.

Jiménez, M. M.; Palace, DA; Palencia, M.; Meléndrez, MF; Rivas, BL. Biobased Polymeric Membranes: Development and Environmental Applications. **Membranes**, 13, 625. 2023. <https://doi.org/10.3390/membranes1307062.5>

Mourão, I. A. C. L. História genealógica do conceito de homologia: uma análise filomemética. Sistemática, Taxonomia e Biodiversidade, Dissertação de Mestrado. **Museu de Zoologia**. 2016.

Muneer, F.; Rasul, I.; Azeem, F. Microbial Polyhydroxyalkanoates (PHAs): Efficient Replacement of Synthetic Polymers. **J. Polym Environ**. 28, 2301–2323. 2020. <https://doi.org/10.1007/s10924-020-01772-1>.

Nascimento, F. X.; Hernández, B. A. G.; Glick, B. R.; Rossib, M. J. Plant growth-promoting activities and genomic analysis of the stress-resistant *Bacillus megaterium* STB1, a bacterium of agricultural and biotechnological interest. **Biotechnology Reports**, v. 25, p. e00406, 2020.

Neoh, S. Z; Chek, M. Z; Hua, T. T.; Javier, A.; Ardra, N. Polyhydroxyalkanoate synthase (PhaC): The key enzyme for biopolyester synthesis, **Current Research in Biotechnology**, Volume 4, Pages 87-101, ISSN 2590-2628. 2022. <https://doi.org/10.1016/j.crbiot.2022.01.002>.

Pakalapati, H.; Chang, C. K.; Show, P. L.; Arumugasamy, S. K.; Lan, J. C. W. Development of polyhydroxyalkanoates production from waste feedstocks and applications. **Journal of Bioscience and Bioengineering**, Volume 126, Issue 3, Pages 282-292. 2018. <https://doi.org/10.1016/j.jbiosc.2018.03.016>.

Pearson, W. R. An introduction to sequence similarity ("homology") searching. **Curr Protoc Bioinformatics**. Chapter 3:3.1.1-3.1.8. 2013. <https://doi:10.1002/0471250953.bi0301s42>.

Prado, C. C. A. do. Degradação de fipronil por bactérias bioprospectadas de solo. 2017. 69 f. TCC (Graduação) – Curso de Biotecnologia, FCBA Universidade Federal da Grande Dourados, Dourados, 2017. Disponível em: <https://repositorio.ufgd.edu.br/jspui/bitstream/prefix/3244/1/CaioCesarAchilesPrado.pdf> . Acesso em: 20 nov. 2023.

Pramanik, Nilkamal. A tool for biomedical application: Synthesis and modification of polyhydroxyalkanoates. **Sustainable Chemistry and Pharmacy**. 32, 101041, ISSN 2352-5541. 2023. <https://doi.org/10.1016/j.scp.2023.101041>.

Pruitt, K. D.; Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. **Nucleic Acids Res.**;35(Database issue):D61-5. 2007. <https://doi:10.1093/nar/gkl842>.

Rahman, M. Z.; Khatun, M. M.; Hoque, M. R. Plastic waste to plastic value: Role of industrial biotechnology. *Biodegradability of Conventional Plastics*. 16, 339-360. 2023. <https://doi.org/10.1016/B978-0-323-89858-4.00016-6>.

Ramya, P.; Aanandhalakshmi, R.; Vanavil, B. Computational Analysis of Alginate Lyases Produced by Different Microorganisms. **Current Trends in Biotechnology and Pharmacy**, 17(3B), 1191–1207. 2023. <https://doi.org/10.5530/ctbp.2023.3s.56>.

Raza, Z. A.; Khalil, S.; Majeed, M. I.; Sarwar, T. Aminolysis of (hydroxybutyrate)-based multicomponent polyfilms for the impregnation of bovine serum albumin, **Polymer Bulletin** 80 (2) 2019–2043 2023. <https://doi.org/10.1007/s00289-022-04165-7>

Reto, M. A. S. Polietilenos Expandem Enquanto Demanda Retrai, **Plástico Moderno**, 26-40. 1998.

Reynolds, C. R.; Islam, S. A.; Sternberg, M. J. E. “EzMol: A web server wizard for the rapid visualisation and image production of protein and nucleic acid structures.” **J Mol Biol**. 2018.

Riaz, A.; Kanwal, F.; Borner, A.; Pillén, K.; Dai, F.; Alqudah, A. M. Advances in Genomics-Based Barley Breeding: Molecular Tools and Genomic Databases. **Agronomy**, 11, 894. 2021. <https://doi.org/10.3390/agronomy11050894>.

Rocha, C, R. F. B.;Garcia, A. O.; Santos, M. G.; Otto, P. I.; Silva, M.V.B.; Martins, M. F.; Machado, M. A.; Panetto, J. C. C.; Calus, M. P. L; Guimarães, S. E. F. Inheritance of genomic regions and genes associated with number of oocytes and embryos in Gir cattle through daughter design. **J. Dairy Sci**. 2023. <https://doi.org/10.3168/jds.2023-24111/>.

Rojo, S. G.; Antolínez, D. R. Production of polyhydroxyalkanoates as a feasible alternative for an integrated multiproduct lignocellulosic biorefinery. **Bioresource Technology**. Volume 386. 2023. <https://doi.org/10.1016/j.biortech.2023.129493>.

Romeiro, A. R. Civilização e Inovação: Porque a Revolução Industrial Foi um fenômeno Dependente de uma Trajetória Evolucionária Civilizacional. **Instituto de Economia** – UNICAMP. p 8-356. 2023. Disponível em:

<https://www.economia.unicamp.br/images/arquivos/ademar-romeiro-1-livro-versao-com-pleta-para-a-edusp.pdf> Acesso em: 02 jan. 2024.

Rodríguez, F. **Principles of Polymer Systems**, Taylor & Francis, Washington. Introduction, p. 14-19. 1996.

Rosenboom, J. G.; Langer, R.; Traverso, G. Bioplastics for a circular economy. **Nat Rev Mater** 7, 117–137. 2022. <https://doi.org/10.1038/s41578-021-00407-8>.

Saeed, M.; Ilyas, N.; Jayachandran, K.; Shabir, S.; Akhtar, N.; Shahzad, A.; Sayyed, R. Z. Bano, A. Advances in Biochar and PGPR engineering system for hydrocarbon degradation: A promising strategy for environmental remediation, *Environmental Pollution*, Volume 305, 119282, ISSN 0269-7491. 2022. <https://doi.org/10.1016/j.envpol.2022.119282>.

Salzano, F. M. The evolution of science in a Latin-American country: genetics and genomics in Brazil. *Genetics*, v. 208, n. 3, p. 823-832, 2018.

Samadhiya, A.; Agrawal, R.; Luthra, S.; Kumar, A.; Garza-Reyes, J. A. and Srivastava, D. K. (2022). Totale Industry 4.0 Productive Maintenance in a sustainability context: Exploring the mediating effect of the circular economy. **The International Journal of Logistics Management**, 34(3), 818–846.

Santos, F. R. F.; Santana, E. S.; ALMEIDA, D. Vaccine design by immunoinformatics against HCV: Chemical characterization and prediction of T-cell epitopes. **Research, Society and Development**, [S. l.], v. 10, n. 16, p. e55101622994, 2021. DOI: 10.33448/rsd-v10i16.22994. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/22994> . Acesso em: 10 nov. 2023.

Saratale, R. G.; Si-Kyung Cho, S. K.; Saratale, D. G.; Kadam, A. A.; Ghodake, G.S.; Kumar, M.; Bharagava, R. N.; Kumarf, G.; Kim, D.S.; Mulla, S. I.; Shin, H. S. A comprehensive overview and recent advances on polyhydroxyalkanoates (PHA) production using various organic waste streams. **Bioresource Technology**. 325, 124685. 2021. <https://doi.org/10.1016/j.biortech.2021.124685>.

Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. **Nucleic Acids Res.** 1;31(13):3381-5. 2003. <https://doi.org/10.1093/nar/31.13.3381>.

Seemann, T. Prokka: rapid prokaryotic genome annotation, **Bioinformatics**, Volume 30, Issue 14, Pages 2068–2069. 2014. <https://doi.org/10.1093/bioinformatics/btu153>.

Shaikh, S. Polyhydroxyalkanoates: Next generation natural biomolecules and a solution for the world's future economy. **International Journal of Biological Macromolecules**, v. 166, p. 297-321, 2020.

Shaikh, S.; Rashid, N.; McKay, G.; Mackey, H.R. Photobioreactor Design for Polyhydroxyalkanoate Production Using Anoxygenic Photoheterotrophs: A Review. **Fermentation**. 9, 778. 2023. <https://doi.org/10.3390/fermentation9080778>

Shikha, K.; Rao, A.; Kaur, N.; Dhania, D. Petroleum-Based Plastics Versus Bio-Based Plastics: A Review. **Nature Environment and Pollution Technology**. Vol. 22 pp.1111-1124. 2023. <https://doi.org/10.46488/NEPT.2023.v22i03.003>

Song, L.; Wang, M.; Yu, D.; Li, Y.; Yu, H.; Han, X. Enhancing Production of Medium-Chain-Length Polyhydroxyalkanoates from *Pseudomonas* sp. SG4502 by *tac* Enhancer Insertion. **Polymers**.15, 2290. 2023. <https://doi.org/10.3390/polym15102290>.

Steenwyk, J.; King, N. From genes to genomes: opportunities and challenges for synteny-based phylogenies. Preprints 2023, 2023090495. <https://doi.org/10.20944/preprints202309.0495.v1>.

Strong, A. B. Plastics – Materials and Processing, Prentice Hall, Columbus. Thermoplastic Materials. **Commodity Plastics**, p. 153-174. 5. 1996.

Tan, G. Y.; Chen, C. L.; Li, L.; Ge, L.; Wang, L.; Razaad, I.; Wang, J. Y. Start research on biopolymer polyhydroxyalkanoate (PHA): A Review. **Polymers**. 6(3), 706–754. 2022.

Tan, I. K. P.; Foong, C. P.; Tan, H. T.; Lim, H.; Zain, N. A. A.; Tan, Y. C.; Hoh, C. C.; Sudesh, K. Polyhydroxyalkanoate (PHA) synthase genes and PHA-associated gene clusters in *Pseudomonas* spp. and *Janthinobacterium* spp. isolated from Antarctica. **J. Biotechnol.** 313, 18–28. 2020. <https://doi.org/10.1016/j.jbiotec.2020.03.006>

Tarazona, N. A.; Hernandez-Arriaga, A. M.; Kniewel, R.; Prieto, M. A. Phasin interactome reveals the interplay of PhaF with the polyhydroxyalkanoate transcriptional regulatory protein PhaD in *Pseudomonas putida*. *Environ. Microbiol.* 22 (9), 3922–3936. 2020. <https://doi.org/10.1111/1462-2920.15175>.

Tennakoon, P.; Chandika, P.; Yi, M.; Jung, W. K. Marine-derived biopolymers as potential bioplastics, an eco-friendly alternative. **iScience Review**. 26, 106404, April 21, 2023. <https://doi.org/10.1016/j.isci.2023.106404>.

Torsten, S. Prokka: anotação rápida do genoma procariótico, **Bioinformatics**, Volume 30, Edição 14, páginas 2068–2069. julho de 2014. <https://doi.org/10.1093/bioinformatics/btu153>.

Tsuge, T. Fundamental factors determining the molecular weight of polyhydroxyalkanoate during biosynthesis. **Polym. J.** 48 (11), 1051–1057. 2016.

Umeda, F.; Kitano, Y.; Murakami, K.; Yagi, K.; Miura, Y.; Mizoguchi, T. Cloning and sequence analysis of the poly(3-hydroxyalkanoic acid)-synthesis genes of *Pseudomonas acidophila*. **Appl. Biochem. Biotechnol.** 70–72, 341–352. 1998.

Vicente, D.; Proença, DN; Morais, PV The Role of Bacterial Polyhydroxyalkanoate (PHA) in a Sustainable Future: A Review on Biological Diversity. **International J. Environment. Public Health Res.** 20, 2959. 2021 <https://doi.org/10.3390/ijerph20042959>.

Wan, J. H.; Ng, L. M.; Neoh, S. Z. Sequência completa do genoma de *Aquitalea pelogenes* USM4 (JCM19919), produtor de polihidroxialcanoato. **Arch Microbiol** 205, 66. 2023. <https://doi.org/10.1007/s00203-023-03406-1>.

Wang, R. C.; Wang, Z. Precision Medicine: Disease Subtyping and Tailored Treatment. **Cancers (Basel)**. 2023 Jul 28;15(15):3837. <https://doi.org/10.3390/cancers15153837>.

Wen, Q.; Liu, B.; Li, F.; Chen, Z. Substrate strategy optimization for polyhydroxyalkanoates producing culture enrichment from crude glycerol. **Bioresource Technology**, 311, 123516. 2020. <https://doi.org/10.1016/j.biortech.2020.123516>.

Wittenborn, E. C.; Jost, M.; Wei, Y.; Stubbe, J. C. L. Structure of the catalytic domain of the class I polyhydroxybutyrate synthase from *Cupriavidus necator*. **J. Biol. Chem.** 291 (48), 25264–25277. 2016. <https://doi.org/10.1074/jbc.M116.756833>.

Xiaolei, P.; Han, L.; Xinrui, Z.; Zhang, Q.; Lei, C.; Yue, Z.; Yin, L. Regulatory and structural mechanisms of PvrA-mediated regulation of the PQS quorum-sensing system and PHA biosynthesis in *Pseudomonas aeruginosa*. **Nucleic Acids Research**, Vol. 51, No. 6 2691–2708. 2023. <https://doi.org/10.1093/nar/gkad059>.

Xie, Y.; Gao, S.; Zhang, D.; Wang, C.; Chu, F. Bio-based polymeric materials synthesized from renewable resources: A mini-review, **Resources Chemicals and Materials**, Volume 2, Issue 3, Pages 223-230, ISSN 2772-4433. 2023. <https://doi.org/10.1016/j.recm.2023.05.001>.

Yin, Y.; Zhang, Q.; Peng, H., Retrospect and prospect of aerobic biodegradation of aniline: Overcome existing bottlenecks and follow future trends, **Journal of Environmental Management**, Volume 330, 117133, ISSN 0301-4797. 2023. <https://doi.org/10.1016/j.jenvman.2022.117133>.

Yu, S; Ali, J; Zhou, S.; Ren, G.; Xie, H.; Xu, J.; Yu, X.; Zhou, F.; Peng, S.; Ma, L.; Yuan, D.; Li, Z.; Chen, D.; Zheng, R.; Zhao, Z.; Chu, C.; You, A.; Yu, W.; Zhan, Q. From Green Super Rice to green agriculture: reaping the promise of functional genomics research. **Molecular Plant**, v. 15, n. 1, p. 9-26, 2022.

Zhang, Z.; Wang, Y.; Wang, X.; Zhang, Y.; Zhu, T.; Peng, L.; Xu, Z.; Chen, X.; Wang, D.; Ni, B.; Liu, Y. Towards scaling-up implementation of polyhydroxyalkanoate (PHA) production from activated sludge: Progress and challenges. **Journal of Cleaner Production**. 447, 141542, ISSN 0959-6526. 2024. <https://doi.org/10.1016/j.jclepro.2024.141542>.

Zhou, W., Bergsma, S.; Colpa, I. D.; Euverink, G.; Krooneman, J. Polyhydroxyalkanoates (PHAs) synthesis and degradation by microbes and applications towards a circular economy, **Journal of Environmental Management**, Volume 341, 118033, ISSN 0301-4797. 2023. <https://doi.org/10.1016/j.jenvman.2023.118033>.

Zou, H.; Shi, M.; Zhang, T.; Li, L.; Li, L.; Xian, M. Natural and engineered polyhydroxyalkanoate (PHA) synthase: key enzyme in biopolyester production. *Appl. Microbiol. Biotechnol.* 101 (20), 7417–7426. 2017. <https://doi.org/10.1007/s00253-017-8485-0>.

Zurron, A. C. B. P. *Biologia molecular e biotecnologia*. Londrina: Editora e Distribuidora **Educacional S.A**, 2016. 224 p.