

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**CAROLINE CAMPOS CORDEIRO**

**ANÁLISE DE PANGENOMA DE BACTÉRIAS DO GÊNERO *Bacillus* spp.  
RELACIONADAS COM BIOINSUMOS AGRÍCOLAS BRASILEIROS**

**DOIS VIZINHOS, PARANÁ**

**2023**

**CAROLINE CAMPOS CORDEIRO**

**ANÁLISE DE PANGENOMA DE BACTÉRIAS DO GÊNERO *Bacillus* spp.  
RELACIONADAS COM BIOINSUMOS AGRÍCOLAS BRASILEIROS**

**Pan-genome analysis of bacteria of genus *Bacillus* spp. related to Brazilian  
agricultural bio-inputs**

Projeto de Trabalho de conclusão de curso de Especialização apresentado como requisito para obtenção do título de Especialização em Biologia Molecular – Habilitação Bioinformática da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup> Juliana Morini Küpper Cardoso Perseguini

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup> Naiana Cristine Gabiatti

**DOIS VIZINHOS, PARANÁ**

**2023**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**CAROLINE CAMPOS CORDEIRO**

**ANÁLISE DE PANGENOMA DE BACTÉRIAS DO GÊNERO *Bacillus* spp.  
RELACIONADAS COM BIOINSUMOS AGRÍCOLAS BRASILEIROS**

Projeto de Trabalho de conclusão de curso de Especialização apresentado como requisito para obtenção do título de Especialização em Biologia Molecular – Habilitação Bioinformática da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup> Juliana Morini Küpper Cardoso Perseguini

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup> Naiana Cristine Gabiatti

Data de aprovação: 16 de Fevereiro de 2023

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Naiana Cristine Gabiatti  
Universidade Tecnológica Federal do Paraná

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Deborah Catharine de Assis Leite  
Universidade Tecnológica Federal do Paraná

---

Prof. Ms.Fábio Antônio Antonelo  
Universidade Tecnológica Federal do Paraná

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Nédia de Castilhos Ghisi  
Universidade Tecnológica Federal do Paraná

**DOIS VIZINHOS, PARANÁ  
2023**

## **AGRADECIMENTOS**

Agradeço a todos os professores do corpo docente do curso de Especialização em Biologia Molecular da Universidade Tecnológica Federal do Paraná, por todos os ensinamentos ao longo do curso.

Agradeço a minha orientadora Prof<sup>a</sup>. Dr<sup>a</sup> Juliana Morini Küpper Cardoso Perseguini, pelas orientações durante essa trajetória.

Agradeço aos meus pais por todo apoio durante toda a minha vida e nessa etapa.

Agradeço ao Laboratório de Análises Biológicas e Biologia Molecular (BioMol) da Universidade Tecnológica Federal do Paraná – Câmpus Dois Vizinhos, por permitir a execução de todas as análises da pesquisa.

Dedico este trabalho à minha família, por acreditar  
que posso sempre melhorar.

## RESUMO

O presente projeto dedica-se a realizar a análise de pangenoma de bactérias do gênero *Bacillus* em bioinsumos agrícolas brasileiros. Atualmente o Brasil conta com 104 bioinsumos registrados à base de *Bacillus*. Tem por objetivo analisar características semelhantes e padrões que distinguem entre si os *Bacillus* utilizados em bioinsumos na agricultura brasileira: *B. amyloliquefaciens*, *B. velezensis*, *B. subtilis*, *B. thuringiensis*, *B. licheniformis* e *B. pumilus*, que são espécies registradas no Brasil para a composição de produtos biológicos. Apesar da grande quantidade de sequências disponíveis desse gênero, há poucos estudos de análises de genoma central e genes específicos. Para análise e interpretação dos resultados, os genomas foram obtidos a partir de sequências depositadas na plataforma NCBI (*National Center for Biotechnology Information*) e, então, analisados quanto à formação de *clusters* e comparação de genes, além da anotação das proteínas. Essas análises foram realizadas por meio das ferramentas PROKKA, ROARY, OrthoVenn2 e PGAweb. Foi possível observar a predominância de *B. thuringiensis* nos resultados. *B. thuringiensis* destacou-se entre as demais, com a maior quantidade de genes acessórios e com a maior quantidade de *clusters* exclusivos de proteínas, além de ser a espécie com maior número de mudanças genéticas em termos filogenéticos.

**Palavras-chave:** Pangenômica; Genoma; *Bacillus*; Bioinsumos no Brasil.

## ABSTRACT

This project is dedicated to performing a pangenome analysis of bacteria of the genus *Bacillus* in Brazilian agricultural bioinputs. There are 104 bioinputs registered based on *Bacillus* spp currently in Brazil. It aims to analyze similar characteristics and patterns that distinguish between themselves the *Bacillus* used in bioinputs in Brazilian agriculture: *B. amyloliquefaciens*, *B. velezensis*, *B. subtilis*, *B. thuringiensis*, *B. licheniformis* and *B. pumilus*, which are species registered in Brazil for the composition of biological products.. Despite the large amount of available sequences of this genus, there are a few studies of core genome and specific genes analysis. For analysis and interpretation of results, genomes were obtained from sequences deposited on the NCBI platform (National Center for Biotechnology Information) and then analyzed for cluster formation and gene comparison, in addition to protein annotation. These analyzes were performed using the PROKKA, ROARY, OrthoVenn2 and PGAweb tools. It was possible to observe the predominance of *B. thurigiensis* in the results. *B. thurigiensis* stood out among the others, with the highest number of accessory genes and the highest number of unique protein clusters, in addition to being the species with the highest number of genetic changes in phylogenetic terms.

**Keywords:** Pangenomics; Genome; *Bacillus*; Bio-inputs in Brazil.

## LISTA DE FIGURAS

Figura 1 - Colonização da superfície da raiz por bactérias benéficas atua no equilíbrio biológico .....	17
Figura 2 - Mecanismo de estímulo local induzido por microrganismos acarreta em aumento da resposta de defesa e resistência sistêmica .....	17
Figura 3 - Mecanismo de ação de toxinas produzidas por <i>B. thuringiensis</i> contra insetos .....	18
Figura 4 - Microscopia de uma cepa de <i>Bacillus</i> .....	20
Figura 5 - Árvore filogenética das sequências analisadas .....	46
Figura 6 - Tamanho do pangenoma e padrão de crescimento .....	47
Figura 7- Novos <i>clusters</i> à medida que são adicionados genomas .....	48
Figura 8 - Genes compartilhados por conservação .....	50
Figura 9 - Diagrama de proteínas ortólogas por <i>overlap number</i> .....	52
Figura 10 - Diagrama de proteínas ortólogas por contagem de proteínas .....	53
Figura 11 - <i>Heat Map</i> de grupos parálogos .....	56
Figura 12 - Diagrama de Venn .....	58



## LISTA DE TABELAS

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero <i>Bacillus</i> como ingrediente ativo .....	23
Tabela 2 - Cepas de bactérias do gênero <i>Bacillus</i> selecionadas para análise.....	36
Tabela 3 - Descrição dos arquivos de <i>output</i> gerados pelo <i>Prokka</i> .....	38
Tabela 4 - Quantidade de arquivos de proteínas obtidos para cada genoma .....	42
Tabela 5 - Quantidade de genes presentes no pangenoma, agrupados em <i>core genes</i> , <i>soft core genes</i> , <i>shell genes</i> e <i>cloud genes</i> .....	43
Tabela 6 - <i>RefSeq</i> dos genomas analisados no <i>OrthoVenn</i> e sua nomenclatura correspondente para melhor compreensão dos resultados .....	51

## LISTA DE GRÁFICOS

Gráfico 1 - Quantidade de arquivos de proteínas dos genomas analisados .....41

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>1.1</b>	<b>Objetivos</b>	<b>14</b>
1.1.1	Objetivo Geral	14
1.1.2	Objetivos Específicos	14
<b>2</b>	<b>REVISÃO DE LITERATURA</b>	<b>16</b>
<b>2.1</b>	<b>Controle biológico</b>	<b>16</b>
<b>2.2</b>	<b>O que são bioinsumos</b>	<b>19</b>
<b>2.4</b>	<b>O genoma das bactérias do Gênero <i>Bacillus</i></b>	<b>32</b>
<b>2.5</b>	<b>O que é pangenoma</b>	<b>32</b>
<b>2.6</b>	<b>Como a comparação dos genomas de <i>Bacillus</i> pode auxiliar na área de bioinsumos</b>	<b>33</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>35</b>
<b>3.1</b>	<b>Instalação sistema operacional</b>	<b>35</b>
<b>3.2</b>	<b>Seleção de genomas</b>	<b>35</b>
<b>3.3</b>	<b>Reanotação dos genomas</b>	<b>38</b>
<b>3.4</b>	<b>Análise do pangenoma</b>	<b>39</b>
3.4.1	Identificação de genes ortólogos	39
3.4.2	Ordenação de genes compartilhados	39
3.4.3	<i>Cluster</i> de genes ortólogos	40
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>41</b>
<b>4.1</b>	<b>Análise geral dos arquivos</b>	<b>41</b>
<b>4.2</b>	<b>Ordenação de genes compartilhados (<i>Roary</i>)</b>	<b>43</b>
<b>4.3</b>	<b>Identificação de genes ortólogos (<i>PGAweb</i>)</b>	<b>46</b>
<b>4.4</b>	<b>Análise de proteínas ortólogas (<i>Orthovenn</i>)</b>	<b>50</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>59</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>60</b>
	<b>APÊNDICE A: ANEXO DE PIPELINES</b>	<b>65</b>
	<b>ANEXO A - Direitos autorais - LEI N. 9.610, DE 19 DE FEVEREIRO DE 1998</b>	<b>69</b>

## 1 INTRODUÇÃO

Os microrganismos podem atuar como produtos industriais em diversas áreas, como, por exemplo, na agricultura. Nessa área, as bactérias fixadoras de nitrogênio são exploradas, como as *Rhizobium*, *Bradyrhizobium* e *Azospirillum*, que infectam as raízes de plantas leguminosas, e o nitrogênio é, então, fixado por um processo simbiótico da planta e da bactéria. Esta relação simbiótica é a base para o desenvolvimento de inoculantes, insumo que contém alta concentração dessas bactérias. Ainda na agricultura, algumas espécies de bactérias do gênero *Bacillus* atuam como patógenos de insetos, ao produzir uma toxina letal a larvas, e assim controlam larvas de insetos que afetam plantações (TORTORA, FUNKE & CASE, 2017). Essa variedade de produtos que utilizam microrganismos como ingredientes ativos em suas formulações constitui os bioinsumos, ou produtos biológicos.

Devido a essas importantes aplicações na agricultura, os bioinsumos estão ganhando cada vez mais força frente ao setor agrícola, visto a necessidade ambiental de substituir produtos químicos por produtos biológicos, como os inoculantes e bioinseticidas. De acordo com o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o Brasil tem mais de 370 produtos biológicos registrados, sendo que mais de 60% desse valor representa produtos formulados a partir de microrganismos (CROPLIFE, 2022). Para produtos biológicos de controle de pragas (inseticida, nematocida, bactericida ou fungicida), há 7 espécies de bactérias do gênero *Bacillus* registrados no Brasil (Agrofit, 2022): *B. amyloliquefaciens*, *B. licheniformis*, *B. methylotrophicus*, *B. pumilus*, *B. subtilis*, *B. thuringiensis* e *B. velezensis*.

Para estudar e avaliar essas bactérias a nível molecular e de genoma, a fim de entender suas funções e gerar dados para o desenvolvimento de soluções de controle de qualidade específicas para cada espécie, a análise de pangenoma é essencial. Na biologia molecular e na bioinformática, a técnica de análise pangenoma fornece a comparação de genomas e o conteúdo genético de um grupo de microrganismos, o que permite estudar a diversidade genética de espécies e prever funções de genes (GUIMARÃES et al., 2015). O pangenoma abrange todo o repertório de genes disponíveis para serem estudados em cepas de uma determinada espécie ou grupo de espécies próximas. Esses genes estão divididos em genoma central (genes que estão presentes em todas as cepas), genes acessórios (genes compartilhados apenas

por um grupo dentre todas as cepas) e genes cepa-específicos (únicos) (CHEN et al., 2018).

Ao obter essa compreensão sobre os genes presentes em todas as cepas, assim como os genes acessórios e únicos, pode-se obter um maior entendimento a respeito de processos bioquímicos em que os organismos estão envolvidos (PAGE et al., 2015). Em pesquisas com bactérias, a análise pangenômica tem importância em muitos campos, como estudos de evolução de espécies, investigação de patogenicidade e mecanismos de resistência, desenvolvimento de vacinas, adaptabilidade em hospedeiros e ambientes, e estruturas de população (CHEN et al., 2018).

A comparação de genomas (análise de pangenoma) e seu estudo têm aplicação junto ao crescente número de produtos biológicos no contexto de que a formulação de um produto biológico pode conter mais de um microrganismo de um mesmo gênero, que por métodos microbiológicos convencionais não é possível diferenciá-los enquanto espécie. Uma análise pangenômica permite encontrar regiões específicas dos *Bacillus* utilizados em bioinsumos e um maior entendimento biológico das espécies, e com isso aplicar esse conhecimento em um controle molecular mais específico, especialmente em produtos que envolvem mais de uma espécie de um mesmo gênero de bactéria.

## 1.1 Objetivos

### 1.1.1 Objetivo Geral

Realizar análises de Pangenoma para os genomas de *Bacillus* presentes em bioinsumos da agricultura brasileira.

### 1.1.2 Objetivos Específicos

- Comparar *B. amyloliquefaciens*, *B. licheniformis*, *B. pumilus*, *B. subtilis*, *B. thuringiensis* e *B. velezensis* quanto aos *clusters* de genes e proteínas ortólogas

- Analisar predominância de genes acessórios entre as espécies

## 2 REVISÃO DE LITERATURA

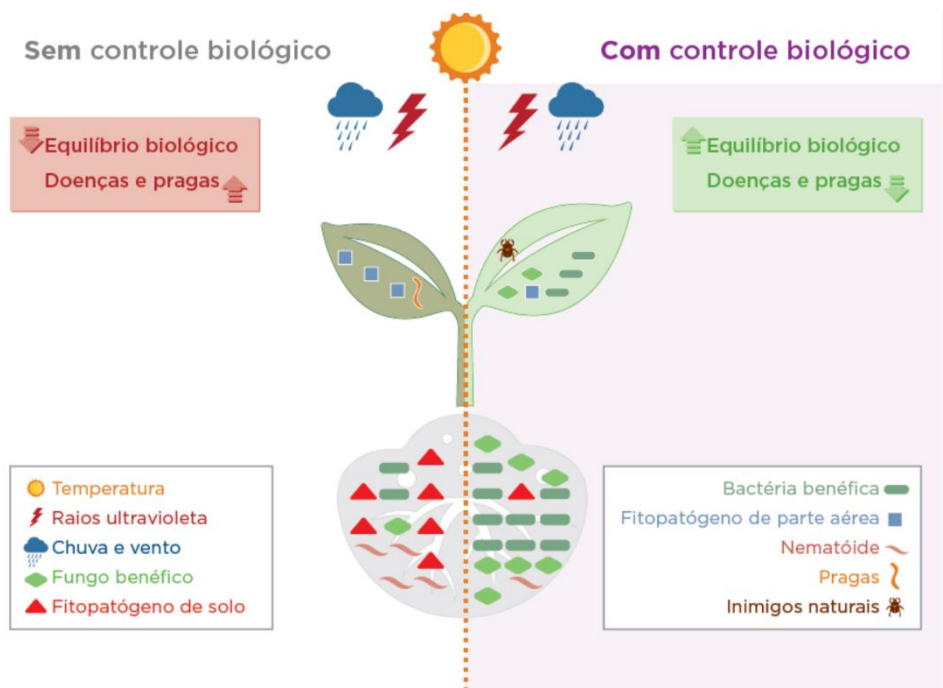
### 2.1 Controle biológico

O controle biológico é definido pela otimização de recursos naturais (como microrganismos benéficos) no desenvolvimento de ferramentas para erradicar e reduzir a ocorrência de fitopatógenos, geralmente insetos que provocam danos e dificultam o avanço da agricultura (CROPLIFE, 2023).

Algumas espécies de *Bacillus* atuam como patógenos de insetos, ao produzir uma toxina letal a larvas, e assim controlam larvas de insetos que afetam plantações. A liberação dessa toxina pelos microrganismos faz com que a multiplicação de pragas em plantas seja controlada e, até mesmo, extinta. Tal ação caracteriza o controle biológico exercido por determinados organismos (TORTORA, FUNKE & CASE, 2017).

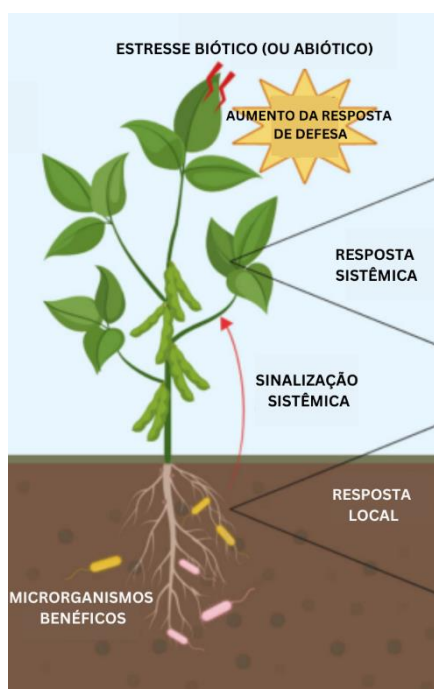
Os metabólitos tóxicos antimicrobianos produzidos por microrganismos agentes de controle biológico podem ser específicos contra um determinado grupo de microrganismos, como fungos, bactérias ou vírus, ou atuarem sobre diferentes grupos de organismos causadores de doenças em plantas. Os agentes de controle biológico podem atuar por diferentes mecanismos de ação: competição por sítios de penetração, por colonização da superfície da raiz (Figura 1), antagonismo direto por meio de metabólitos nocivos e indução de resistência sistêmica (Figura 2), ao desencadear resposta imune no metabolismo da planta (FONTES & VALADARIS-INGLIS, 2020).

Figura 1 - Colonização da superfície da raiz por bactérias benéficas atua no equilíbrio biológico



Fonte: CropLife, 2023

Figura 2 - Mecanismo de estímulo local induzido por microrganismos acarreta em aumento da resposta de defesa e resistência sistêmica

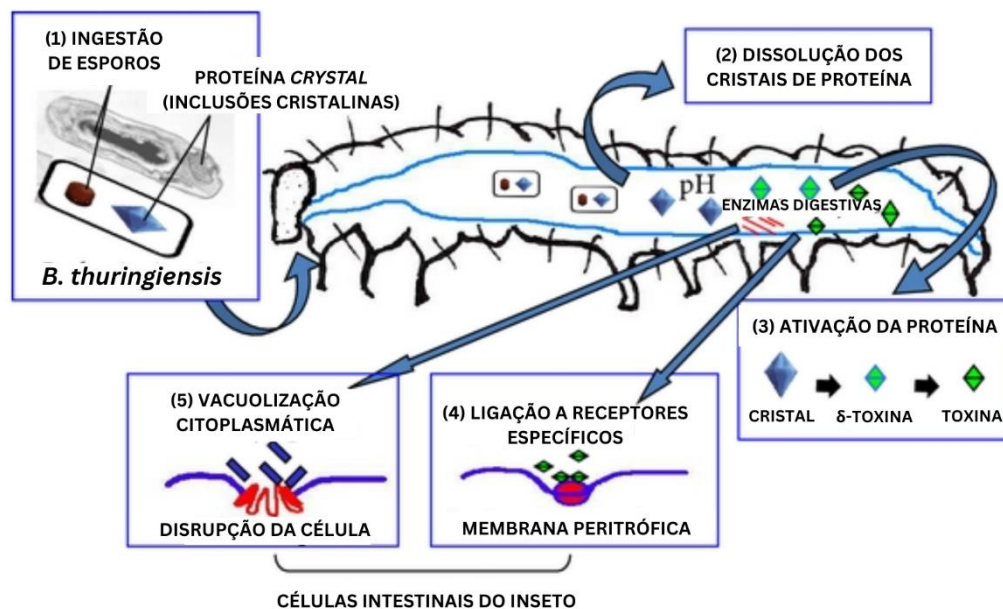


Fonte: Traduzido de Tonelli *et al*, 2020



Um exemplo de antagonismo direto por meio de metabólitos nocivos consiste nas toxinas produzidas por *B. thuringiensis* (Figura 3). O microrganismo que compõe um produto com ação inseticida, no caso *B. thuringiensis*, produz uma proteína durante a fase de esporulação na forma de inclusões cristalinas, denominada de  $\delta$ -endotoxina (proteína Cry). Essas proteínas Cry interagem diretamente com a membrana lipídica de células do intestino dos insetos. Quando o inseto ingere o produto que contém os cristais formados pelas proteínas Cry, essas proteínas são solubilizadas no intestino do inseto, o que resulta em toxinas ativas. As toxinas em sua forma ativa ligam-se com receptores presentes na parede intestinal, inserem-se na membrana, causam poros e conseqüente rompimento da parede celular. Esse processo ocasiona desequilíbrio osmótico e extravasamento do conteúdo intestinal para o aparelho circulatório, e permite a proliferação de esporos, o que resulta na septicemia severa e conseqüente morte do inseto (BRAVO *et al.*, 2007; DAS *et al.*, 2021; HECKEL, 2020).

**Figura 3 - Mecanismo de ação de toxinas produzidas por *B. thuringiensis* contra insetos**



Fonte: Traduzido de Schünemann *et al.*, 2014

## 2.2 O que são bioinsumos

Os produtos que utilizam microrganismos como ingredientes ativos em suas formulações constituem os bioinsumos, ou produtos biológicos, de maneira que os microrganismos podem atuar como produtos industriais em diversas áreas. Biopesticidas incluem moléculas que ocorrem naturalmente e apresentam a finalidade de controle de fitopatógenos (pesticidas bioquímicos), microrganismos que controlam patógenos (pesticidas microbiológicos), e até mesmo substâncias produzidas por plantas que apresentam ação pesticida (DAMALAS & KOUTROUBAS, 2018; EPA, 2022). Atualmente esses bioinsumos atraíram a atenção para o manejo de fitopatógenos e têm se tornado uma próspera alternativa aos pesticidas sintéticos (DAMALAS & KOUTROUBAS, 2018).

Bioinseticidas recebem muita atenção devido às suas vantagens a favor da sustentabilidade, pois são inerentemente menos tóxicos e agem somente sobre seus alvos específicos e organismos proximamente relacionados, em contrapartida aos pesticidas convencionais que apresentam um amplo espectro, afetando outros organismos que não são alvo do produto. Os bioinseticidas geralmente são eficazes em baixas quantidades e decompõem-se rapidamente, o que evita a contaminação do solo e do lençol freático (EPA, 2022).

Os pesticidas microbiológicos contêm microrganismos (bactérias, vírus de insetos, fungos, actinomicetos e protozoários) cuja função é agir como um agente de controle biológico, que afeta o fitopatógeno alvo de forma direta ou indireta, por meio de compostos que produzem. O microrganismo mais conhecido e mais amplamente utilizado é a bactéria *Bacillus thuringiensis*, que é comercializada há mais de 70 anos (MARRONE, 2019).

A implantação de bioinsumos não se restringe apenas ao campo de controle biológico. Muito é pesquisado e desenvolvido em relação à substituição de fertilizantes químicos por biofertilizantes e inoculantes. A Revolução Verde, que aconteceu especialmente nos anos de 1960 e 1980 no período após a Segunda Guerra Mundial, aumentou a produção agrícola, mas também exigiu uma grande quantidade de fertilizantes sintéticos. No início dos anos 1990, já havia regiões em desenvolvimento em que fertilizantes químicos não estavam disponíveis ou acessíveis. Além disso, o uso continuado de fertilizantes nitrogenados contribuiu substancialmente para a

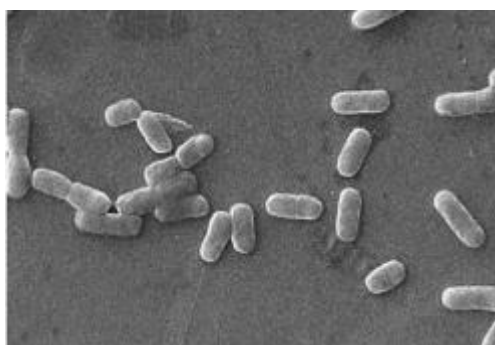
poluição do meio ambiente. Dessa forma, o processo de fixação biológica de nitrogênio oferece uma alternativa, com a comercialização de inoculantes formulados a partir de bactérias fixadoras de nitrogênio (BOHLOOL *et al.*, 1992; PEOPLES & CRASWELL, 1992).

### 2.3 As bactérias do Gênero *Bacillus* utilizadas em controle biológico de pragas e doenças

Um dos agentes de controle biológico mais conhecidos e utilizados no setor agrícola são as bactérias do gênero *Bacillus* (Crop Life, 2022).

As bactérias do gênero *Bacillus* apresentam formato de bastonete (Figura 4), são aeróbicas ou anaeróbicas facultativas, gram-positivas e formadoras de esporos. A maioria das espécies de bactérias do gênero *Bacillus* exibem uma ampla quantidade de habilidades fisiológicas que permitem com que sobrevivam em vários ambientes naturais. Os esporos são resistentes ao calor, frio, radiação e desinfetantes (BARON, 1996).

Figura 4 - Microscopia de uma cepa de *Bacillus*



Fonte: Filho *et al.*, 2018

Atualmente o Brasil conta com 7 espécies de bactérias do gênero *Bacillus* registradas no Ministério da Agricultura, Pecuária e Abastecimento (MAPA) (Agrofit, 2022). *B. amyloliquefaciens*, *B. licheniformis*, *B. methylotrophicus*, *B. pumilus*, *B. subtilis*, *B. thuringiensis* e *B. velezensis* são ingredientes ativos de produtos biológicos comercializados no Brasil. No total, há 104 produtos biológicos agrícolas registrados que contam com *Bacillus* em sua composição (dados obtidos a partir de busca no site Agrofit, do Ministério da Agricultura, Pecuária e Abastecimento, em 2022).

*B. amyloliquefaciens* pertence a um grupo de bactérias do solo de vida livre, que apresentam a característica de atuar como promotora do crescimento de plantas e de suprimir patógenos de plantas. Também é capaz de produzir substâncias antibacterianas e antifúngicas que protegem as plantas contra patógenos (NCBI, 2023). *B. amyloliquefaciens* está na classe de inseticida microbiológico, e está registrado na formulação de 34 produtos microbiológicos, dentre esses, 6 produtos apresentam uma composição mista com fungos do gênero *Trichoderma*: *T. harzianum* e *T. asperellum*. Há também o registro da cepa *B. amyloliquefaciens*, isolado CCT 7901 para apenas um produto.

*B. licheniformis* é uma bactéria formadora de endósporos, que pode ser isolada de amostras de solo e de plantas. Pode ser utilizada como um vetor de expressão para a produção de enzimas e apresenta potencial para novos estudos como ferramenta de manipulação genética (MURAS et al, 2021). *B. licheniformis* está na classe de nematicida microbiológico, e está registrado na composição de 6 produtos, todos esses em conjunto com *B. subtilis*, e alguns ainda contam com o fungo filamentosos *Paecilomyces lilacinus*.

*B. methylotrophicus* conta com apenas um registro, como ingrediente ativo de nematicida microbiológico em suspensão concentrada. Ainda nessa espécie, há o registro da cepa *B. methylotrophicus*, isolado UFPEDA20, também presente em apenas um nematicida.

*B. pumilus* é um microrganismo natural do solo. Esta bactéria coloniza a zona radicular de algumas plantas, em que inibe doenças fúngicas e oriundas de nematóides (NCBI, 2023). *B. pumilus* está presente em produtos microbiológicos da classe fungicida. Faz parte da formulação de 5 produtos registrados, dos quais 2 produtos apresentam uma formulação mista em conjunto com *B. subtilis* e *B. velezensis*.

*B. subtilis* corresponde a uma das primeiras bactérias estudadas. É uma bactéria do solo, não-patogênica, gram-positiva e em formato de bastonete, que produz esporos que são resistentes a fatores ambientais (PIGGOT, 2009). *B. subtilis* é um ingrediente ativo de produtos microbiológicos de 3 classes, bactericida microbiológico, fungicida microbiológico e nematicida microbiológico, presente na formulação de 23 produtos, dos quais 10 contam com outros microrganismos em sua composição. Ainda na espécie *B. subtilis*, há outras 2 cepas registradas, *B. subtilis*, isolado UFPEDA764 e *B. subtilis* linhagem QST 713.

*B. thuringiensis* é uma bactéria gram-positiva, que ocorre naturalmente no solo, caracterizada pela produção de proteínas cristais com atividade inseticida (toxinas *Cry*) durante a esporulação (WU, 2014). *B. thuringiensis* é outra espécie presente em uma grande quantidade de produtos, 32 no total, com a função de inseticida microbiológico. Há outra cepa registrada, *B. thuringiensis* var. *kurstaki*, isolado HD-1 (S1450), presente em 5 inseticidas.

Por fim, *B. velezensis*, registrado *B. velezensis*, isolado CNPSo 3602, é ingrediente ativo de 6 produtos microbiológicos das classes fungicida e nematicida, dos quais 4 deles são formulações mistas com outras espécies de *Bacillus*. *B. velezensis* é uma bactéria aeróbica, gram-positiva e formadora de esporos, com capacidade de promover o crescimento de plantas e supressão de patógenos, por meio da biossíntese de metabólitos secundários (RABBEE et al, 2019).

Todos os dados desse tópico foram obtidos por meio da busca no site Agrofit - Consulta aberta. A fim de melhor visualização dos produtos registrados no MAPA que contam com bactérias do gênero *Bacillus* em sua formulação, os dados foram compilados na Tabela 1.

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continua)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. amyloliquefaciens</i>	AgTecmmon	Massen Produtos Biologicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Amanzi	Agrobiológica Sustentabilidade S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Amitrix SC	Cooperativa Mista de Desenvolvimento do Agronegócio/COMDEAGRO	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	AmyloTrop	Cooperativa Mista de Desenvolvimento do Agronegócio/COMDEAGRO	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Amylo-X SL	Mitsui & Co (Brasil) S.A.	Concentrado Solúvel	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Aveo EZ	Sumitomo Chemical do Brasil Representações Ltda.	Suspensão Concentrada p/ Trat. Sementes	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Bactel	Dillon Biotecnologia Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i> + <i>T. asperellum</i> + <i>T. harzianum</i>	BF20.001	Ballagro Agro Tecnologia Ltda.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Biagro Proteção	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Boneville	Koppert do Brasil Holding Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Chevelle	Koppert do Brasil Holding Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Duravel	Basf S.A.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Eco-Shot	Iharabras S.A. Indústria Químicas	Grânulos Dispersíveis em Água	Inseticida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. amyloliquefaciens</i>	Faciens Protection	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Fx Protection	Bioma Indústria Comércio e Distribuição - EIRELI	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Inlayon	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Inlayon Eco	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Lumialza	Sumitomo Chemical do Brasil Representações Ltda.	Suspensão Concentrada p/ Trat. Sementes	Inseticida microbiológico
<i>B. amyloliquefaciens</i> + <i>T. harzianum</i>	Native	Agrivalle Brasil Industria e Comercio de Produtos Agrícolas S.A.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Nema-Attack	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Nemacontrol	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Nemacontrol Super	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Nema-Guard	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	No-Nema	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Inseticida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. amyloliquefaciens</i> + <i>T. asperellum</i> + <i>T. harzianum</i>	Pardella	Ballagro Agro Tecnologia Ltda.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Quartz SC	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Restrict	Massen Produtos Biologicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i> + <i>T. harzianum</i>	Shocker	Agrivalle Brasil Industria e Comercio de Produtos Agrícolas S.A.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i> + <i>T. asperellum</i> + <i>T. harzianum</i>	Tanus	Biota Innovations Industria e Comercio de Bioprodutos Ltda	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. amyloliquefaciens</i> + <i>T. harzianum</i>	Torpeno	Massen Produtos Biologicos S.A.	Pó Molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Trunemco	Sumitomo Chemical Brasil Indústria Química S.A. - Maracanaú/CE	Suspensão Concentrada p/ Trat. Sementes	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	TWIXX-A	Agrivalle Brasil Indústria e Comércio de Produtos Agrícolas S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Velez	Biota Innovations Industria e Comercio de Bioprodutos Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. amyloliquefaciens</i>	Veraneio	Koppert do Brasil Holding Ltda.	Pó molhável	Inseticida microbiológico
<i>B. amyloliquefaciens</i> , isolado CCT 7901 + <i>T. asperellum</i> isolado URM 8120 + <i>Trichoderma harzianum</i> , isolado URM 8119	Tricozak	Cooperativa Mista de Desenvolvimento do Agronegócio/COMDEAGRO	Grânulos Dispersíveis em Água	Inseticida microbiológico



Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. licheniformis</i> + <i>B. subtilis</i> + <i>P. lilacinus</i>	AgDommon	Massen Produtos Biologicos S.A.	Pó Molhável	Nematicida microbiológico
<i>B. licheniformis</i> + <i>B. subtilis</i> + <i>P. lilacinus</i>	Messenger	Massen Produtos Biologicos S.A.	Pó Molhável	Nematicida microbiológico
<i>B. licheniformis</i> + <i>B. subtilis</i>	Presence	FMC Química do Brasil Ltda. Agrivalle Brasil Indústria e Comércio de Produtos Agrícolas S.A.	Pó para Preparação de Pasta em Água	Nematicida microbiológico
<i>B. licheniformis</i> + <i>B. subtilis</i> + <i>P. lilacinus</i>	Profix	FMC Química do Brasil Ltda. Agrivalle Brasil Indústria e Comércio de Produtos Agrícolas S.A.	Pó Molhável	Nematicida microbiológico
<i>B. licheniformis</i> + <i>B. subtilis</i>	Quartzo	FMC Química do Brasil Ltda. Agrivalle Brasil Indústria e Comércio de Produtos Agrícolas S.A.	Pó para Preparação de Pasta em Água	Nematicida microbiológico
<i>B. licheniformis</i> + <i>B. subtilis</i> + <i>P. lilacinus</i>	Tramo®	FMC Química do Brasil Ltda. Agrivalle Brasil Indústria e Comércio de Produtos Agrícolas S.A.	Pó Molhável	Nematicida microbiológico
<i>B. methylotrophicus</i>	Onix	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. methylotrophicus</i> , isolado UFPEDA20*	Onix OG	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. pumilus</i> + <i>B. subtilis</i> + <i>Bacillus velezensis</i> , isolado CNPSo 3602	Bombardeiro	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Fungicida microbiológico
<i>B. pumilus</i>	BTP 005-19	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Fungicida microbiológico
<i>B. pumilus</i>	Caravan	Koppert do Brasil Holding Ltda.	Suspensão Concentrada	Fungicida microbiológico
<i>B. pumilus</i> + <i>B. subtilis</i> + <i>Bacillus velezensis</i> , isolado CNPSo 3602	Lastro	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Fungicida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. pumilus</i>	Sonata	Bayer S.A. - São Paulo/ SP	Suspensão Concentrada	Fungicida microbiológico
<i>B. subtilis</i> + <i>B. velezensis</i> , isolado CNPSO 3602	Ataplan	FMC Química do Brasil Ltda. Vittia Fertilizantes e Biológicos	Suspensão Concentrada	Fungicida microbiológico
	<i>B. subtilis</i>	Baci-Attack B S.A.	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i>	Baci-Guard	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i>	Baktillis	Mezfer BR Soluções Agrícolas Ltda	Suspensão Concentrada	actericida/fungicida microbiológico
<i>B. subtilis</i>	Biobac	UPL do Brasil Indústria e Comércio de Insumos Agropecuários S.A.	Pó Molhável	actericida/fungicida microbiológico
<i>B. subtilis</i>	Biobaci; Baci-Attack;	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i>	Bio-imune	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	actericida/fungicida microbiológico
<i>B. subtilis</i>	Furatrop	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	actericida/fungicida microbiológico
<i>B. subtilis</i>	Multi-Attack	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	actericida/fungicida microbiológico
<i>B. subtilis</i>	Multi-Guard	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Fungicida microbiológico
<i>B. subtilis</i>	Nematrop	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i>	Paladyo	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Nematicida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. subtilis</i>	Promobio	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i> + <i>B. velezensis</i> , isolado CNPSo 3602	Provilar	FMC Química do Brasil Ltda. - Campinas	Suspensão Concentrada	Fungicida microbiológico
<i>B. subtilis</i>	Rizos	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i> , isolado UFPEDA764*	Rizos OG	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Nematicida microbiológico
<i>B. subtilis</i> linhagem QST 713	Serenade	Bayer S.A. - São Paulo/ SP	Suspensão Concentrada	actericida/fungicida microbiológico
<i>B. thuringiensis</i>	Able	Mitsui & Co (Brasil) S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Acera	Ballagro Agro Tecnologia Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	Agree	Bio Controle - Métodos de Controle de Pragas Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Bac Control Max EC	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	Bac-Control Max WP	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Bac-Control WP	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Bettus Org	Nooa Ciencia e Tecnologia Agricola Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Biagro Cristalles	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. thuringiensis</i>	BioBac T	Vital Brasil Chemical Industria e Comercio de Produtos Quimicos Ltda – ME	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Biossela	Endeavour Biologicos Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	BI73.002/17	Ballagro Agro Tecnologia Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	BTControl	Simbiose Indústria e Comércio de Fertilizantes e Insumos Microbiológicos Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	BTFERT	Serquíbio Biotecnologia Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Btkill JCO	JCO Indústria e Comércio de Fertilizantes Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	BT-Turbo Max	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Costar	Mitsui & Co (Brasil) S.A.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. thuringiensis</i>	Crystal	Lallemand Soluções Agrobiológicas Ltda	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Dipel	Sumitomo Chemical do Brasil Representações Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Dipel ES-NT	Sumitomo Chemical do Brasil Representações Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Dipel WG	Sumitomo Chemical do Brasil Representações Ltda.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. thuringiensis</i>	Dipel WP	Sumitomo Chemical do Brasil Representações Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Helymax EC	Ballagro Agro Tecnologia Ltda.	Concentrado Emulsionável	Inseticida microbiológico

Tabela 1- Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(continuação)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. thuringiensis</i>	Javelin WG	Mitsui & Co (Brasil) S.A.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. thuringiensis</i>	Ponto Final	União Química Farmacêutica Nacional S.A	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Stregga EC	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	Sympatico	Sumitomo Chemical do Brasil Representações Ltda.	Dispersão de óleo ou Suspensão Concentrada em óleo	Inseticida microbiológico
<i>B. thuringiensis</i>	Tarik EC	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	Tarik WP	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Thuricide	Bio Controle - Métodos de Controle de Pragas Ltda.	Pó Molhável	Inseticida microbiológico
<i>B. thuringiensis</i>	Thuricide SC	Mitsui & Co (Brasil) S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i>	Winner Max EC	Vectorcontrol Industria e Comercio de Produtos Agropecuários Ltda.	Concentrado Emulsionável	Inseticida microbiológico
<i>B. thuringiensis</i>	Xentari	Sumitomo Chemical do Brasil Representações Ltda.	Grânulos Dispersíveis em Água	Inseticida microbiológico
<i>B. thuringiensis</i> var. <i>kurstaki</i> , isolado HD-1 (S1450)	Bacillus Thuringiensis Bom Futuro	Bom Futuro Agrícola Ltda.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i> var. <i>kurstaki</i> , isolado HD-1 (S1450)	Bacmix BTKSC	Cooperativa Mista de Desenvolvimento do Agronegócio/COMDEAGRO	Suspensão Concentrada	Inseticida microbiológico

Tabela 1 - Produtos biológicos comercializados no Brasil com bactérias do gênero *Bacillus* como ingrediente ativo

(conclusão)

Ingrediente ativo	Produto	Titular do registro	Formulação	Classe
<i>B. thuringiensis</i> var. <i>kurstaki</i> , isolado HD-1 (S1450)	BT Protection	Bioma Indústria Comércio e Distribuição	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i> var. <i>kurstaki</i> , isolado HD-1 (S1450)	BT-Turbo	Vittia Fertilizantes e Biológicos S.A.	Suspensão Concentrada	Inseticida microbiológico
<i>B. thuringiensis</i> var. <i>kurstaki</i> , isolado HD-1 (S1450)	Ecobaci T	Vital Brasil Chemical Industria e Comercio de Produtos Quimicos Ltda - ME	Suspensão Concentrada	Inseticida microbiológico
<i>B. velezensis</i> , isolado CNPSo 3602 (biológico)	Bionema	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	ngicida/Nematicida microbiológi
<i>B. velezensis</i> , isolado CNPSo 3602 (biológico)	Rudder	Biotrop Soluções Biológicas e Participações Ltda	Suspensão Concentrada	ngicida/Nematicida microbiológi

Fonte: Agrofit, 2022

## 2.4 O genoma das bactérias do Gênero *Bacillus*

Desde que a primeira cepa de *Bacillus*, *B. subtilis*, foi sequenciada em 1997, esse gênero de microrganismo tornou-se um dos mais estudados, e possui um dos maiores conjuntos de genomas sequenciados. Porém, esse dado ainda é conflitante com os estudos de análises de genoma central, e de genes cepa-específicos, que é limitado a poucos estudos. A maioria das publicações de *Bacillus* é focada em cepas patogênicas (KIM *et al.*, 2017).

## 2.5 O que é pangenoma

O termo pangenoma microbiano foi descrito pela primeira vez em 2005, para descrever a união de genes compartilhados por genomas de interesse (MEDINI *et al.*, 2005; PAGE *et al.*, 2015). O pangenoma refere-se ao conjunto de genes compartilhados por várias cepas de um determinado organismo, que inclui o genoma central (*core*) e o genoma acessório. O genoma central é definido pelos genes que existem e são compartilhados por todas as cepas, que geralmente desempenham funções essenciais para a espécie, são os genes que são responsáveis pelos aspectos básicos e pela maioria dos principais traços fenotípicos. O genoma acessório consiste nos genes que estão presentes em apenas algumas das cepas analisadas, mas desempenham papel importante ao contribuir para a diversidade das espécies e pela possibilidade de codificar vias bioquímicas suplementares, que não são essenciais à espécie, mas conferem vantagens seletivas, como resistência a antibióticos, adaptação em diferentes nichos e até colonização em um novo hospedeiro. Esses genes específicos estão, geralmente, agrupados em ilhas genômicas, flanqueadas por pequenas sequências de DNA repetidas e são caracterizados por um conteúdo anormal de bases guanina e timina (MEDINI *et al.*, 2005; KIM *et al.*, 2017).

Dentro do pangenoma, o número de genes centrais diminui conforme o número de sequências de genoma em uma espécie aumenta, enquanto o número de

sequências específicas aumenta junto com a quantidade de genomas (KIM *et al.*, 2017).

A análise de pangenoma fornece uma estrutura para determinar a diversidade genômica de um conjunto de dados de genomas de espécies de interesse (VERNIKOS *et al.*, 2015). A partir da aquisição de múltiplos dados de genoma da mesma ou de espécies relacionadas, a análise de pangenoma pode ser feita como um catálogo de todas as sequências e variantes da população (MARCUS *et al.*, 2014).

No estudo de genômica comparativa é fundamental realizar a análise de cluster de genes (grupos de genes) para a compreensão de funções e evolução dos organismos. Nos genomas bacterianos, os genes, na maioria dos casos, não são distribuídos randomicamente, mas sim organizados em grupos de *operons* (genes reguladores com funções relacionadas) (Yi *et al.*, 2007; OSBOURN & FIELD, 2009). Na compreensão dos genes presentes no genoma, genes homólogos são originados de um ancestral comum, de maneira que podem ou não ter a mesma função. Na categoria de genes homólogos, há a classificação de homólogos ortólogos e homólogos parálogos. Os genes ortólogos derivam-se de uma origem comum, que por um evento de especiação, cada cópia do gene foi destinada a espécies distintas. Esses genes possuem funções iguais e são importantes para o estudo da função de novos genes. Enquanto os parálogos são originados a partir de um evento de duplicação, em que um gene presente no genoma de determinada espécie é duplicado, porém as novas cópias ocupam posições diferentes dentro do mesmo genoma da espécie. Portanto, os genes parálogos possuem funções diferentes e são importantes no estudo de evolução (KOONIN, 2005).

## **2.6 Como a comparação dos genomas de *Bacillus* pode auxiliar na área de bioinsumos**

Ao comparar diferentes genomas de espécies relacionadas, é possível desenvolver sondas de qPCR (*Quantitative Polymerase Chain Reaction*) capazes de distinguir as espécies e as cepas analisadas. O *design* de um *primer* ou uma sonda, é um aspecto crítico para as análises de qPCR. Se os *primers* e as sondas forem desenhados incorretamente, a seletividade e a sensibilidade da detecção podem ser reduzidas, o que acarreta um resultado errado ou inespecífico. Por meio da análise



de pangenoma, ao identificar e selecionar genes cepa-específicos, torna-se possível desenvolver *primers* e sondas específicas para diferenciar as espécies estudadas (KWON *et al.*, 2021).

O desenvolvimento de produtos biológicos passa por quatro etapas principais: pesquisa de novos ativos biológicos, produção e formulação, avaliações no campo e, por fim, análise de segurança, aprovação e registro (Crop Life, 2022). Dentro da etapa de produção e formulação, faz-se necessário seguir um protocolo para controle de qualidade, e é nessa etapa que o desenvolvimento de *primers* e sondas específicas tem aplicabilidade. A detecção do ingrediente ativo de forma específica por qPCR permite assegurar a formulação, além de permitir diferenciar duas ou mais espécies relacionadas presentes no mesmo produto. Conforme a Tabela 1, alguns dos produtos comercializados no Brasil apresentam mais de um microrganismo do mesmo gênero (*Bacillus*) em sua composição, que poderiam ser detectados separadamente com alta taxa de especificidade com a utilização de sondas desenhadas a partir de sequências cepa-específicas encontradas na análise de pangenoma.

A investigação de genomas de *Bacillus* utilizados na produção de bioinsumos para a agricultura permite identificar propriedades presentes no genoma central e nos genes específicos, de forma que os genes de cada espécie podem apresentar diferenças significativas em suas funções e assinatura genômica (KIM *et al.*, 2017). Essas informações são essenciais para futuras pesquisas, aperfeiçoamento e desenvolvimento de novos produtos, posto que podem gerar ferramentas para a compreensão das funções dos genes.

### 3 MATERIAIS E MÉTODOS

#### 3.1 Instalação sistema operacional

Para a correta e completa execução das análises, foi necessária a instalação do sistema operacional *Linux*, por meio da habilitação do *Windows Power Shell* (WSL) e da instalação do sistema operacional de distribuição *Linux Ubuntu 20.04.4 LTS*.

No sistema operacional, foi instalado o gerenciador de ambientes computacionais virtuais ANACONDA 3-2021.11. A partir da instalação do ANACONDA, foi realizada a configuração do repositório (*channel*) BIOCONDA, que permite a instalação de pacotes de *software* usando o pacote de gerenciamento CONDA.

#### 3.2 Seleção de genomas

A fim de obter as sequências dos genomas de bactérias do gênero *Bacillus* utilizadas em produtos biológicos comercializados no Brasil como ingrediente ativo registrados no MAPA (Agrofit - consulta realizada em Abril de 2022), buscou-se no banco de dados do NCBI as sequências depositadas que possuíam o genoma completamente sequenciado. Na busca por espécie no campo “*Genome*”, ao selecionar “*Genome Assembly and Annotation report*”, foram obtidas as seguintes quantidades (consulta realizada em Abril de 2022):

- 1) *B. amyloliquefaciens* - 1 genoma de referência (RefSeq NZ\_CP072120.1) e 62 genomas completos;
- 2) *B. licheniformis* - 1 genoma de referência (RefSeq NZ\_CP014842.1) e 42 genomas completos;
- 3) *B. pumilus* - 1 genoma de referência (RefSeq GCF\_003020795.1) e 15 genomas completos;
- 4) *B. subtilis* - 1 genoma de referência (RefSeq NC\_000964.3) e 196 genomas completos;
- 5) *B. thuringiensis* - 1 genoma de referência (RefSeq NZ\_CM000753.1) e 74 genomas completos;

6) *B. velezensis* - 1 genoma de referência (RefSeq NZ\_CP009679.1) e 206 genomas completos.

*B. methylotrophicus* corresponde a *B. velezensis*, portanto essa espécie será retratada e agrupada como *B. velezensis* no presente estudo.

Em vista da capacidade da máquina (Intel(R) Core(TM) i5-2537M CPU @ 1.40GHz 1.40 GHz, Sistema operacional de 64 bits, processador baseado em x64) em operar com um grande número de sequências, foi definido que seriam utilizadas 5 sequências (escolhidas de maneira aleatória dentre as que possuíam genoma completo em “*Assembly Level*”) de cada espécie para compor as análises e adequar-se a um conteúdo representativo. Dessa forma, foram selecionados os genomas descritos na Tabela 2 para seguimento das análises.

**Tabela 2 - Cepas de bactérias do gênero *Bacillus* selecionadas para análise**

(continua)

<b>Identificação cromossomo (RefSeq)</b>	<b>Organismo</b>	<b>Cepa</b>	<b>Bioproject</b>
CP072120	<i>B. amyloliquefaciens</i>	GKT04	PRJNA716416
CP038028	<i>B. amyloliquefaciens</i>	FS1092	PRJNA243331
CP089530	<i>B. amyloliquefaciens</i>	LS1-002-014s	PRJNA787446
CP044132	<i>B. amyloliquefaciens</i>	ZKY01	PRJNA569143
CP071970	<i>B. amyloliquefaciens</i>	XJ5	PRJNA715331
CP090312	<i>B. licheniformis</i>	NWMCC0046	PRJNA792462
CP049330	<i>B. licheniformis</i>	CP6	PRJNA608985
CP035405	<i>B. licheniformis</i>	SRCM103608	PRJNA516719
CP041154	<i>B. licheniformis</i>	CSL2	PRJNA548835
CP065029	<i>B. licheniformis</i>	LCDD6	PRJNA661689

Tabela 2 - Cepas de bactérias do gênero *Bacillus* selecionadas para análise

(conclusão)

Identificação cromossomo (RefSeq)	Organismo	Cepa	Bioproject
LT906438	<i>B. pumilus</i>	NCTC10337	PRJEB6403
CP027116	<i>B. pumilus</i>	145	PRJNA377620
CP010997	<i>B. pumilus</i>	SH-B11	PRJNA276290
CP058951	<i>B. pumilus</i>	UAMX	PRJNA645214
CP085037	<i>B. pumilus</i>	BIM B-171	PRJNA770178
CP020102	<i>B. subtilis</i>	NCIB 3610	PRJNA377766
CP068982	<i>B. subtilis</i>	pb2441	PRJNA694795
CP016767	<i>B. subtilis</i>	CW14	PRJNA330772
CP028202	<i>B. subtilis</i>	SRCM102754	PRJNA438180
CP104097	<i>B. subtilis</i>	GL-4	PRJNA877089
CP106947	<i>B. thuringiensis</i>	4F5	PRJNA884171
CP039721	<i>B. thuringiensis</i>	BT-59	PRJNA534189
CP045030	<i>B. thuringiensis</i>	JW-1	PRJNA574201
CP083134	<i>B. thuringiensis</i>	BMP144	PRJNA757250
CP040782	<i>B. thuringiensis</i>	HM-311	PRJNA545808
CP097467	<i>B. velezensis</i>	HMB26553	PRJNA839221
CP082262	<i>B. velezensis</i>	BIM B-454D	PRJNA753960
CP103856	<i>B. velezensis</i>	SRCM116265	PRJNA872352
CP059405	<i>B. velezensis</i>	MV2	PRJNA647609
CP076290	<i>B. velezensis</i>	Q12	PRJNA734319

Fonte: NCBI, 2022

Para ser possível o *download* da base de dados do NCBI, foi instalada a ferramenta *Entrez Direct*, que permite acesso ao conjunto de banco de dados interconectados do NCBI (*Entrez Programming Utilities Help*, NCBI, 2023). Por meio do Sistema operacional Ubuntu.20.04.4 LTS, foi feito o *download* dessas sequências em formato FNA, FAA e GBK. Os arquivos FNA indicam arquivos em formato FASTA de *contigs* de entrada originais (nucleotídeos), os arquivos FAA indicam arquivos FASTA de genes codificadores traduzidos (proteínas), e arquivos GBK indicam arquivos do *GenBank* contendo sequências e anotações (SEEMANN, 2014).

### 3.3 Reanotação dos genomas

*Prokka* (versão 1.14.6) é um *software* que coordena um conjunto de ferramentas de *software* existentes para obter uma anotação ampla e confiável de sequências bacterianas genômicas. A anotação consiste no processo de identificação e marcação de características importantes em uma sequência de genoma (SEEMANN, 2014). Portanto, no sistema operacional Ubuntu.20.04.4 LTS, os genomas foram submetidos ao *software Prokka* para a predição dos genes. O *software Prokka* gera *outputs* (saídas) nos formatos: err, faa, ffn, fna, fsa, gbk, gff, log, sqn, tbl, txt, descritos na Tabela 3.

**Tabela 3 - Descrição dos arquivos de *output* gerados pelo *Prokka***

Sufixo	Descrição do conteúdo dos arquivos
.faa	Arquivo FASTA de genes codificadores traduzidos (proteínas)
.ffn	Arquivo FASTA de características genômicas (nucleotídeos)
.fna	Arquivo FASTA de <i>contigs</i> de entrada originais (nucleotídeos)
.fsa	Sequências de <i>contigs</i> para submissão (nucleotídeos)
.gbk	Arquivo dos <i>GenBank</i> contendo sequências e anotações
.gff	Arquivo contendo sequências e anotações
.log	Arquivo de saída de processamento do <i>Prokka</i>
.sqn	Arquivo editável para submissão
.tbl	Tabela de recursos para submissão
.txt	Sumário de estatísticas das anotações

Fonte: Seemann, 2014

### 3.4 Análise do pangenoma

#### 3.4.1 Identificação de genes ortólogos

*Pan-Genome Analysis Web Server – PGAweb* é uma ferramenta *online* gratuita que integra os *softwares* PGAP, PanGP e PGAP-X. Essa ferramenta permite a análise de dados de sequências inseridas, fornece diversos métodos para identificação de genes ortólogos, e apresenta características estruturais e dinâmicas do genoma. A organização da ferramenta é feita em 3 passos principais: 1) inserção de dados; 2) configuração de parâmetros; 3) análise de dados e visualização dos resultados, em que a ferramenta fornece um arquivo de texto correspondente a um gráfico (CHEN *et al.*, 2018).

Para os genomas serem submetidos à análise da ferramenta *PGAweb*, os arquivos em formato GBK tiveram seus formatos alterados para GB, formato de entrada aceito pela ferramenta. No site <http://pgaweb.vlcc.cn/analyze>, os arquivos em formato GB foram inseridos, definiu-se cobertura (*coverage*): 0,5 e identidade (*identity*): 0,8, e a função *Pan-genome analysis* foi selecionada.

#### 3.4.2 Ordenação de genes compartilhados

*Roary* (versão 3.13.0) é uma ferramenta que constrói o pangenoma de procariotos em larga escala a partir de uma ampla coleção de genomas bacterianos. Um gráfico das relações dos *clusters* é construído com base na ordem de ocorrência das sequências de entrada, o que permite que os *clusters* sejam ordenados e forneçam um contexto para cada gene. Os isolados são agrupados com base na presença de conjuntos de genes compartilhados apenas entre alguns genomas (PAGE *et al.*, 2015).

Os arquivos obtidos em formato GFF pelo *software Prokka* foram submetidos à análise com o *Roary*.

*ETE Toolkit* é uma ferramenta *online* para visualização de árvore filogenética (formato *Newick*) que permite o alinhamento de múltiplas sequências. Os dados em formato *Newick* são obtidos a partir das análises da ferramenta *Roary*.

### 3.4.3 *Cluster* de genes ortólogos

*OrthoVenn2* é um servidor *online* para a comparação de genomas inteiros e anotação de *clusters* ortólogos ao longo de múltiplas espécies, que permite a visualização de ocorrência de padrões. Os resultados são mostrados em formato de tabela de *clusters* para múltiplas espécies associados a diagramas de Venn (XU *et al.*, 2019).

O formato dos arquivos em formato FAA obtidos como *output* do *Prokka* foi alterado para formato FASTA, pois o *OrthoVenn2* aceita arquivos de entrada de proteína em formato FASTA para sua execução. Na página <https://orthovenn2.bioinfotoolkits.net/task/create>, os arquivos de proteína em formato FASTA foram submetidos à análise, com *e-value*: 1e-2 e *Inflation value*: 1,5.

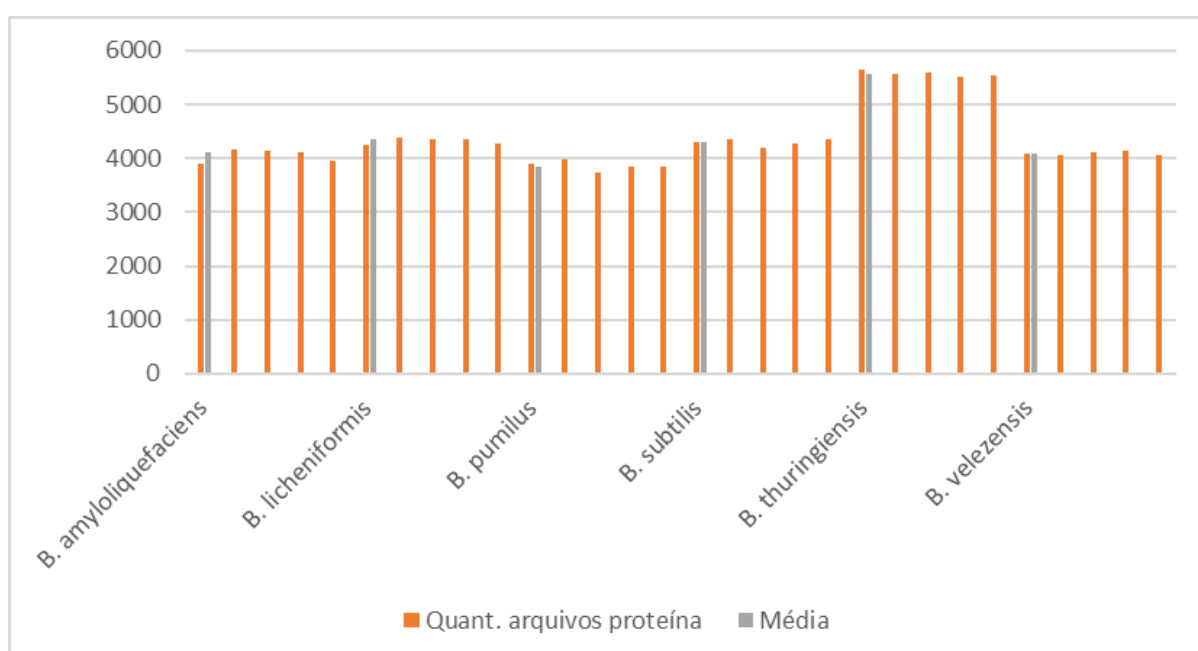
## 4 RESULTADOS E DISCUSSÃO

### 4.1 Análise geral dos arquivos

Para essa análise do pangenoma, 6 espécies de bactérias do gênero *Bacillus* foram analisadas para verificação da estrutura e funções de seus genomas. Um total de 30 genomas (Tabela 2), 5 representantes de cada espécie (escolhidos aleatoriamente) foram utilizados para determinar características de genoma *core* e genes específicos. Somente foram incluídos genomas de cromossomos marcados como completos na plataforma NCBI, para evitar dados incorretos oriundos de genomas incompletos que não possuem todos os conjuntos de genes, e consequentemente não permitiriam a análise de comparação dos genomas.

A partir do *download* das sequências nos formatos FNA, FAA e GBK, obteve-se as quantidades de proteínas encontradas nos arquivos FAA (arquivo FASTA de genes codificadores traduzidos), pois os arquivos em formato FAA contêm as sequências de proteínas dos genes. A fim de verificar a ocorrência e a amplitude das sequências trabalhadas em termos de quantidade de proteínas, o Gráfico 1 foi elaborado para visualização da quantidade de proteínas de cada cepa, e os números absolutos estão expressos na Tabela 4.

**Gráfico 1 - Quantidade de arquivos de proteínas dos genomas analisados**



Fonte: A autora, 2022



Tabela 4 - Quantidade de arquivos de proteínas obtidos para cada genoma

(continua)

Organismo	RefSeq	Quant. de proteína
<i>B. amyloliquefaciens</i>	CP072120	3901
<i>B. amyloliquefaciens</i>	CP038028	4158
<i>B. amyloliquefaciens</i>	CP089530	4148
<i>B. amyloliquefaciens</i>	CP044132	4103
<i>B. amyloliquefaciens</i>	CP071970	3959
<i>B. licheniformis</i>	CP090312	4255
<i>B. licheniformis</i>	CP049330	4385
<i>B. licheniformis</i>	CP035405	4359
<i>B. licheniformis</i>	CP041154	4356
<i>B. licheniformis</i>	CP065029	4287
<i>B. pumilus</i>	LT906438	3905
<i>B. pumilus</i>	CP027116	3968
<i>B. pumilus</i>	CP010997	3742
<i>B. pumilus</i>	CP058951	3850
<i>B. pumilus</i>	CP085037	3852
<i>B. subtilis</i>	CP020102	4296
<i>B. subtilis</i>	CP068982	4368
<i>B. subtilis</i>	CP016767	4195
<i>B. subtilis</i>	CP028202	4277
<i>B. subtilis</i>	CP104097	4343
<i>B. thuringiensis</i>	CP106947	5637
<i>B. thuringiensis</i>	CP039721	5580
<i>B. thuringiensis</i>	CP045030	5584
<i>B. thuringiensis</i>	CP083134	5512
<i>B. thuringiensis</i>	CP040782	5538
<i>B. velezensis</i>	CP097467	4092
<i>B. velezensis</i>	CP082262	4067
<i>B. velezensis</i>	CP103856	4102

Tabela 4 - Quantidade de arquivos de proteínas obtidos para cada genoma

Organismo	RefSeq	Quant. de proteína
<i>B. velezensis</i>	CP059405	4135
<i>B. velezensis</i>	CP076290	4062

Fonte: A autora, 2022

## 4.2 Ordenação de genes compartilhados (*Roary*)

Os arquivos de entrada na ferramenta *Roary* são arquivos em formato GFF3, provenientes da anotação de todos os genomas usando o *software Prokka*. As regiões codificadoras do genoma são extraídas a partir do genoma e convertidas em sequências de proteínas, que são filtradas para remover sequências parciais e previamente agrupadas em *clusters*. Isso resulta em um conjunto reduzido de sequências de proteínas, ao remover redundância dos dados (FU *et al.*, 2012; PAGE *et al.*, 2015). A análise de comparação dessas sequências reduzidas realizadas com BLASTP (*Basic Local Alignment Search Tool of Proteins*) por meio da ferramenta *Roary*, com percentual de identidade das sequências de 40%, está apresentada na Tabela 5. Para melhor visualização dos resultados, o código inicial formado por letras das sequências *RefSeq* foram substituídos pelas iniciais de suas respectivas espécies (*B. amyloliquefaciens* - BA, *B. licheniformis* - BL, *B. pumilus* - BP, *B. subtilis* - BS, *B. thuringiensis* - BT e *B. velezensis* - BV).

Tabela 5 - Quantidade de genes presentes no pangenoma, agrupados em *core genes*, *soft core genes*, *shell genes* e *cloud genes*

Classificação	Presença nos genomas	Quantidade de genes
<i>Core genes</i>	100% <= cepas <= 100%	1428
<i>Soft core genes</i>	95% <= cepas < 100%	30
<i>Shell genes</i>	15% <= cepas < 95%	7302

**Tabela 5 - Quantidade de genes presentes no pangenoma, agrupados em *core genes*, *soft core genes*, *shell genes* e *cloud genes***

(conclusão)

Classificação	Presença nos genomas	Quantidade de genes
<i>Cloud genes</i>	0% <= cepas < 15%	6895
Total genes	0% <= cepas <= 100%	15655

**Fonte: Roary, 2023**

O genoma *core* (*core genes*) é definido pelos genes presentes em todos os genomas analisados, geralmente relacionados a *housekeeping genes*, que são genes que sintetizam proteínas, essência para o funcionamento e manutenção do metabolismo do organismo vivo. Os *soft core* genes correspondem aos genes presentes em 95% à 100% dos genomas, ou seja, ao ampliar a faixa de genes, há um adicional de 30 genes que estão presentes na grande maioria dos genomas, porém não estão presentes em todos (PAGE *et al*, 2015).

Ao analisar os genes acessórios (*shell* e *cloud genes*), a quantidade de genes codificadores aumenta. Os *shell genes*, genes compartilhados por 15% à 95% dos genomas, contêm 7302 genes, enquanto os *cloud genes*, presentes em um número mais restrito de genomas (até 15%) correspondem a 9895 genes.

O total de genes encontrados, correspondentes ao genoma *core* e ao genoma acessório, somam 15655 genes codificadores de proteínas nas 6 espécies de *Bacillus* analisadas.

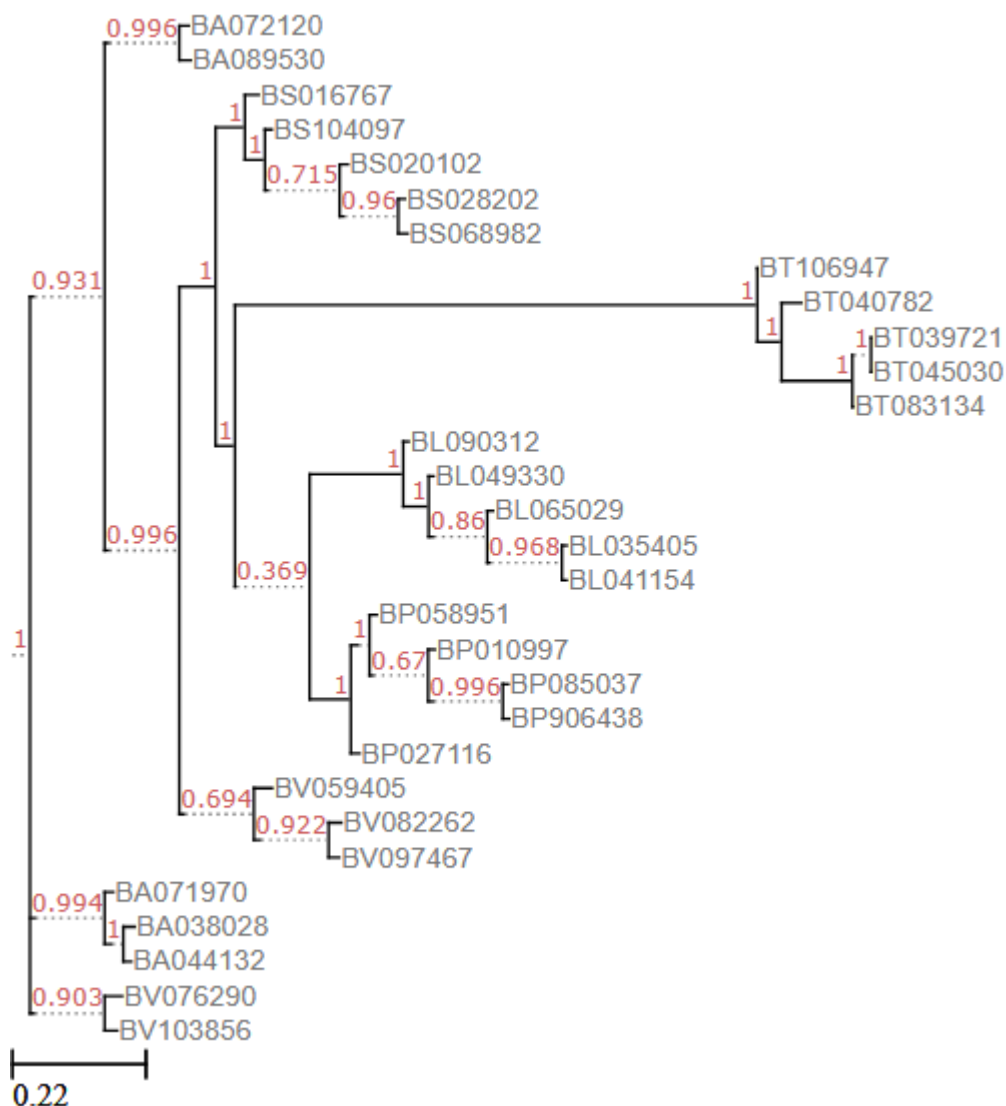
A planilha de presença e ausência de genes gerada pelo *Roary* apresenta genes compartilhados por todos (genoma *core*) correspondentes a características importantes para o desempenho das bactérias *Bacillus* contra fitopatógenos. Por exemplo, os genes *GroI* (anotação correspondente a *chaperonin* GroEL) e *tuf* (anotação correspondente a *elongation factor* Tu), ambos presentes no genoma *core*, desempenham importante papel ao facilitar a invasão das bactérias por meio de mobilidade e adesão (WANG *et al*, 2022).

Na árvore filogenética obtida pelos dados de *Roary* e da ferramenta *ETEToolkit* (Figura 5), quanto maior o ramo horizontal, maiores as quantidades de mudanças evolucionárias ocorridas. Devido ao maior comprimento do ramo que contém o nó interno das cepas analisadas da espécie *B. thurigiensis*, essa espécie apresenta o

maior número de mudanças genéticas, ou seja, mudanças ou substituições de nucleotídeos ao longo da sequência (cerca de quatro vezes o tamanho da barra inferior equivalente a 0,22). Ainda em relação ao *cluster* de *B. thurigiensis*, os números ao lado de cada nó, que representam a medida de suporte, são todos iguais a 1, o que corresponde a uma maior significância dos ramos que partem daquele nó.

*B. licheniformis* e *B. pumilus* possuem o mesmo ancestral comum, mas *B. licheniformis* apresenta maiores mudanças genéticas devido ao maior comprimento de seu ramo horizontal. E ao analisar o ancestral comum, essas duas espécies são as mais próximas a *B. thurigiensis* em termos evolucionários genéticos. As cepas analisadas das espécies *B. amyloliquefaciens* e *Bacillus velezensis* não estão todas ligadas a partir do mesmo nó interno, portanto, possivelmente, são as espécies que possuem maiores diferenças em sequências de nucleotídeos entre as cepas de uma mesma espécie.

Figura 5 - Árvore filogenética das sequências analisadas

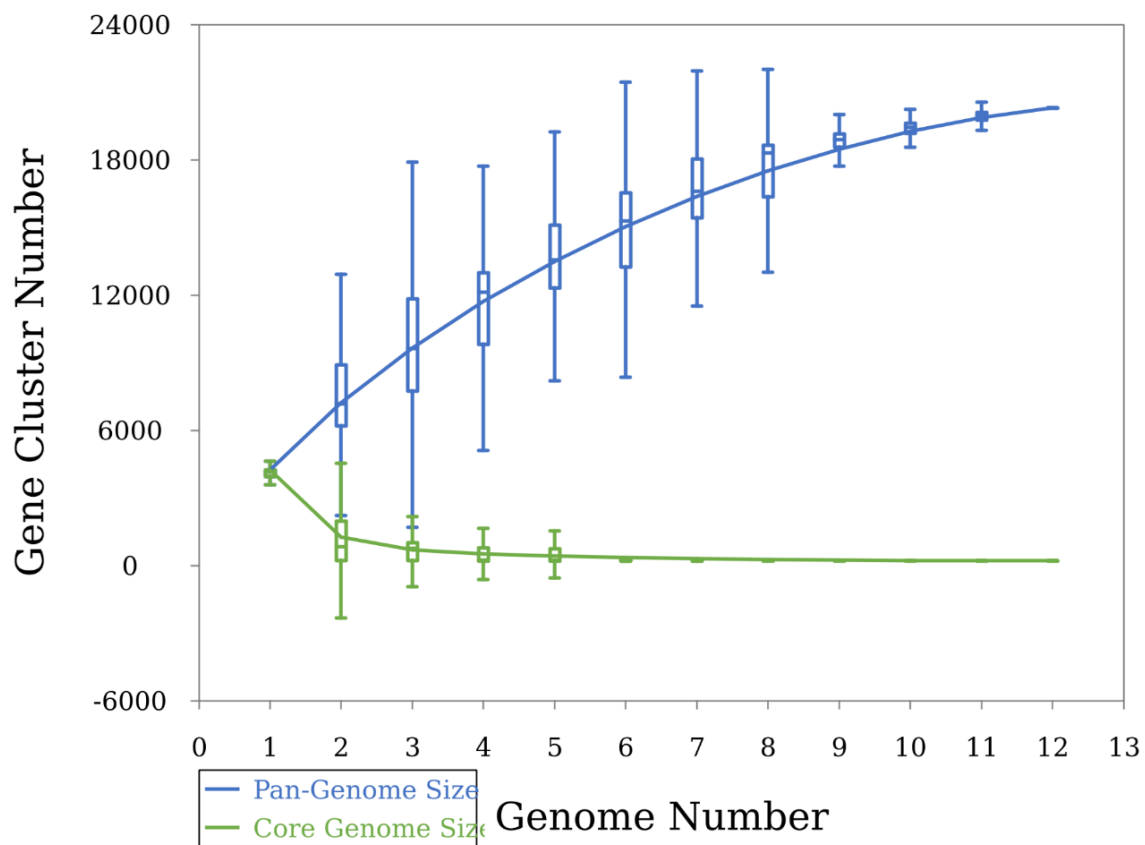


Fonte: ETEToolkit, 2023

### 4.3 Identificação de genes ortólogos (*PGAweb*)

Conforme o número de genomas analisados aumenta, diminui o número de genes pertencentes ao genoma *core*. O tamanho do pangenoma, ou seja, o conjunto de todos os genes segue a tendência contrária conforme o número de genomas analisados aumenta. O pangenoma acompanha o crescimento, pois a maior quantidade de genomas acarreta no aumento da soma da quantidade de genes específicos que cada um carrega (CHEN et al, 2018). Esse padrão de crescimento para as espécies analisadas está representado na Figura 6.

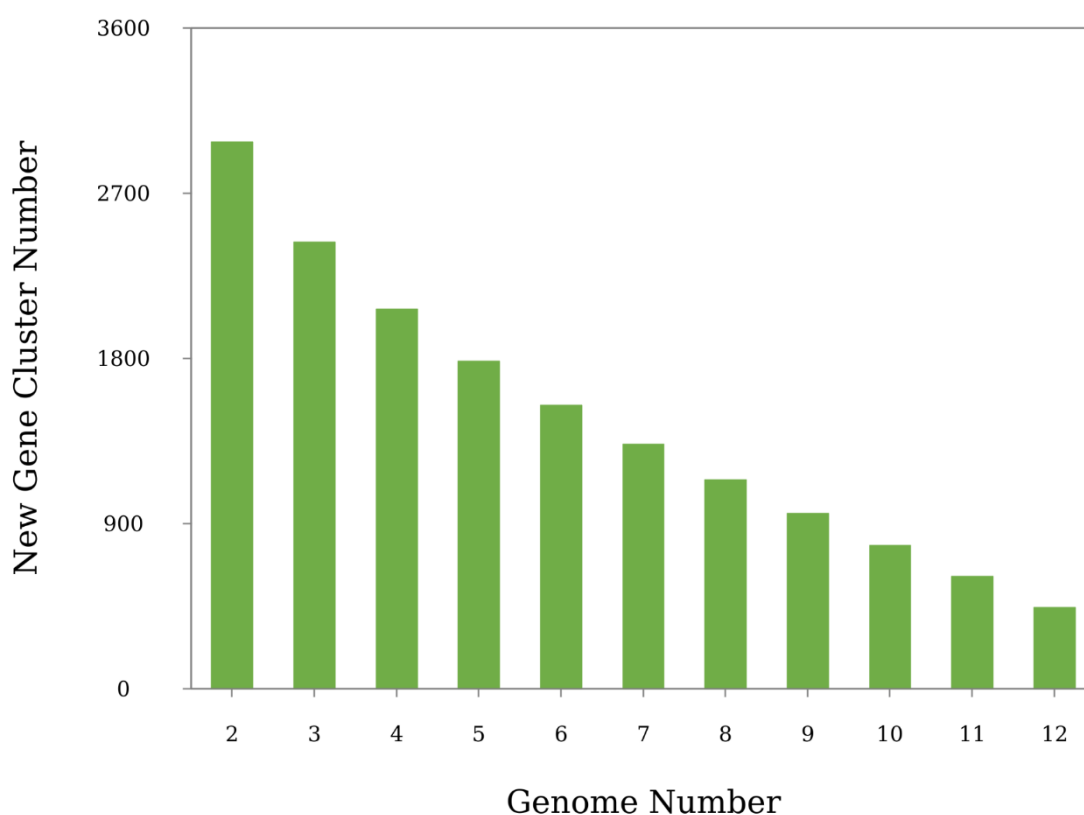
Figura 6 - Tamanho do pangenoma e padrão de crescimento



Fonte: PGAweb, 2023

O número de *clusters* que são formados na medida que são adicionados novos genomas segue uma tendência decrescente, conforme pode ser visto na Figura 7.

**Figura 7- Novos clusters à medida que são adicionados genomas**



**Fonte: PGAweb, 2023**

A análise de genes compartilhados por conservação indica o número de genes em cada genoma que são compartilhados entre os genomas analisados. Cerca de 220 genes são compartilhados entre os 12 genomas das 6 espécies, fato que está representado pelas barras uniformes à direita da Figura 8. Ao analisar o número de genes compartilhados entre 2 genomas, o par que se destaca corresponde aos genomas de *B. thuringiensis*, com 4200 genes compartilhados (barras maiores em tons acinzentados no segundo grupo da Figura 8). As espécies representantes de *B. licheniformis* ficam na segunda posição em número de genes compartilhados entre 2 clusters (2800), o que indica que *B. licheniformis* apresentam grande quantidade de genes acessórios, mas ainda apresentam muitos genes em comum com as demais espécies. *B. thuringiensis* também destaca-se quando são analisados os números de genes compartilhados por 10 genomas, apenas 11, contra cerca de 530 referentes aos genomas das outras 5 espécies analisadas. Esses resultados indicam que a espécie *B. thuringiensis* (por meio das 2 cepas representadas nessa análise) possui

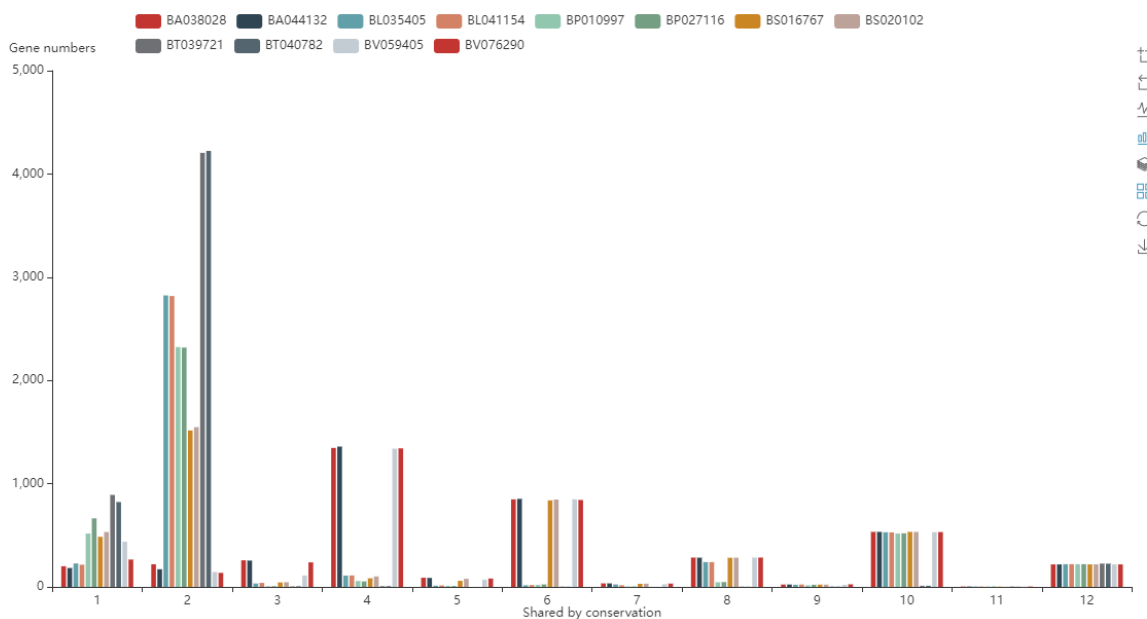
maior quantidade de genes acessórios (específicos) em comparação às outras espécies.

Tais resultados corroboram com o estudo realizado por Kim et al (2017), em que *B. thuringiensis* apresentou o maior número de genes específicos (7087) em comparação com as demais espécies analisadas no estudo (*B. amyloliquefaciens*, com 1364 genes específicos, *B. anthracis*, com 156, *B. cereus*, com 8617, e *B. subtilis*, com 3013). Em relação aos genes ortólogos, os resultados permanecem de acordo entre o presente trabalho e o estudo de Kim et al (2017), em que mais uma vez *B. thuringiensis* demonstrou maior quantidade de genes ortólogos (20882) comparado aos demais (*B. amyloliquefaciens*, com 6656 genes ortólogos, *B. anthracis*, com 6135, *B. cereus*, com 21675, e *B. subtilis*, com 10116). A maior quantidade de genes específicos e ortólogos encontrados nesse estudo deve-se, possivelmente, a maior quantidade de cepas totais analisadas, um total de 238 cepas (20 de *B. amyloliquefaciens*, 52 de *B. anthracis*, 58 de *B. cereus*, 58 de *B. subtilis* e 50 de *B. thuringiensis*) contra 12 cepas analisadas no *PGAweb* e 30 cepas analisadas no *Roary* no presente trabalho.

A novidade no presente trabalho está na comparação entre as seis espécies de *Bacillus* utilizadas como ingredientes ativos em bioinsumos brasileiros registrados, que, para os limites de meu conhecimento, consiste no primeiro trabalho de análise de pangenoma com as múltiplas espécies de *Bacillus* com foco nos efeitos de fitopatogenicidade e agricultura.



**Figura 8 - Genes compartilhados por conservação**



Fonte: PGAweb, 2023

#### 4.4 Análise de proteínas ortólogas (*OrthoVenn*)

A ferramenta *OrthoVenn* gera resultados de *output* de *clusters* de proteínas, em que cada *cluster* consiste em ortólogos ou parálogos das espécies. Enquanto os *singletons* são definidos como genes em que não foram encontrados ortólogos em outras espécies, ou seja, sequências de proteínas que não foram incluídas em nenhum *cluster*.

Como o *OrthoVenn* só permite a análise de 12 arquivos, as 12 sequências analisadas (duas de cada espécie, escolhidas aleatoriamente entre as 30 sequências iniciais descritas na Tabela 2) no *OrthoVenn* estão apresentadas na Tabela 6. Para melhor visualização dos resultados, o código inicial formado por letras das sequências *RefSeq* foi substituído pelas iniciais de suas respectivas espécies (*B. amyloliquefaciens* - BA, *B. licheniformis* - BL, *B. pumilus* - BP, *B. subtilis* - BS, *B. thuringiensis* - BT e *B. velezensis* - BV).

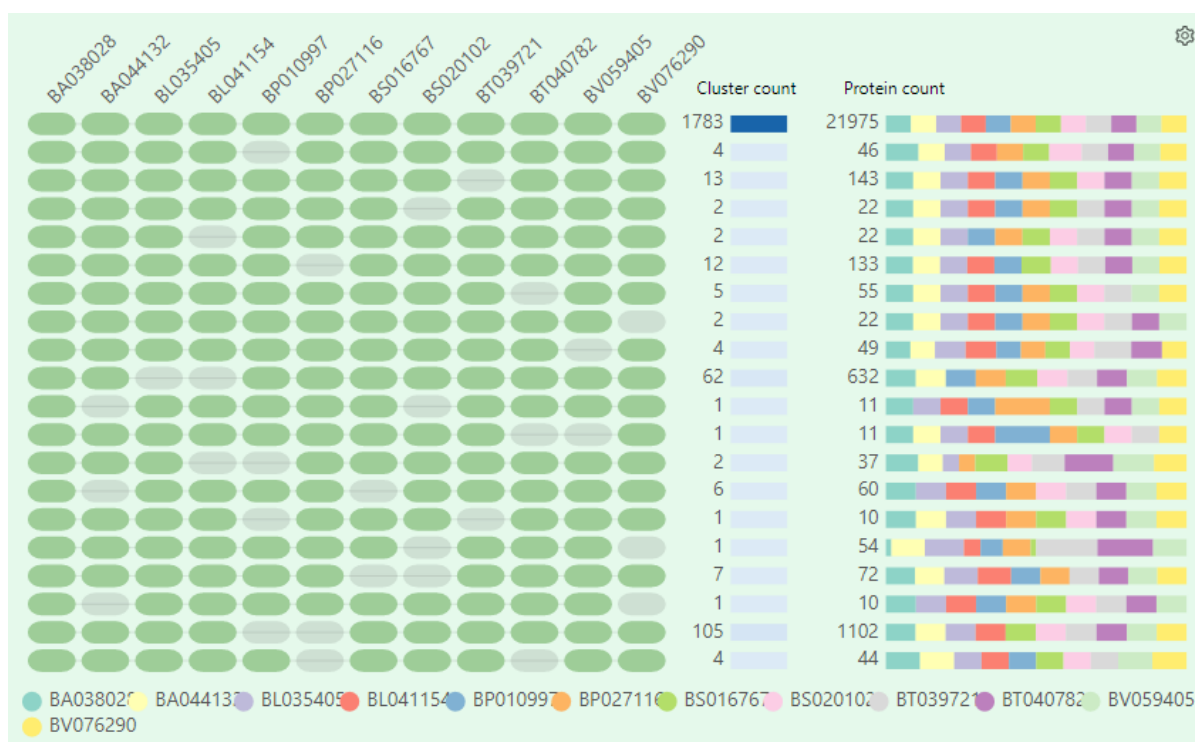
**Tabela 6 - RefSeq dos genomas analisados no *OrthoVenn* e sua nomenclatura correspondente para melhor compreensão dos resultados**

RefSeq	Genoma analisado
CP038028	BA038028
CP044132	BA044132
CP035405	BL035405
CP041154	BL041154
CP010997	BP010997
CP027116	BP027116
CP016767	BS016767
CP020102	BS020102
CP039721	BT039721
CP040782	BT040782
CP059405	BV059405
CP076290	BV076290

**Fonte: A autora, 2023**

A Figura 9 apresenta os resultados gerados pelo *OrthoVenn*, em que cada linha representa um *cluster* ortólogo e cada coluna indica a espécie. Uma célula de cor verde indica a presença de um *cluster* na espécie correspondente, enquanto uma barra de cor cinza indica a ausência de um *cluster* naquela espécie (CHEN et al, 2018). O gráfico de barras empilhadas presente na direita da Figura 9 representa o número cumulativo de sequências presentes em um *cluster* compartilhado. A coluna identificada como *Cluster count* indica o número de *clusters* compartilhados entre as espécies.

**Figura 9 - Diagrama de proteínas ortólogas por overlap number**



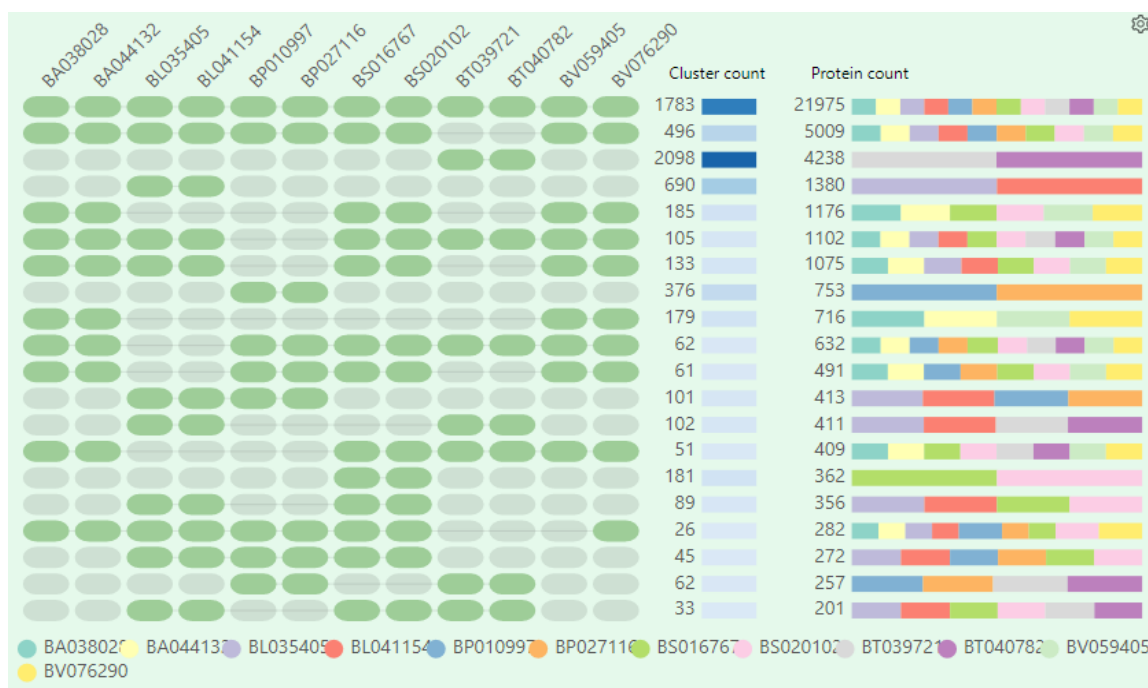
Fonte: Orthoveen, 2023

Ao analisar o mesmo diagrama resultante da análise do *OrthoVenn* selecionado para contagem de proteínas (Figura 10), as 2 sequências de *B. thuringiensis* compartilham de forma exclusiva 2098 *clusters* (que contêm 4238 proteínas no total). Em contrapartida, não compartilham 496 *clusters* que as outras espécies compartilham. Ao visualizar a primeira linha, é possível verificar a presença de 1787 clusters que são compartilhados entre todas as espécies que agrupam 21975 proteínas no total.

O fato de a espécie *B. thuringiensis* apresentar uma grande variedade de genes específicos e suas cepas compartilharem maior quantidade de *clusters* exclusivos corrobora com sua característica de possuir genes codificadores de proteínas de importância inseticida, tais como os genes *cry*, *cyt*, *vip* (proteínas inseticidas vegetativas, do inglês *vegetative insecticidal protein*) e genes de endotoxinas. Isolados de *B. thuringiensis* podem capturar genes essenciais para sua sobrevivência de outras bactérias presentes na mesma comunidade microbiana por meio de transferência horizontal de genes, portanto manifestam grande diversidade para adaptação em seus

ambientes (FANG et al, 2011; ALVES et al, 2011) e, com isso, adquirem e passam a apresentar genes únicos quando comparados a outros *Bacillus*.

**Figura 10 - Diagrama de proteínas ortólogas por contagem de proteínas**



Fonte: OrthoVenn, 2023

Os *clusters* compartilhados apenas por *B. thuringiensis* têm como dados de *Gene Ontology* (GO) *Enrichment*: resposta a antibióticos (GO:0046677), hemólise em outros organismos (GO:0044179), ligação de quinona (GO:0048038) e componente da membrana plasmática (GO:0005886).

O *cluster* de resposta a antibióticos (GO:0046677) pode estar relacionado à capacidade de resposta à resistência de insetos a toxinas existentes. O desenvolvimento de resistência em insetos alvos contra toxinas *Bt* (*B. thuringiensis*) pode ser mitigado pelo papel da toxina *Cyt* em superar a resistência a toxinas *Cry*. Além de ocorrer um mecanismo de sinergia entre essas duas toxinas de *B. thuringiensis*, em que determinado modelo de proteína *Cyt* age como receptor ligado à membrana da toxina *Cry*. Dessa forma, esse sinergismo evita a resistência de insetos e promove o aumento da atividade inseticida de toxinas *Cry* (SOBERÓN et al., 2013).

A presença de genes que promovem resposta a antibióticos soma vantagens a *B. thuringiensis* em seu papel de biodefensivo na agricultura. A microbiota intestinal presente em insetos, além de estar relacionada com as atividades fisiológicas e metabolismo do hospedeiro, influencia a resistência a proteínas *Cry*. Em caso de desregulação ou eliminação da microbiota intestinal de insetos considerados pragas por ação de antibióticos, as proteínas *Cry* têm capacidade alterada de ligação aos receptores. A eficácia das toxinas de *B. thuringiensis* dependem também da presença da comunidade bacteriana presente na microbiota intestinal dos insetos alvo, e a indução de resistência por reguladores como antibióticos sinaliza a necessidade de estratégias para manter uma produção agrícola sustentável (VISWESHWAR et al., 2015, LIU et al, 2021).

Em relação a hemólise em outros organismos (GO:0044179), esse resultado está de acordo com as características toxicológicas encontradas nas bactérias da espécie *B. thuringiensis*, que são capazes de sintetizar as proteínas *Cry* e *Cyt*, que exibem efeitos tóxicos e atividade hemolítica, respectivamente. Essas toxinas pertencem a uma classe de toxinas formadoras de poros (ou disruptoras de membranas) que, sob mudanças conformacionais, agem na membrana do hospedeiro (DAS et al., 2021). As toxinas *Cyt*, produzidas por *B. thuringiensis*, apresentam atividade citolítica (destruição de uma célula viva, devido ao desequilíbrio osmótico causado pela destruição da membrana) contra diferentes culturas de células e acarretam na hemólise de eritrócitos (hemácias) (SOBERÓN et al., 2013).

Para analisar o *cluster* de ligação de quinona (GO:0048038), deve-se observar que a ação inseticida de *B. thuringiensis* está relacionada com a presença de outras bactérias presentes na microbiota intestinal dos insetos alvos (QADRI et al., 2020, VISWESHWAR et al., 2015). Esse *cluster* deve-se, possivelmente, ao fato de que as quinonas podem agir como fatores de crescimento para a microbiota intestinal, ao influenciar o crescimento de bactérias existentes no meio e assim promover eventos de simbiose (FENN et al., 2017).

Em relação ao cluster de componente da membrana plasmática (GO:0005886), a análise parte das características das toxinas produzidas por *B. thuringiensis*. As bactérias *B. thuringiensis* tem ação pesticida conhecida, devido a ação das suas toxinas produzidas, *Cry* e *Cyt*. As proteínas *Cyt* interagem com lipídeos de membrana não-saturados, como a fosfatidilcolina. Uma das hipóteses para a sua inserção na

membrana corresponde a formação de poro em que a proteína *Cyt* liga-se à membrana celular, o que induz a formação de canais cátions-seletivos e tem como consequência a lise celular por desequilíbrio osmótico. Além disso, há a associação proteínas *Cry* com *Cyt*, que facilita a formação de oligômeros *Cry*. Esses oligômeros formados na presença de *Cyt* são capazes de formar poros quando inseridos na membrana. Portanto, formam um mecanismo de sinergia entre as duas toxinas produzidas, que sugere que uma função de *Cyt* é agir como receptor de ligação de membrana para as toxinas *Cry* (SOBERÓN et al., 2013). O *cluster* de componente da membrana plasmática possivelmente foi resultante dessas características das toxinas em promover a lise celular por meio das disrupções da membrana celular dos insetos alvo.

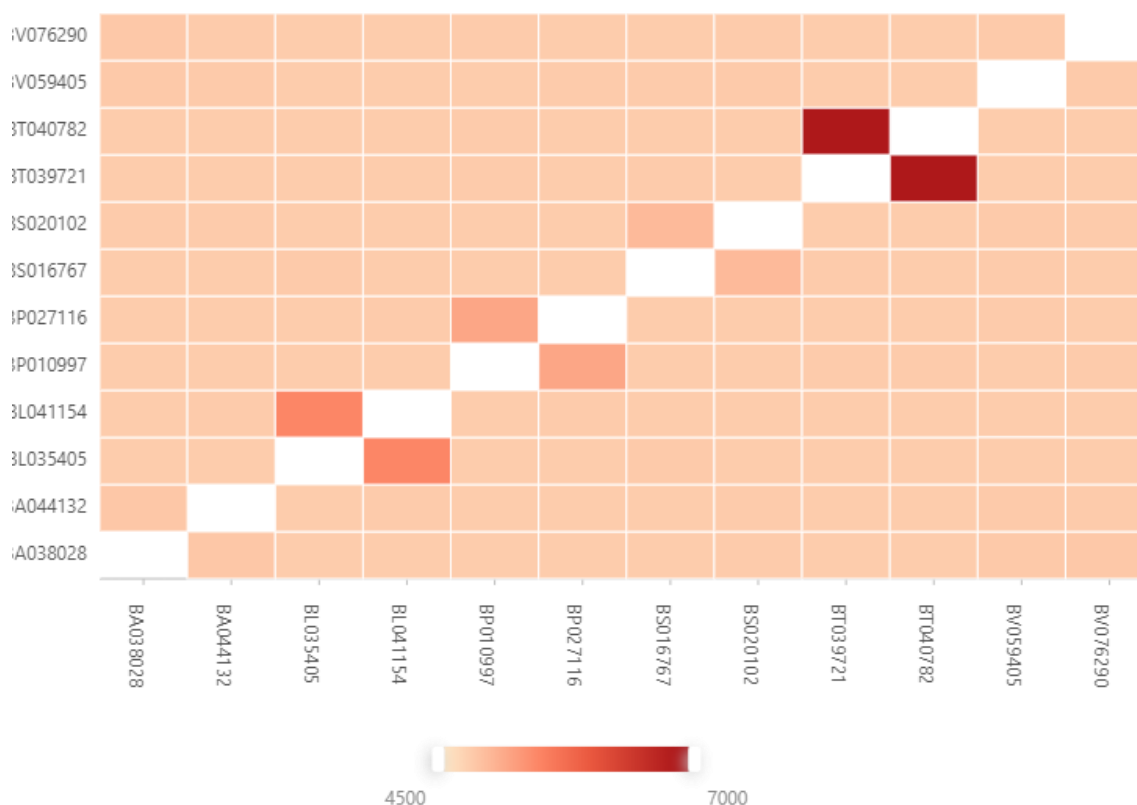
Os dados de GO descrevem o aspecto biológico de proteínas de acordo com sua função molecular, o processo biológico do qual participa e sua participação como componente celular (Uniprot, 2023).

A Figura 11 representa o *Heat Map* das 12 sequências analisadas no *OrthoVenn*, de forma que os genomas que compartilham maior número de parálogos entre eles são apresentados com coloração vermelha em tom mais escuro. Essa figura permite a visualização dos *clusters* sobrepostos para as sequências analisadas. Os números de *clusters* sobrepostos são indicados por meio de um gradiente de cores. *Clusters* sobrepostos abaixo do limite mínimo são ignorados e indicados por células em branco na figura, ou seja, quando o pareamento ocorre entre uma sequência com sua própria sequência.

Os ortogrupos compartilhados pelas duas espécies de *B. thuringiensis* são os mais destacados, de modo que mais uma vez essa espécie apresenta características genéticas diferentes das demais espécies avaliadas dentro do mesmo gênero. Pela visualização do *Heat Map*, é possível inferir que os *clusters* sobrepostos entre as sequências de BT039721 e BT040782 alcançam aproximadamente 7000 (6876) *clusters* sobrepostos, ou seja, compartilham aproximadamente 7000 *clusters* entre eles. As duas espécies de *B. licheniformis* analisadas (BL035405 e BL041154) também apresentaram grande quantidade de *clusters* compartilhados (5468) quando comparados aos demais, porém com um número inferior às sequências de *B. thuringiensis*. Por fim, *B. pumilus* (5154) e *B. subtilis* (4959) apresentaram baixas quantidades de *clusters* sobrepostos quando comparados a *B. thuringiensis* e *B.*

*licheniformis*. Os representantes das espécies *B. amyloliquefaciens* e *B. velezensis* apresentaram as menores quantidades de clusters compartilhados entre as espécies.

Figura 11 - Heat Map de grupos parálogos



Fonte: OrthoVenn, 2023

O *OrthoVenn* disponibiliza apenas a criação de um Diagrama de Venn a partir de seis sequências. Portanto, das doze sequências analisadas, foram selecionadas apenas 6 (uma de cada espécie dentre as listadas na Tabela 6) para a visualização do diagrama. O diagrama de Venn (Figura 12) representa a distribuição de *clusters* ortólogos compartilhados entre as seis espécies.

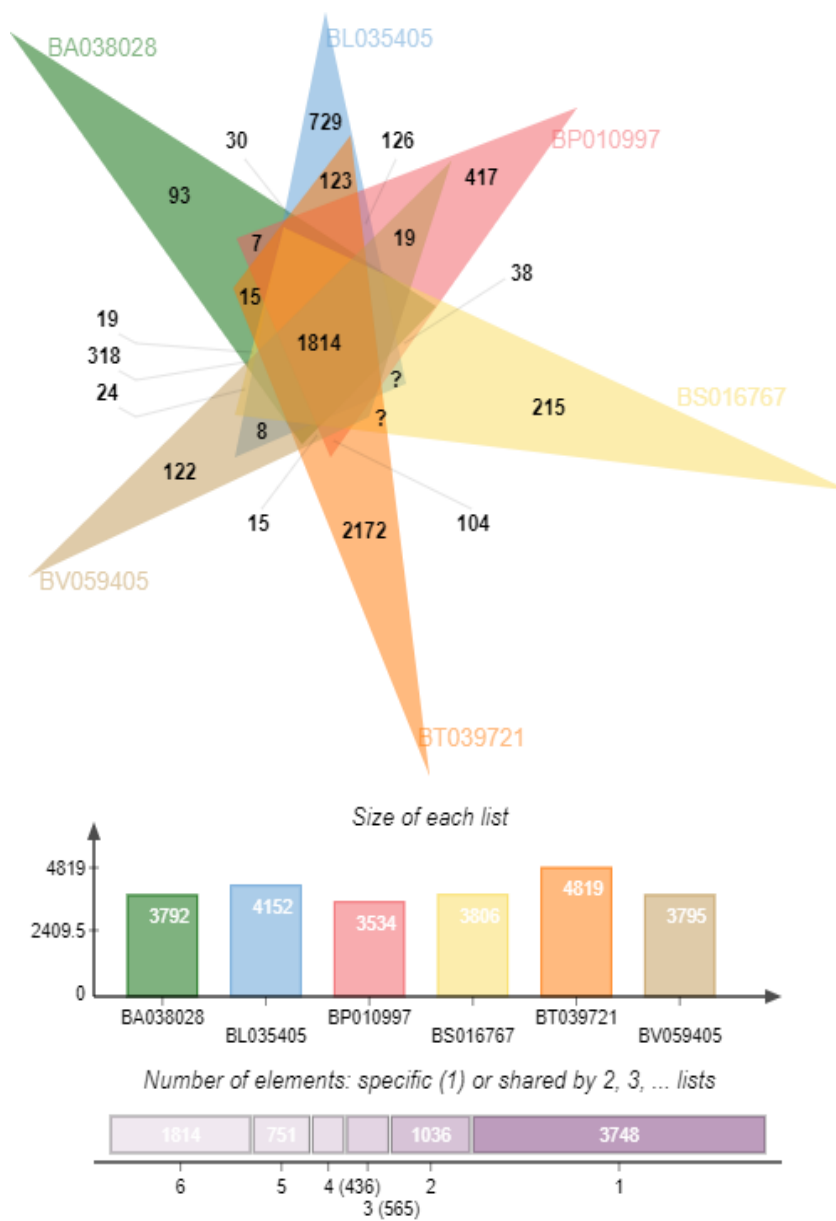
A barra abaixo do diagrama apresenta o número de clusters encontrados em cada espécie. Por exemplo, se um determinado cluster tem uma quantidade de proteínas de uma espécie (exemplo, *B. thuringiensis*) e outra quantidade de outra espécie (exemplo, *B. subtilis*), isso significa que o cluster em questão está sobreposto para *B. thuringiensis* e *B. subtilis*, e também significa que essas duas espécies compartilham esse mesmo *cluster*. De acordo com o diagrama e a barra, há 1814 *clusters* compartilhados pelas seis espécies, enquanto há 751 *clusters* compartilhados

por 5 espécies, 436 *clusters* compartilhados por 4 espécies, 565 *clusters* compartilhados por 3 espécies, 1036 *clusters* compartilhados por 2 espécies e 3748 *clusters* compartilhados por 1 única espécie (somatório de todos os *clusters* únicos de todas as espécies).

Novamente, *B. thuringiensis* e *B. licheniformis* destacam-se entre as demais espécies avaliadas dentro do mesmo gênero. A cepa representante de *B. thuringiensis* apresenta 4819 *clusters* de proteínas e *B. licheniformis* apresenta 4152 *clusters*, os maiores valores entre as cepas das espécies analisadas. Também são essas espécies que apresentam a maior quantidade de *clusters* compartilhados (123) entre duas espécies, ou *cluster* que estão sobrepostos para essas duas espécies.



Figura 12 - Diagrama de Venn



Fonte: OrthoVenn, 2023

## 5 CONCLUSÃO

Foi possível realizar as análises de Pangenoma para os genomas de *Bacillus* presentes em bioinsumos da agricultura brasileira. Em todas as análises, desde a quantidade total de proteínas até identificação de genes ortólogos e ordenação de genes compartilhados, a espécie *B. thurigiensis* destaca-se entre as seis espécies analisadas.

*B. thurigiensis* apresentou a maior quantidade de genes acessórios. Tal resultado é condizente ao fato de a espécie ser a mais antiga a ser aplicada na agricultura no manejo de pragas, e melhor descrita na literatura quanto aos mecanismos de ação das suas toxinas produzidas, *Cry* e *Cyt*. Esses resultados foram seguidos por *B. licheniformis*, que foi a espécie que apresentou a segunda maior quantidade de genes específicos e *clusters* compartilhados.

Porém, estudos posteriores são necessários para encontrar genes específicos e para maior comparação dos genomas de cada espécie para alcançar novos desdobramentos de pesquisas e aplicações, tanto no desenvolvimento de produtos, quanto na detecção molecular dos ingredientes ativos. Faz-se necessário também uma análise mais profunda em relação, especialmente, aos genes acessórios. Essa ação permite o detalhamento de produto da transcrição e tradução dos genes, assim como seus papéis e diferenças para a espécie, o que implica em como isso pode conferir uma vantagem para o bioinsumo que contém a espécie que expressa o gene em questão.

## REFERÊNCIAS BIBLIOGRÁFICAS

Agrofit Consulta Aberta - **Ministério da Agricultura, Pecuária e Abastecimento**.

Disponível em: <[http://agrofit.agricultura.gov.br/agrofit\\_cons/principal\\_agrofit\\_cons](http://agrofit.agricultura.gov.br/agrofit_cons/principal_agrofit_cons)>.

Acesso em: Abril, 2022.

ALVES, M. C.; ROSSI, J. R.; RODRIGUES, M.G. F.; ALVES, E. C. C.; FERRAUDO, A. S.; LEMOS, M. V. F.; DESIDERIO, J. A.; FERNANDES, O. A. Identificação e caracterização de genes vip e cry coleóptero-específicos em isolados de *Bacillus thuringiensis*. **Pesq. agropec. bras.**, v.46, n.9, p.1053-1060, 2011

BOHLOOL, B. B.; LADHA, J. K.; GARRITY, D. P.; GEORGE, T. Biological nitrogen fixation for sustainable agriculture: A perspective. **Plant and Soil**, p. 1-11, v.141, 1992.

BARON, S. **Medical Microbiology**. 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. PMID: 21413252.

BRAVO, A.; GILL, S.; SOBERON, M. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. **Toxicon**, v.49(4), p.423-435, 2007

CHEN, X.; ZHANG, Y. ZHANG, Z.; ZHAO, Y.; SUN, C.; YANG, M.; WANG, J.; LIU, Q.; ZHANG, B.; CHEN, M.; YU, J.; WU, J.; JIN, Z.; XIAO, J. PGAweb: A Web Server for Bacterial Pan-Genome Analysis. **Frontiers in Microbiology**, v.9, 2018.

Crop Life. **Como funciona o controle biológico na agricultura?** Disponível em: <<https://croplifebrasil.org/conceitos/controle-biologico/>>. Acesso em: Abril, 2022.

Crop Life. **Conheça os protagonistas dos produtos biológicos disponíveis no Brasil**. Disponível em: <<https://croplifebrasil.org/conceitos/conheca-os-protagonistas-dos-produtos-biologicos-disponiveis-no-brasil/>>. Acesso em: Abril, 2022.

Crop Life. **O que é controle biológico?** Disponível em: <<https://croplifebrasil.org/perguntas-frequentes/o-que-e-controle-biologico/>>. Acesso em: Janeiro, 2023

Crop Life. **Produtos biológicos registrados**. Disponível em: <<https://croplifebrasil.org/publicacoes/produtos-biologicos-registrados/>>. Acesso em: Abril, 2022.

Crop Life. **Regulamentação de produtos biológicos**. Disponível em: <<https://croplifebrasil.org/produtos-biologicos/regulamentacao-de-produtos-biologicos/>>. Acesso em: Abril, 2022.

DAMALAS, C. A.; KOUTROUBAS, S. D. Current Status and Recent Developments in Biopesticide Use. **MDPI Agriculture**, 8, 13, 2018.

DAS, S. K.; PRADHAN, S. K.; SAMAL, K. C.; SINGH, N. R. Structural, functional, and evolutionary analysis of Cry toxins of *Bacillus thuringiensis*: an in silico study. **Egyptian Journal of Biological Pest Control**, p.31-44, 2021

Entrez Programming Utilities Help. **NCBI Help Manual**. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK25501/>>. Acesso em: Janeiro, 2023.

EPA - United States Environmental Protection Agency. **Biopesticides**. Disponível em: <<https://www.epa.gov/pesticides/biopesticides>>. Acesso em: Abril, 2022.

ETE Toolkit: A Python Environment for Tree Exploration. **Phylogenetic tree (newick) viewer**. Disponível em: <<http://etetoolkit.org/treeview/>>. Acesso em: Janeiro, 2023

FANG, Y.; LI, Z.; LIU, J.; SHU, C.; WANG, X.; ZHANG, X.; YU, X.; ZHAO, D.; LIU, G.; HU, S.; ZHANG, J.; MSSLLEM, I.; YU, J. A pangenomic study of *Bacillus thuringiensis*. **Journal of Genetics and Genomics**, v.38, p.567-576, 2011

FENN, K.; STRANDWITZ, P.; STEWART, E. J.; DIMISE, E., RUBIN, S.; GURUBACHARYA, S.; CLARDY, J.; LEWIS, K. Quinones are growth factors for the human gut microbiota. **Microbiome**, v.5, 2017

FILHO, J. C. *et al.* Complete genome sequence of native *Bacillus cereus* strains isolated from intestinal tract of the crab *Ucides* sp. **Data in Brief**, v. 16, p. 381-385, 2018

FONTES, E. M. G.; VALADARIS-INGLIS, M. C. **Controle Biológico de Pragas da Agricultura**. Editoras Técnicas, Embrapa, 2020.

FU, L.; NIU, B.; ZHU, Z.; WU, S.; LI, W. CD-HIT: accelerated for clustering the next-generation sequencing data. **Bioinformatics**, v.28, p.3150-3152, 2012

UniProt. **Gene Ontology (GO)**. Disponível em:

<[https://www.uniprot.org/help/gene\\_ontology](https://www.uniprot.org/help/gene_ontology)>. Acesso em: Janeiro, 2023

GUIMARÃES, L. C.; JESUS, L. B.; VIANA, M. V. C.; SILVA, A.; RAMOS, R. T. J.; SOARES, S. C.; AZEVEDO, C. Inside the Pan-genome - Methods and Software Overview. **Current Genomics**, v.16, p.245-252, 2015

HECKEL, D. G. How do toxins from *Bacillus thuringiensis* kill insects? An evolutionary perspective. **Insect Biochemistry and Physiology**, v.104, 2020

KIM, Y.; KOH, I.; LIM, M.; CHUNG, W.; RHO, M. Pan-genome analysis of *Bacillus* for microbiome profiling. **Scientific Reports**, 2017.

KOONIN, E. V. Orthologs, Paralogs, and Evolutionary Genomics. **Annual Review of Genetics**, v.39, p.309-338, 2005

KWON, E.; LEE, J.; PARK, J. W.; KIM, S. Application of comparative genomics in the development of DNA probes to detect *Bacillus cereus* and *Bacillus subtilis*. **Food Science and Technology**, v.142, 2021.

LIU, L.; LI, Z.; LUO, X.; ZHANG, X.; CHOU, S.; WANG, J.; HE, J. Which Is Stronger? A Continuing Battle Between Cry Toxins and Insects. **Frontiers in Microbiology**, v.12, 2021

MARCUS, S.; LEE, H.; SCHATZ, M. C. SplitMEM: a graphical algorithm for pan-genome analysis with suffix skips. **Bioinformatics**, p.3476-3483, v.30, 2014.

MARRONE, P. G. Pesticidal natural products – status and future potential. **Pest Management Science**, p. 2325-2340, v. 75, 2019.

MEDINI, D.; DONATI, C.; TETTELIN, H.; MASIGNANI, V.; RAPPUOLI. The microbial pan-genome. **Current Opinion in Genetics & Development**, p.589-594, v. 15, 2005.

NCBI. **Genome**. Disponível em: <<https://www.ncbi.nlm.nih.gov/>>. Acesso em: Abril, 2022.

MURAS, A.; ROMERO, M.; MAYER, C.; OTERO, A. Biotechnological applications of *Bacillus licheniformis*. **Critical Reviews in Biotechnology**, v.41, n.4, p. 609-627, 2021

OSBOURN, A. E.; FIELD, B. Operons. **Cellular and Molecular Life Sciences**, v.66, p.3755-3775, 2009

PAGE, A. J.; CUMMINS, C. A.; HUNT, M.; WONG, V. K.; REUTER, S.; HOLDEN, M.; FOOKES, M.; FALUSH, D.; KEANE, J. A.; PARKHILL, J. Roary: rapid large-scale prokaryote pan genome analysis. **Bioinformatics**, v. 31(22), p.3691–3693, 2015.

PEOPLES, M. B.; CRASWELL, E. T. Biological nitrogen fixation: Investments, expectations and actual contributions to agriculture. **Plant and Soil**, p.13-39, v.141, 1992.

PIGGOT, P. J. *Bacillus subtilis*. Encyclopedia of Microbiology. **Academic Press**, p. 45-56, 2009

QADRI, M.; SHORT, S.; GAST, K.; HERNANDEZ, J.; WONG, A, C. Microbiome Innovation in Agriculture: Development of Microbial Based Tools for Insect Pest Management. **Frontiers in Sustainable Food Systems**, v.4, 2020

RABBEE, M. F.; ALI, M. S.; CHOI, J.; HWANG, B. S.; JEONG, S. C.; BAEK, K. *Bacillus velezensis*: A Valuable Member of Bioactive Molecules within Plant Microbiomes. **Molecules**, 24(6), 2019

SCHÜNEMANN, R.; KNAAK, N.; FIUZA, L. M. Mode of Action and Specificity of *Bacillus thuringiensis* Toxins in the Control of Caterpillars and Stink Bugs in Soybean Culture. **ISRN Microbiology**, v.2014, p. 1-12, 2014

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, v.30, 2014.

SEPPEY, M.; MANNI, M.; ZDOBNOV, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. **Methods in Molecular Biology**, v.1962, p.227-245, 2019.

SHI, H.; XU, X. Learning the Computing Quality Statistics Method of the Sequence Reads. **Advances in Intelligent Systems Research**, v.135, p. 255-259, 2016.

SOBERÓN, M.; LÓPEZ-DÍAZ, J. A.; BRAVO, A. Cyt toxins produced by *Bacillus thuringiensis*: A protein fold conserved in several pathogenic microorganisms.

**Peptides**, v.41, p.87-93, 2013

Tange, O. (2022, September 22). **GNU Parallel 20220922 ('Elizabeth') released**. Zenodo. <https://doi.org/10.5281/zenodo.7105792>

TONELLI, M. L.; FIGUEREDO, M. S.; RODRÍGUEZ, J.; FABRA, A.; IBANEZ, F. Induced systemic resistance -like responses elicited by rhizobia. **Plant and Soil**, v.448, p.1-14, 2020

TORTORA, G. J.; FUNKE, B. R.; CASE, C. L. **Microbiologia**; tradução: Danielle Soares de Oliveira Daian, Luis Fernando Marques Dorvillé; revisão técnica: Flávio Guimarães da Fonseca, Ana Paula Guedes Frazzon, Jeverson Frazzon. – 12. ed. – Porto Alegre : Artmed, 2017

VERNIKOS, G.; MEDINI, D.; RILEY, D. R.; TETTELIN, H. Ten years of pan-genome analyses. **Current Opinion in Microbiology**, p. 148-154, v. 23, 2015.

VISWESHWAR, R.; SHARMA, H. C.; AKBAR, S. M. D.; SREERAMULU, K. Elimination of Gut Microbes with Antibiotics Confers Resistance to *Bacillus thuringiensis* Toxin Proteins in *Helicoverpa armigera* (Hubner). **Applied Biochemistry and Biotechnology**, v.177, p. 1621-1637, 2015

WANG, Q.; ZHANG, L.; ZHANG, Y.; CHEN, H.; SONG, J.; LYU, M.; CHEN, R.; ZHANG, L. Comparative genomic analyses reveal genetic characteristics and pathogenic factors of *Bacillus pumilus* HM-7. **Front. Microbiol**, v.13, 2022

WU, Y. Chapter Six - Detection and Mechanisms of Resistance Evolved in Insects to Cry Toxins from *Bacillus thuringiensis*. **Advances in Insect Physiology**, v.47, p. 297-342, 2014

XU, L.; DONG, Z.; FANG, L. LUO, Y.; WEI, Z.; GUO, H.; ZHANG, G.; GU, Y.; COLEMAN-DERR, D.; XIA, Q.; WANG, Y. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. **Nucleic Acids Research**, v. 47, 2019.

YI, G.; SZE, S.; THON, M. R. Identifying clusters of functionally related genes in genomes. **Bioinformatics**, v.23, p.1053-1060, 2007

## APÊNDICE A: ANEXO DE PIPELINES



## 1. **DOWNLOAD DE SEQUÊNCIAS**

```
source ~/anaconda3/bin/activate
```

```
yes | conda create --name ncbi
```

```
conda activate ncbi
```

```
conda config --add channels defaults
```

```
conda config --add channels bioconda
```

```
conda config --add channels conda-forge
```

```
yes | conda install entrez-direct
```

```
nano
```

```
#!/bin/bash
```

```
genomas=(CP072120 CP038028 CP089530 CP044132 CP071970 CP082283
CP041691 CP071042 CP021505 CP082243 CP090312 CP049330 CP014795
CP035405 CP035188 CP038186 CP041154 CP065029 CP023729 CP014781
LT906438 CP027116 CP010997 CP058951 CP085037 CP011007 CP007436
CP027034 AP014928 CP016784 CP020102 CP068982 CP016767 CP028202
CP104097 CP061870 CP092369 CP045820 CP103350 CP089148 CP106947
CP039721 CP045030 CP083134 CP040782 CP052061 CP017577 CP083129
CP019230 CP053972 CP097467 CP082262 CP024203 CP103856 CP059405
CP076290 CP083764 CP065791 CP079099 CP015443)
```

```
for i in ${!genomas[@]};
```

```
do
```

```
echo ${genomas[i]}
```

```
esearch -db nuccore -query ${genomas[i]} | efetch -format fasta > ${genomas[i]}.fna
```

```
esearch -db nuccore -query ${genomas[i]} | efetch -format gbwithparts >
${genomas[i]}.gbk
```

```
esearch -db nuccore -query ${genomas[i]} | efetch -format fasta_cds_aa -id
${genomas[i]} > ${genomas[i]}.faa
```

```
done
```

(após salvar e sair de “nano”) bash download.bash

## 2. ANOTAÇÃO DOS GENOMAS COM PROKKA

```
nano script.prokka.bash
```

```
#!/bin/bash
```

```
genomas=( CP072120 CP038028 CP089530 CP044132 CP071970 CP082283
CP041691 CP071042 CP021505 CP082243 CP090312 CP049330 CP014795
CP035405 CP035188 CP038186 CP041154 CP065029 CP023729 CP014781
LT906438 CP027116 CP010997 CP058951 CP085037 CP011007 CP007436
CP027034 AP014928 CP016784 CP020102 CP068982 CP016767 CP028202
CP104097 CP061870 CP092369 CP045820 CP103350 CP089148 CP106947
CP039721 CP045030 CP083134 CP040782 CP052061 CP017577 CP083129
CP019230 CP053972 CP097467 CP082262 CP024203 CP103856 CP059405
CP076290 CP083764 CP065791 CP079099 CP015443)
```

```
for i in ${!genomas[@]};
```

```
do
```

```
echo “Analisando o genoma ‘${genomas[i]}’;”
```

```
prokka --compliant --addgenes --locustag “${genomas[i]}” \
```

```
--genus ‘Bacillus’ --kingdom ‘Bacteria’ --gcode 11 \
```

```
--proteins references.gbk --evaluate 1e-10 --cpus 15 \
```

```
--cdsrnaolap \
```

```
--outdir prokka --force --prefix ${genomas[i]} \
```

```
fna2/${genomas[i]}.fna
```

```
done
```

### **3. ORDENAÇÃO DE GENES COM *ROARY***

```
roary -e --mafft -l 40 -cd 100 -t 11 -p 4 \
```

```
> -f resultadoroary40 prokka/*gff
```

**ANEXO A - Direitos autorais - Lei n. 9.610, de 19 de fevereiro de 1998**



**Presidência da República  
Casa Civil  
Subchefia para Assuntos Jurídicos**

**LEI Nº 9.610, DE 19 DE FEVEREIRO DE 1998<sup>1</sup>.**

Mensagem de veto

**Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências.**

**O PRESIDENTE DA REPÚBLICA** Faço saber que o Congresso Nacional decreta e eu sanciono a seguinte Lei:

Título I - Disposições Preliminares

Art. 1º Esta Lei regula os direitos autorais, entendendo-se sob esta denominação os direitos de autor e os que lhes são conexos.

Art. 2º Os estrangeiros domiciliados no exterior gozarão da proteção assegurada nos acordos, convenções e tratados em vigor no Brasil.

Parágrafo único. Aplica-se o disposto nesta Lei aos nacionais ou pessoas domiciliadas em país que assegure aos brasileiros ou pessoas domiciliadas no Brasil a reciprocidade na proteção aos direitos autorais ou equivalentes.

Art. 3º Os direitos autorais reputam-se, para os efeitos legais, bens móveis.

Art. 4º Interpretam-se restritivamente os negócios jurídicos sobre os direitos autorais.

Art. 5º Para os efeitos desta Lei, considera-se:

I - publicação - o oferecimento de obra literária, artística ou científica ao conhecimento do público, com o consentimento do autor, ou de qualquer outro titular de direito de autor, por qualquer forma ou processo;

II - transmissão ou emissão - a difusão de sons ou de sons e imagens, por meio de ondas radioelétricas; sinais de satélite; fio, cabo ou outro condutor; meios óticos ou qualquer outro processo eletromagnético;

III - retransmissão - a emissão simultânea da transmissão de uma empresa por outra;

IV - distribuição - a colocação à disposição do público do original ou cópia de obras literárias, artísticas ou científicas, interpretações ou execuções fixadas e fonogramas, mediante a venda, locação ou qualquer outra forma de transferência de propriedade ou posse;

V - comunicação ao público - ato mediante o qual a obra é colocada ao alcance do público, por qualquer meio ou procedimento e que não consista na distribuição de exemplares;

VI - reprodução - a cópia de um ou vários exemplares de uma obra literária, artística ou científica ou de um fonograma, de qualquer forma tangível, incluindo qualquer armazenamento permanente ou temporário por meios eletrônicos ou qualquer outro meio de fixação que venha a ser desenvolvido;

VII - contrafação - a reprodução não autorizada;

VIII - obra:

a) em co-autoria - quando é criada em comum, por dois ou mais autores;

b) anônima - quando não se indica o nome do autor, por sua vontade ou por ser desconhecido;

c) pseudônima - quando o autor se oculta sob nome suposto;

d) inédita - a que não haja sido objeto de publicação;

e) póstuma - a que se publique após a morte do autor;

f) originária - a criação primígena;

g) derivada - a que, constituindo criação intelectual nova, resulta da transformação de obra originária;

h) coletiva - a criada por iniciativa, organização e responsabilidade de uma pessoa física ou jurídica, que a publica sob seu nome ou marca e que é constituída pela participação de diferentes autores, cujas contribuições se fundem numa criação autônoma;

i) audiovisual - a que resulta da fixação de imagens com ou sem som, que tenha a finalidade de criar, por meio de sua reprodução, a impressão de movimento, independentemente dos processos de sua captação, do suporte usado inicial ou posteriormente para fixá-lo, bem como dos meios utilizados para sua veiculação;

IX - fonograma - toda fixação de sons de uma execução ou interpretação ou de outros sons, ou de uma representação de sons que não seja uma fixação incluída em uma obra audiovisual;

X - editor - a pessoa física ou jurídica à qual se atribui o direito exclusivo de reprodução da obra e o dever de divulgá-la, nos limites previstos no contrato de edição;

XI - produtor - a pessoa física ou jurídica que toma a iniciativa e tem a responsabilidade econômica da primeira fixação do fonograma ou da obra audiovisual, qualquer que seja a natureza do suporte utilizado;

XII - radiodifusão - a transmissão sem fio, inclusive por satélites, de sons ou imagens e sons ou das representações desses, para recepção ao público e a transmissão de sinais codificados, quando os meios de decodificação sejam oferecidos ao público pelo organismo de radiodifusão ou com seu consentimento;

XIII - artistas intérpretes ou executantes - todos os atores, cantores, músicos, bailarinos ou outras pessoas que representem um papel, cantem, recitem, declamem, interpretem ou executem em qualquer forma obras literárias ou artísticas ou expressões do folclore.

Art. 6º Não serão de domínio da União, dos Estados, do Distrito Federal ou dos Municípios as obras por eles simplesmente subvencionadas.