



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

GUSTAVO MANOEL TEIXEIRA

**SEQUENCIAMENTO GENÔMICO DE *Bacillus velezensis***  
**cepa LABIM22:**  
UM ESTUDO *in silico* DO POTENCIAL BIOTECNOLÓGICO

GUSTAVO MANOEL TEIXEIRA

**SEQUENCIAMENTO GENÔMICO DE *Bacillus velezensis***  
**cepa LABIM22:**  
UM ESTUDO *in silico* DO POTENCIAL BIOTECNOLÓGICO

Defesa apresentada ao Programa de Pós-Graduação em Microbiologia da Universidade Estadual de Londrina como parte dos requisitos para a obtenção do título de Mestre em Microbiologia

Orientador: Prof. Dr. Admilton Gonçalves de Oliveira Junior

Co-orientador: Prof. Dr. Ulisses de Padua Pereira

Londrina  
2020

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Teixeira, Gustavo Manoel.

Sequenciamento genômico de *Bacillus velezensis* cepa LABIM22: um estudo *in silico* do potencial biotecnológico / Gustavo Manoel Teixeira. - Londrina, 2020.  
59 f. : il.

Orientador: Admilton Gonçalves De Oliveira Junior.

Coorientador: Ulisses de Padua Pereira .

Dissertação (Mestrado em Microbiologia) - Universidade Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Microbiologia, 2020.

Inclui bibliografia.

1. Genômica - Tese. 2. Biotecnologia - Tese. 3. Biocontrole - Tese. 4. microrganismos - Tese. I. De Oliveira Junior, Admilton Gonçalves . II. Pereira , Ulisses de Padua . III. Universidade Estadual de Londrina. Centro de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia. IV. Título.

CDU 579

GUSTAVO MANOEL TEIXEIRA

**SEQUENCIAMENTO GENÔMICO DE *Bacillus velezensis*  
cepa LABIM22:**

UM ESTUDO *in silico* DO POTENCIAL BIOTECNOLÓGICO

Defesa apresentada ao Programa de Pós-Graduação em Microbiologia da Universidade Estadual de Londrina como parte dos requisitos para a obtenção do título de Mestre em Microbiologia

**BANCA EXAMINADORA**

---

Orientador: Prof. Dr. Admilton Gonçalves de  
Oliveira Junior  
Universidade Estadual de Londrina – UEL

---

Prof. Dr. Leandro Simões Azeredo Gonçalves  
Universidade Estadual de Londrina – UEL

---

Prof. Dr. Youssef Khamis Youssef Ahmed  
Agricultural Research Center – ARC  
Cairo, Egypt

Londrina, 23 de março de 2020

## **AGRADECIMENTOS**

Agradeço a todos que fizeram parte do desenvolvimento da minha vida acadêmica até aqui e me proporcionaram subsídio para poder continuar contribuindo com a ciência no Brasil.

Agradeço primeiramente minha mãe, que sempre me deu suporte e me apoiou em minhas decisões e fazendo seu trabalho de mãe da melhor maneira possível para me transformar no ser humano que me tornei atualmente. Ao meu pai que sempre soube, por algum motivo, que eu iria me interessar pela área da Microbiologia mesmo antes de eu ter contato com este campo de estudo.

A todos os meus que me fizeram acreditar que eu seria capaz de passar por esta fase da minha vida mesmo em momentos difíceis.

Gostaria de agradecer a equipe do LABIM, LABBEP e do SBFGE (UNESP – Botucatu) por contribuírem com espaço físico e apoio para o desenvolvimento e conclusão desta dissertação.

Ao meu orientador Admilton Gonçalves de Oliveira que me deu todo o suporte para dar início aos estudos com bioinformática, ao meu co-orientador Ulisses de Padua Pereira que auxiliou com sua expertise em bioinformática.

TEIXEIRA, Gustavo Manoel. **Sequenciamento genômico de *Bacillus velezensis* cepa LABIM22**: um estudo *in silico* do potencial biotecnológico. 2020. 56 f. Dissertação (Mestrado em Microbiologia) - Universidade Estadual de Londrina, Londrina, 2020.

## RESUMO

A exploração de novas estratégias para o controle de doenças no campo que causam perdas consideráveis nas produções cresceu nas últimas décadas. Diferentes alternativas vem ganhando destaque com o avanço tecnológico exponencial do final do século XX e que continua até os dias de hoje, dentre elas o uso de microrganismos como agentes de biocontrole vem ganhando destaque por se tratar de uma alternativa natural que utiliza o processos biológicos já existentes no ecossistema para combater doenças, além de contribuir para o manejo integrado de doenças (MID) que leva a uma diminuição do uso de agrotóxicos que causam danos ao meio ambiente quando usados como única estratégia de controle. As bactérias do gênero *Bacillus* chamam a atenção devido a características importantes como a formação de endósporos e uma diversa maquinaria molecular que produz uma vasta gama de metabólitos secundários com diferentes atividades como diminuição do crescimento micelial de fungos fitopatogênicos, promoção de crescimento, indução de sistemas de defesas contra doenças, atividade nematocida dentre outros ainda não descritos. Neste trabalho o genoma de um *Bacillus velezensis* cepa LABIM22 foi estudado para buscar entender melhor os mecanismos genéticos relacionados as atividades antifúngicas apresentadas pela cepa em estudos *in vitro*. Para isto seu genoma foi sequenciado, comparado e analisado buscando genes relacionados a esta atividade bem como sua capacidade de formação de biofilme, característica importante para colonização da rizosfera e estabelecimento do microrganismo. Seu genoma possui 3.996.396 pb e 4042 CDS (sequencias codificantes) e um teor GC de 46,4%. Utilizando a plataforma antiSMASH foram encontradas 13 BGCs (Biosynthetic Gene Clusters), alguns ligados a síntese de importantes famílias de moléculas antifúngicas como Fengicinas e Surfactinas. Diversos genes importantes ligados a regulação/formação de biofilme foram encontrados no genoma como o operon *epsA-O* ligado a síntese de exopolissacarídeos, *abrB* um repressor de genes de indução de esporulação e *degQ*, um gene ligado a degradação de compostos extracelulares. Sua diversa maquinaria metabólica, alta similaridade com cepas referência de biocontrole e presença de genes importantes para esta atividade, fazem de LABIM22 um excelente candidato a agente de biocontrole e dá subsídio para mais estudos deste isolado.

**Palavras-chave:** microrganismos; biocontrole; *Bacillus*; antifúngicas; genoma; biofilme.

TEIXEIRA, Gustavo Manoel. **Genomic sequencing of *Bacillus velezensis* strain LABIM22: an *in silico* study of the biotechnological potential.** 2020. 56 p. Master's in Microbiology – State University of Londrina, Londrina, 2020.

## ABSTRACT

The exploration of new strategies to disease control on crops that cause considerable losses on net production have grown in the past decades. Different alternatives have been highlighted with the exponential technological advance that took place by the end of the XXth century and that continues until nowadays, among them the use of microorganisms as biocontrol agents have gained attention since it is a natural alternative that utilizes biological processes already existing in ecosystems to fight diseases, in addition to contribute to the Integrated Disease Management, which brings a decrease on the use pesticides that harms the environment when used as the only control strategy. Bacteria belonging to the *Bacillus* genus have drawn attention due to important characteristics as the endospore formation and a diverse molecular machinery able to produce a vast range of secondary metabolites with different activities such as decrease in mycelial growth of phytopathogens fungi, growth promoting, induction of defense systems against diseases, nematicide activity among others yet to be described. In this work the *Bacillus velezensis* strain LABIM22's genome was studied in the aim of better understand the genetic mechanisms related to the antifungal activity shown by the strain on *in vitro* assays. To accomplish this its genome was sequenced, compared and analyzed looking for genes related to this activity as well as its capacity of biofilm formation, important characteristic for root and rhizosphere colonization and microorganism establishment. Its genome possesses 3.996.396 pb and 4042 CDS (Coding Sequences) with a GC content of 46,4%. Using the antiSMASH platform 13 BGCs (Biosynthetic Gene Cluster) were found within the genome, some associated with the synthesis of important families of antifungal molecules like fengycin and surfactins. Several important genes related to biofilm regulation/formation were found within the genome like the *epsA-O* operon linked to the synthesis of exopolysaccharides, *abrB* a gene repressor of the sporulation induction and *degQ* which is related to the degradation of extracellular compounds. Its diverse metabolic machinery, high similarity to reference biocontrol strains and presence of important genes to this activity make LABIM22 an excellent candidate as biocontrol agent and allows further studies of this isolated.

**Keywords:** microorganisms; biocontrol; *Bacillus*; antifungal; genome; biofilm.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1.</b> Ilustrando a leitura dos dados do sequenciamento de Sanger e Maxam Gilbert .....	16
<b>Figura 2.</b> PCR em ponte realizado em plataforma Illumina de segunda geração para amplificação dos fragmentos.....	19
<b>Figura 3.</b> Representação gráfica dos poços de detecção de luz usando a tecnologia ZMV (Zero Mode Waveguide), o diâmetro do poço é menor que o comprimento de onda de luz visível diminuindo problemas como ruído de fundo fazendo com que o sinal de luz emitido por apenas uma molécula de DNA seja detectado com mais facilidade .....	20
<b>Figura 4.</b> Oxford Nanopore Minlon, em aumento a membrana com o nanoporo associado, a flow cell contém 2048 destes nanoporos. Esta tecnologia se destaca pela portabilidade e rapidez no fornecimento de dados .....	21
<b>Figura 5.</b> Layout do site do NCBI onde se encontram todas as informações a respeito do GenBank.....	22
<b>Figura 6.</b> Layout da plataforma on-line BLAST para comparação de seqüências de nucleotídeos e aminoácidos .....	22
<b>Figura 7.</b> Web Server antiSMASH utilizado para busca de BGCs ligados ao metabolismo secundário do genoma desejado .....	25

## FIGURAS DO ARTIGO

<b>Figure 1.</b> QCAST report showing the main metrics chosen to evaluate the better assembly among the ones made, better results are highlighted in blue and worst results in red.....	34
<b>Figure 2.</b> Phylogenomic tree representing the similarity between strains from the <i>Bacillus subtilis</i> group. The matrix was generated by Gegenees and exported to SplitsTree. Highlighted in green are the most similar strains to LABIM22 belonging to the <i>Bacillus velezensis</i> species. In yellow are the <i>B. amyloliquefaciens</i> species, less similar to	

LABIM22 and highlighted in red are the <i>B. subtilis</i> species with a <i>B. cereus</i> FRI35 as an extern group .....	36
<b>Figure 3.</b> Circular representation of LABIM22 genome using the program BRIG. From the inside to the outside the legends are the following: (1) GC Content, (2) GC Skew, (3) BGCs position on the genome indicated by antiSMASH, (4-10) respectively LABIM40, FZB42, S141, QST713, Bac57, NKG-1 and Y2 .....	38
<b>Figure 4.</b> Pellicle formation on liquid-air interface on static liquid medium. Columns A and C: Medium Control; Columns B and D: inoculated with 10 µL of solution containing LABIM22 cells adjusted to the tube 0,5 of MacFarland scale .....	39
<b>Figure 5.</b> In vitro tests made with LABIM22, its motility capacity was tested on a TSA 0,8%. A cell inoculum was put in the middle of the dish and its growth was evaluated after 24h. LABIM22 was able to colonize the whole surface of the medium because of its ability to move thanks to flagella that its genome showed that has the capacity to synthetize .....	39
<b>Figure 6.</b> LABIM22 root colonization. Set of 15 images showing LABIM22 in vivo colonization capacity .....	43
<b>Figure 7.</b> Representation of the main genes found within the LABIM22 genome related with the antifungal activity and sporulation inducing the biofilm formation/regulation aiming the root colonization .....	44
<b>Figure 7.</b> Inoculation effect of bacteria on the germination percentage (%) on soybean culture. The bars are averages of 4 repetitions; each repetition contained 50 seed per treatment. Bars with different letters above differs significantly between them at $p < 0,05$ , according with the Tukey test .....	45
<b>Figure 8.</b> Effect of bacteria inoculation on total root length and root surface area of corn (a) and soybean (b and c) roots. The bars are averages of 4 repetitions; each repetition contained 50 seed per treatment. Bars with different letters above differs significantly between them at $p < 0,05$ , according with the Tukey test .....	45

## LISTA DE TABELAS

<b>Table 1.</b> Main genomic data from LABIM22 genome.....	34
<b>Table 2.</b> Genome comparison of LABIM22 with other Bacillus species.....	35
<b>Table 3.</b> Biosynthetic Gene Clusters (BGCs) found within LABIM22 genome using the webserver antiSMASH 5.1.0 .....	40
<b>Table 4.</b> Genes related to biofilm formation/regulation on LABIM22.....	41

## LISTA DE ABREVIATURAS

MID	Manejo Integrada de Doenças
BLAST	Basic Local Alignment Search Tool
CDS	Coding Sequences
antiSMASH	Antibiotics and Secondary Metabolite Analysis Shell
GC	Guanina e Citosina
PGPR	Plant Growth Promoter Rhizobacteria
NCBI	National Center for Biotechnology Information
ANI	Average Nucleotide Identity
dDDH	Digital DNA-DNA Hybridization
NRPS	Non-Ribossomal Peptide Synthetase
TransATPKS	trans-Acyl Transferase Polyketide Synthetase
rRNA	RNA ribossomal
NRPS	Non Ribossomal Peptide Synthetase
NRP	Non Ribossomal Peptide
PKS	Polyketide Sinthetase
T3PKS	PKS do Tipo 3
AT	ACETYL TRANSFERASE
PCR	Polimerase Chain Reaction
USB	Universal Serial Bus
MIBIG	Minimum Information about a Biosynthetic Gene cluster
ZMW	Zero Mode Waveguide

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO .....</b>	<b>12</b>
<b>2.</b>	<b>OBJETIVOS .....</b>	<b>15</b>
2.1.	OBJETIVO GERAL .....	15
2.2.	OBJETIVOS ESPECÍFICOS.....	15
<b>3.</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>16</b>
3.1.	SEQUENCIAMENTO DE ACIDOS NUCLEICOS.....	16
3.2.	BANCOS DE DADOS INTEGRADOS.....	21
3.3.	IDENTIFICAÇÃO DE MICROORGANISMOS E RELAÇÕES TAXONÔMICAS .....	23
3.4.	MINERAÇÃO DO GENOMA ATRAVÉS DE FERRAMENTAS IN SILICO.....	24
3.5.	BIOFILME .....	25
3.6.	PGPR (PLANT GROWTH-PROMOTING RHIZOBACTERIA).....	27
3.7.	<i>BACILLUS</i> spp.....	28
<b>4.</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>29</b>
<b>5.</b>	<b>CONCLUSÃO.....</b>	<b>50</b>

## 1. INTRODUÇÃO

A exploração de microrganismos de vida livre do solo na busca de novas estratégias para controle de doenças e pragas na agricultura vem ganhando força na última década. Esses microrganismos conhecidos como PGPR (*Plant-growth Promoting Rhizobacteria*) auxiliam no crescimento da planta de maneira direta e indireta, direta através da síntese de fitoestimulantes que exercem o papel dos fitormônios [1] e indireta através do controle contra fitopatógenos a partir da síntese de moléculas com atividade antimicrobiana e indução de sistemas de defesa na planta [2].

Bactérias dos gêneros *Bacillus* e *Pseudomonas* já são conhecidas pelo seu potencial de colonização rizosférica e produção de biomoléculas com potencial antifúngico e de promoção de crescimento, o que fazem com que estes grupos sejam utilizados como agentes de biocontrole em alguns produtos [2].

O gênero *Bacillus* é amplamente explorado como agente de biocontrole devido a sua importante atividade antimicrobiana contra fitopatógenos e promoção de crescimento da planta bem como indução de resistência contra estes [3]–[5]. Devido a sua capacidade de formar endósporos, bioformulações que contenham este microrganismo acabam se tornando mais estáveis e resistentes ao longo do tempo, podendo manter sua eficiência depois de um período prolongado de prateleira ou mesmo após sua aplicação no solo que pode passar por estresses bióticos e abióticos de diferentes tipos como estresse hídrico, mudanças bruscas de pH e temperaturas elevadas. Espécies do gênero *Bacillus* também possuem registros de atividade nematocida, um problema que também gera um prejuízo elevado em diferentes cultivares [6].

*Bacillus licheniformis*, *B. megaterium*, *B. mucilaginous*, *B. pumilus*, *B. subtilis*, *B. subtilis* var. *amyloliquefaciens* são as principais espécies deste amplo gênero que são utilizadas comercialmente como PGPR [7]. Devido ao aumento na exploração deste grupo se percebeu a complexidade das relações filogenéticas entre bactérias do grupo *Bacillus subtilis* como *Bacillus subtilis*, *B. velezensis*, *B. amyloliquefaciens*, *B. pumilus* e *B. methylotrophicus*. Análises comparativas mais finas mostram que a comparação do DNA ribossomal 16S para distinção entre essas espécies se mostra ineficiente haja visto que há uma similaridade tanto fenotípica quanto genotípica muito grande entre as bactérias deste grupo, fazendo assim com que novas estratégias de comparação sejam

1 necessárias para definir a espécie de um isolado deste gênero, tais estratégias  
2 requerem os dados do genoma completo destes microrganismos [8]–[11]  
3 ilustrando a importância da evolução das tecnologias de sequenciamento ao  
4 longo dos anos.

5 O avanço nas tecnologias de sequenciamento bem como a diminuição do  
6 seu custo possibilitou que um número muito maior de genomas sejam  
7 sequenciados com maior qualidade e cobertura, fazendo com que a quantidade  
8 de dados genéticos relacionados a microrganismos aumentasse  
9 exponencialmente ao longo dos últimos anos [12]–[15], chegando a  
10 aproximadamente 4699 genomas sequenciados deste gênero (aqui incluem  
11 *contigs*, *scaffolds* e cromossomos completos fechados) e enviados à base de  
12 dados NCBI até o dia 03/12/2020.

13 O desafio para os bioinformatas é tratar esta grande quantidade de dados  
14 provenientes de sequenciamentos completos de genomas e montar sequências  
15 que representam o cromossomo circular completo da bactéria. Muitos dos  
16 genomas depositados, devido muitas vezes a limitações do processo de  
17 sequenciamento, não são fechados e possuem lacunas em suas sequências,  
18 são enviados aos banco de dados em forma de *contigs* ou *scaffolds*, o que limita  
19 em partes os avanços de estudos *in silico* do isolado principalmente para  
20 análises voltadas à relações filogenômicas onde se comparam regiões  
21 conservadas e não conservadas e suas posições [16]–[18]. Além disso *contigs*  
22 ou *scaffolds* desordenados podem não conter todas as informações visto que  
23 alguns genes, ou parte deles, podem estar contidos nestas lacunas. Diferentes  
24 estratégias são exploradas para buscar o fechamento do genoma para que  
25 contenha toda a informação sequenciada com qualidade.

26 O aumento na disponibilidade destas informações genéticas possibilitou  
27 também uma maior exploração do genoma de isolados deste gênero com  
28 relação a produção de metabólitos complexos oriundos de seu metabolismo  
29 secundário que possuem atividades antimicrobianas, ecológicas e fisiológicas de  
30 elevada importância [4], [19], [20]. Estes metabólitos secundários se apresentam  
31 no genoma em forma de BGCs (*Biosynthetic Gene Clusters*) e uma parte  
32 considerável do genoma é dedicada a esses metabólitos [4]. Dentro do gênero  
33 *Bacillus* os principais grupo de moléculas produzidas são peptídeos  
34 antimicrobianos [4], Surfactinas [21]–[24], Fengicinas [24]–[28], Iturinas [29] e

1 Bacteriocinas [30]. Estudos em larga escala visando compreender a diversidade  
2 das BGCs e seus papéis na interação e desenvolvimento destas bactérias  
3 sugerem que as BGCs mais conservadas estejam relacionadas com regulações  
4 fisiológicas do microrganismo, enquanto os *clusters* menos conservados e mais  
5 diversos estariam ligados à interação com outros organismos encontrados nos  
6 mais diversos ambientes que este gênero já foi descrito [19].

7

## 1 **2. OBJETIVOS**

### 2 **2.1. OBJETIVO GERAL**

3 Avaliar o potencial biotecnológico de *Bacillus velezensis* cepa LABIM22  
4 sequenciando seu genoma, explorando seu metabolismo secundário e  
5 minerando-o em busca por genes ligados a regulação/formação de biofilme.

6

### 7 **2.2. OBJETIVOS ESPECÍFICOS**

8 - Sequenciar e montar o genoma de LABIM22;

9 - Anotar o genoma;

10 - Analisar os metabólitos secundários através do antiSMASH;

11 - Realizar comparações filogenômicas com outras espécies de *Bacillus*;

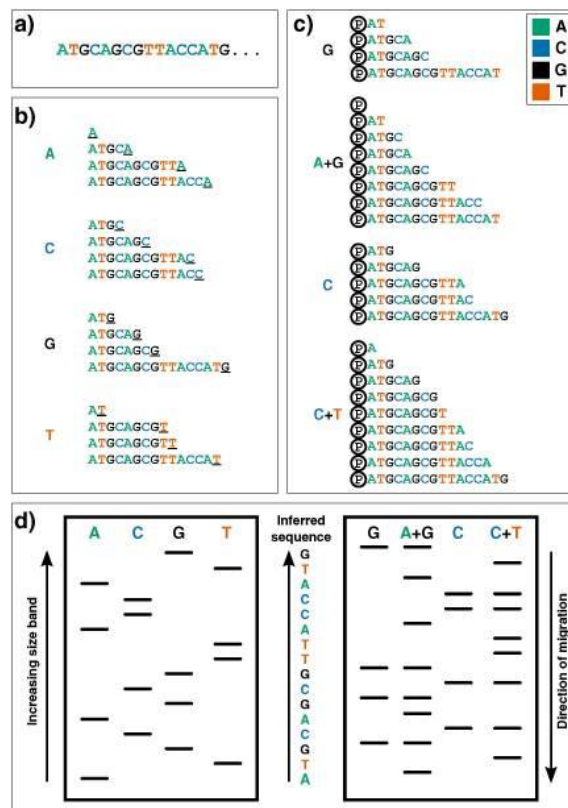
12 - Minerar o genoma buscando genes ligados a regulação/formação de  
13 biofilme.

14

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1. SEQUENCIAMENTO DE ACIDOS NUCLEICOS

Por volta dos anos 70 foram descritas as primeiras técnicas de decodificação do DNA [31], sendo assim possível ler a sua sequência de nucleotídeos e compreender melhor os ácidos nucleicos. A primeira geração de sequenciamentos começou com o método de Sanger e posteriormente aprimorada com os trabalhos de Maxam e Gilbert [31], esta utiliza quatro reações diferentes para cada um dos quatro nucleotídeos, a reação gera fragmentos de diferentes tamanhos pois este método utiliza nucleotídeos comuns e didesoxinucleotídeos, estes quando adicionados a uma reação de polimerase impossibilitam sua continuação devido a uma hidroxila em sua parte terminal 3', são formados fragmentos de diferentes tamanhos, estes são então colocados em uma corrida em gel de poliacrilamida e separados por tamanho, são realizadas 4 corridas, uma para cada didesoxinucleotídeo (A, T, G e C), com isso a sequência pode ser lida de “baixo pra cima” no gel por diferença de tamanho dos fragmentos (Figura 1).



**Figura 1.** ilustrando a leitura dos dados do sequenciamento de Sanger e Maxam Gilbert. Fonte: [32].

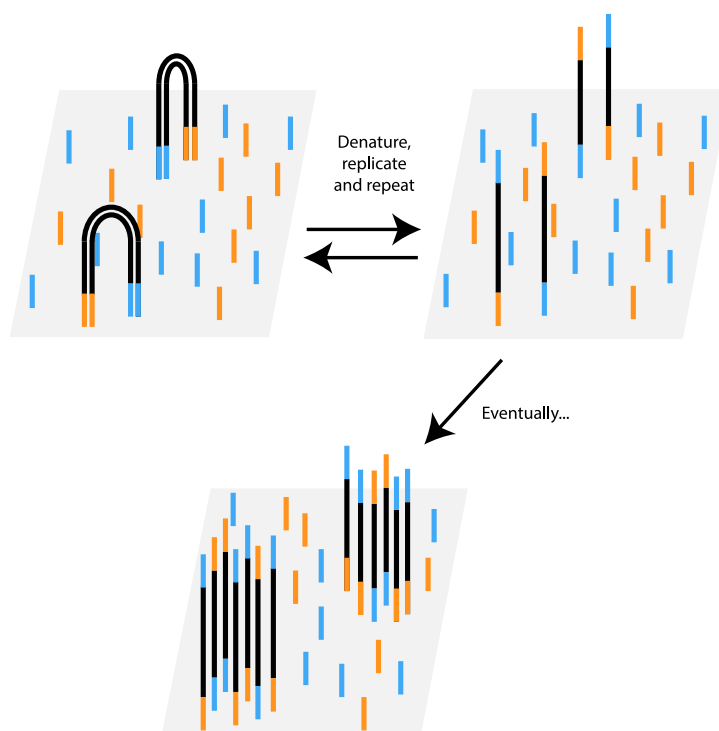
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

Esta tecnologia foi o deu início aos estudos de sequencias genéticas e propiciou subsídio para importantes estudos na época. Mesmo com o passar dos anos os princípios desta técnica ainda são utilizados para sequenciamentos de pequenas regiões e genes já conhecidos. Técnicas mais recentes que utilizam os princípios do sequenciamento de Sanger/Maxam Gilbert sem uso de marcadores radioativos [33] substituídos por agente fluoróforo que emitem picos quando excitados com um laser e detectados por um aparelho, associando isso a tecnologia de eletroforese em capilares foi possível diminuir o número de reações de 4 para 1 uma vez que a máquina registra a ordem e intensidade destes picos fluorescentes [34], cada um dos quatro nucleotídeos emite um comprimento de onda diferente podendo então diferenciar qual foi adicionado a reação e conseqüentemente a sequência de nucleotídeos [35]. Estes avanços permitiram aumentar consideravelmente a quantidade de dados gerada (ex: tamanho das reads) e futuramente foram associadas às técnicas de *shotgun sequencing* que fazem a leitura de vários fragmentos criados de maneira aleatória através de enzimas de digestão e clonados em vetores para criação de uma biblioteca, a sequência estudada é sequenciada de maneira aleatória e montada em computadores que buscam por sobreposições das fitas aleatoriamente geradas e clonadas [32], o primeiro rascunho do genoma humano foi obtido através desta técnica.

Com os avanços tecnológicos crescendo exponencialmente as técnicas de sequenciamento começaram a associar reações de PCR aos fragmentos de DNA escolhidos para sequenciamento, facilitando a detecção sinais emitidos de acordo com a tecnologia usada, são os chamados sequenciamentos de segunda geração. O pirosequenciamento realiza reações de PCR em pequenos fragmentos de DNA ligados a uma *bead* que se encontra dentro de um poço em uma placa, a luz é gerada através da enzima Luciferase que utiliza ATP resultado da adição de dNTP na fita sintetizada, 4 reações diferentes são realizadas com essas *beads* com lavagens entre elas, o sequenciador observa todas as reações de maneira independente levando em conta a intensidade da luz emitida e em qual momento ela foi detectada, gerando assim informações sobre a sequência em cada *bead* [36]. Esta tecnologia foi introduzida no mercado pela empresa 454 Life Sciences, posteriormente adquirida pela Roche.

1           Depois de anos de pesquisa e aperfeiçoamento a técnica da 454 foi  
2 melhorada associando a tecnologia de semicondutor de metal-óxido  
3 complementar utilizada na fabricação de circuitos integrados e processadores,  
4 com isso foi possível medir a alteração do pH no poço com a *bead* devido a  
5 liberação de íons H<sup>+</sup> durante a adição de nucleotídeos na fita molde tornando o  
6 processo mais rápido e eficiente [37]. Contudo mesmo com a adição da  
7 tecnologia de semicondutores aos aparelhos (454), ainda se tinha dificuldade  
8 para identificar sequencias de homopolímeros (mesmo nucleotídeo repetidas  
9 vezes em seguida), a intensidade do sinal em casos de muitas repetições não  
10 fornecia um dado seguro sobre a quantidade de nucleotídeos repetidos.

11           A plataforma Illumina utiliza um princípio similar de sequenciamento por  
12 síntese, associando reações de PCR em ponte com a detecção de luz por cada  
13 nucleotídeo adicionado a fita sintetizada (Figura 2). Os fragmentos selecionados  
14 para sequenciamento através da criação de uma biblioteca a partir da digestão  
15 do DNA total são ligados a adaptadores nas duas extremidades, o adaptador da  
16 extremidade 3' é diferente do encontrado na extremidade 5' para auxiliar na  
17 identificação da orientação da fita, estes adaptadores são oligonucleotídeos que  
18 possuem complementariedade com adaptadores conectados a superfície da  
19 placa (*flow cell*), uma vez ligados a placa começa o processo de PCR em ponte  
20 que amplifica o fragmento de DNA através da enzima DNA polimerase, esta atua  
21 assim que uma extremidade da fita se liga ao adaptador correspondente na  
22 placa, deixando o fragmento com uma extremidade livre e com um sítio de  
23 iniciação para a DNA polimerase livre, o aparelho controla a hibridização dos  
24 adaptadores através da temperatura para que novas reações com os novos  
25 fragmentos sintetizados possam ser realizados. De início os fragmentos com a  
26 extremidade 5' livres são lavados para que apenas sequências da mesma  
27 orientação estejam na placa, com isso a reação de adição de nucleotídeos  
28 fluorescentes a fita molde se inicia, cada adição gera um pico de luz que é  
29 amplificado devido à grande quantidade de cópias próximas e detectado pelo  
30 aparelho [15].



1

2 **Figura 2.** PCR em ponte realizado em plataforma Illumina de segunda geração  
 3 para amplificação dos fragmentos. Fonte: Binf Snipcademy

4

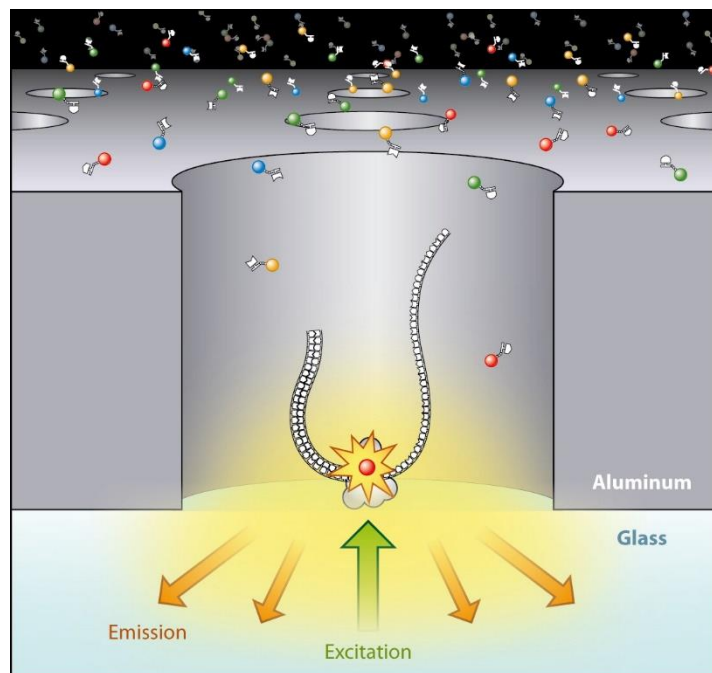
5 A terceira geração de sequenciadores tem contradições quanto a sua  
 6 definição, a mais aceita envolve a capacidade de gerar dados sobre a sequência  
 7 a partir de uma única molécula e sem o uso de tecnologias de amplificação [32].  
 8 Atualmente são 3 as principais plataformas de sequenciamento de terceira  
 9 geração, Pacific Biosciences (PacBio) Single Molecule Real Time (SMRT), o  
 10 Illumina Tru-seq Synthetic Long-Read technology e o Oxford Nanopore  
 11 Technologies sequencing platform.


12

13 PacBio utiliza uma estratégia que não envolve amplificação de material e  
 14 aumenta a detecção do sinal de luz utilizando uma tecnologia chamada *Zero-  
 15 Mode Waveguide* (ZMW) (Figura 3) que atenua pequenos sinais de luz através  
 16 de uma câmara onde não há nenhuma interferência de luz externa e apenas a  
 17 luz emitida dentro dela é detectada, com isso a reação que ocorre dentro desta  
 18 câmara é facilmente detectada e realizada com uma quantidade menor de  
 19 amostra, além de gerar leituras maiores que os sequenciamentos de segunda  
 20 geração [38]. As leituras desta plataforma fornecem *reads* de até 10.000 pb [38]  
 que auxiliam na montagem de genomas produzindo *contigs/scaffolds* grandes

1 que resolvem problemas de montagem com relação a regiões repetitivas nos  
2 genomas [39].

3 Como esta plataforma ainda apresenta limitações como maior número de  
4 erros e baixa cobertura, estratégias híbridas são adotadas associando técnicas  
5 de segunda e terceira geração para fechamento de genomas.  
6

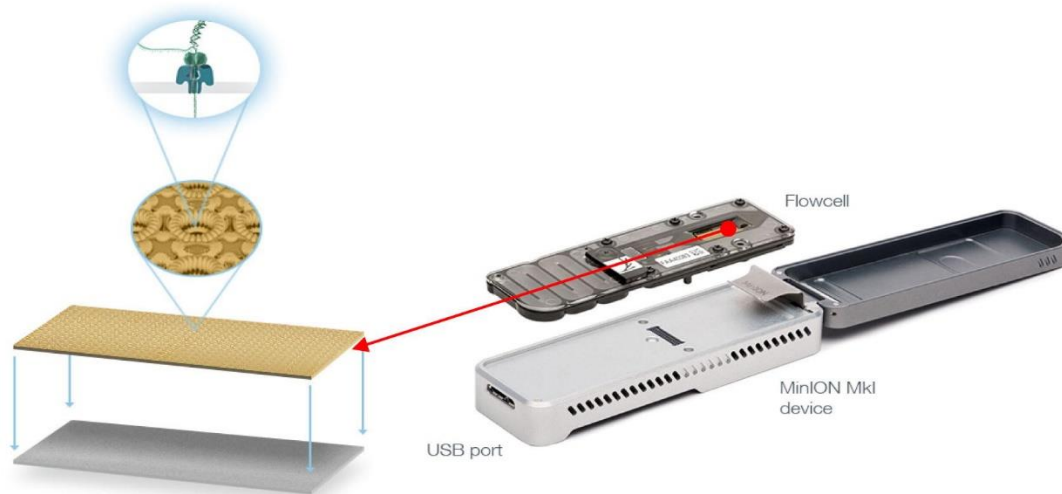


7  Zhu P, Craighead HG. 2012.  
8 Annu. Rev. Biophys. 41:269–93

9 **Figura 3.** Representação gráfica dos poços de detecção de luz usando a  
10 tecnologia ZMW (Zero Mode Waveguide), o diâmetro do poço é menor que o  
11 comprimento de onda de luz visível diminuindo problemas como ruído de fundo  
12 fazendo com que o sinal de luz emitido por apenas uma molécula de DNA seja  
13 detectado com mais facilidade [40]

14 Oxford Nanopore MinÍon troca a detecção de sinais de luz pela detecção  
15 da diferença de fluxo de Íons através de um nanoporo ligado a uma membrana  
16 por onde passa a molécula de DNA ou RNA, cada nucleotídeo altera de maneira  
17 individual a passagem de íons pelo nanoporo, gerando assim informações sobre  
18 a sequência que o atravessa. Dentre as vantagens desta plataforma temos a  
19 mobilidade e rapidez de obtenção de dados, o sequenciamento é realizado em  
20 um aparelho portátil pesando cerca de 90g e analisa os dados gerados em tempo

1 real ligado via USB a algum dispositivo [41], [42], auxiliando na velocidade dos  
 2 resultados, importantes em casos clínicos [43]. Leituras podem chegar a 150  
 3 kilobases, auxiliando na montagem de regiões repetitivas [41] (Figura 4).



4  
 5 **Figura 4.** Oxford Nanopore Minlon, em aumento a membrana com o nanoporo  
 6 associado, a *flow cell* contém 2048 destes nanoporos. Esta tecnologia se  
 7 destaca pela portabilidade e rapidez no fornecimento de dados.

8

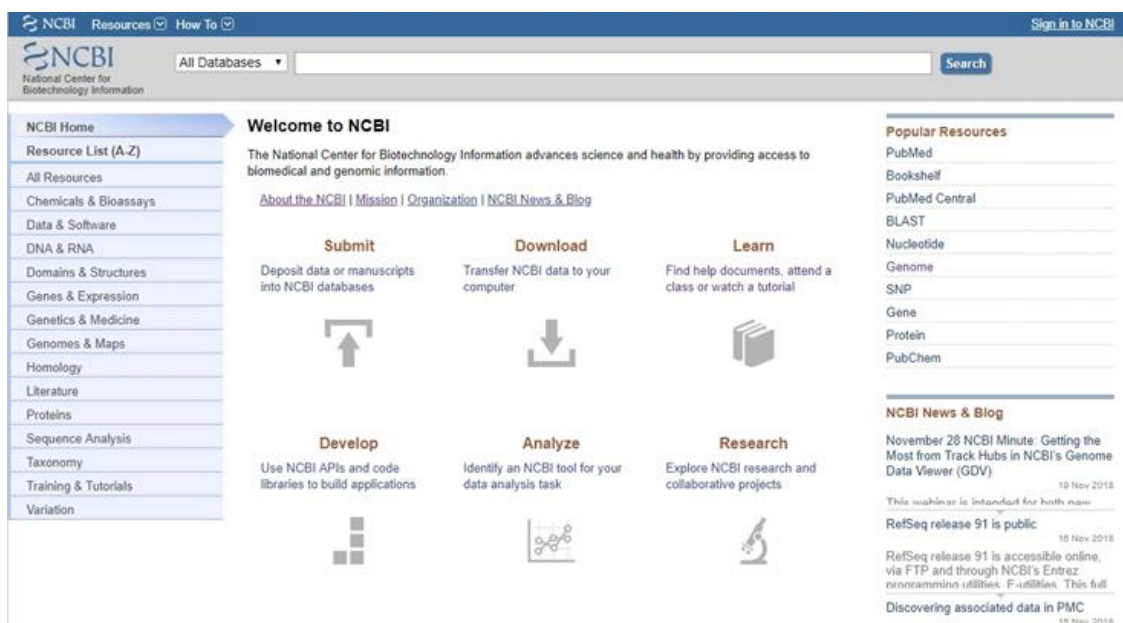
### 9 **3.2. BANCOS DE DADOS INTEGRADOS**

10 Devido a elevada quantidade de informação gerada com relação as  
 11 seqüências de nucleotídeos e aminoácidos nas últimas décadas foi necessário  
 12 a criação de bases de dados integradas para facilitar o acesso a essas  
 13 informações. O GenBank é um banco de dados de biologia molecular disponível  
 14 on-line disponível em uma das divisões da National Center for Biotechnology  
 15 Information (NCBI) (Figura 5), que associado European Molecular Biology  
 16 Laboratory Nucleotide Sequence Database (EMBL) e o DNA Databank of Japan  
 17 (DDBJ) formam o International Nucleotide Sequence Database Collaboration  
 18 (INSDC) (Benson et al. 2007). Esta base de dados compreende dados  
 19 moleculares de projetos do mundo todo, que vão desde pequenos fragmentos  
 20 de DNA ou cadeias peptídicas até genes e operons completos além de  
 21 seqüências de genoma completo de vários grupos de seres vivos como animais,  
 22 plantas e microrganismos [44].

23 Uma ferramenta amplamente utilizada por pesquisadores no mundo todo  
 24 é o Basic Local Alignment Search Tool (BLAST) (Figura 6). O BLAST é um  
 25 algoritmo de busca de similaridade que utiliza a base de dados fornecida pelo

1 NCBI para buscar semelhanças em sequências de nucleotídeos e aminoácidos  
 2 alinhando a sequência desejada (query) com todas as sequências que  
 3 apresentem alguma similaridade na base de dados [45]. Esta ferramenta se  
 4 mostrou de extrema importância pois possibilitou que sequências obtidas  
 5 possam ser comparadas com dados de outros estudos auxiliando na exploração  
 6 dos resultados.

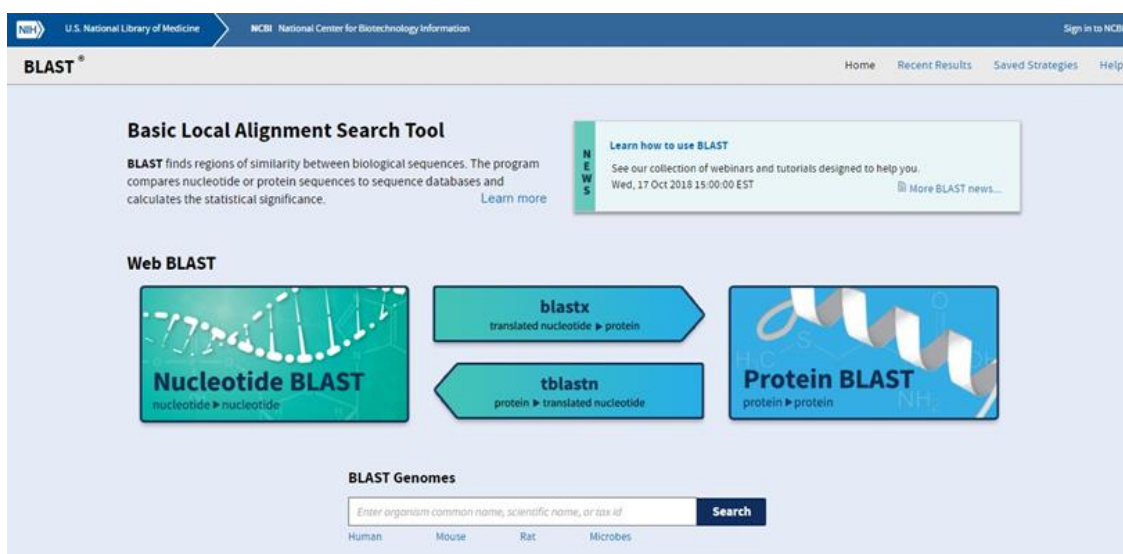
7



8

9 **Figura 5.** Layout do site do NCBI onde se encontram todas as informações a  
 10 respeito do GenBank.

11



12

13 **Figura 6.** Layout da plataforma on-line BLAST para comparação de sequências  
 14 de nucleotídeos e aminoácidos.

### 3.3. IDENTIFICAÇÃO DE MICROORGANISMOS E RELAÇÕES TAXONÔMICAS

Uma das tarefas mais importantes ao se estudar algum microrganismo é a sua correta identificação e posição taxonômica. As classificações mais antigas se baseavam apenas em características fenotípicas de crescimento do microrganismo e testes bioquímicos [12]. Contudo, com o advento das tecnologias de sequenciamento, ferramentas que estabeleciam relações com bases genômicas ganharam mais espaço na pesquisa e são o padrão para construção de árvores filogenéticas.

Uma das primeiras análises comparativas de genoma foi a DNA-DNA Hybridization (DDH) que consiste na quantificação da energia necessária para dissociar duas fitas de DNA de origens diferentes, quanto mais forte a ligação maior à similaridade entre os dois genomas, logo a energia necessária para separar as duas fitas é maior. Valores de similaridade entre dois isolados acima de 70% indicam que os mesmos são da mesma espécie [46].

Um problema enfrentado por esta técnica é incapacidade de deposição de dados em bancos de livre acesso, o que vai na contramão do que a era da bioinformática vem apresentando mostrando que a técnica precisou ser substituída [46]. Devido a esse fato, abordagens como a comparação do gene da subunidade 16S do ribossomo, que pode ser depositado e utilizado para comparações, a utilização da ANI que representa a média de identidade dos nucleotídeos entre os isolados estudadas, se apresentam como bons substitutos para a técnica de DDH.

Outra característica genética utilizada para identificação de isolados é a comparação do gene responsável pela transcrição da subunidade 16S do ribossomo, valores de similaridade de 95% a 97% quando comparadas duas regiões estabelece as duas como sendo da mesma espécie [46]. A região do rDNA 16S é um fragmento altamente conservado entre os procariotos do mesmo gênero e espécie [47], esta região se repete pelo genoma podendo ocorrer repetições seguidas entre si ou espaçadas no genoma. Por ser uma região repetitiva sua amplificação é possível através do desenho de primers que flanqueiam a região para que ocorra amplificação do fragmento, o que facilita seu sequenciamento. O gene responsável pela transcrição dessa região é

1 formado por regiões conservadas intercaladas por segmentos hipervariáveis que  
2 ilustram a diversidade entre os procariotos. Levando em consideração a  
3 comparação desse gene para construção de árvores filogenéticas chegou-se a  
4 divisão dos três principais grupos de seres vivos, Procariotos, Eucariotos e  
5 Archea [48]. O uso desta região não se limitou apenas a fins de pesquisa e é  
6 também utilizada na identificação de isolados em laboratórios de microbiologia  
7 clínica para fornecer tratamentos com antibióticos mais direcionados e com  
8 maior eficiência [47].

### 9 10 **3.4. MINERAÇÃO DO GENOMA ATRAVÉS DE FERRAMENTAS *IN SILICO***

11 Diversos estudos em relação a diversidade genética presente nos grupos  
12 de procariotos vêm sendo desenvolvidos, ao longo dos anos essa diversidade  
13 vem sido aproveitada na forma de descoberta de produtos naturais de origem  
14 microbiana com atividade antimicrobiana. Porém com o advento das tecnologias  
15 de sequenciamento de alta cobertura e com preços cada vez mais acessíveis,  
16 vem se descobrindo que a diversidade no que concerne a produção de  
17 metabólitos com potencial biotecnológico presente nos procariotos é muito maior  
18 do que se imaginava [13], nesse contexto se tornou necessário o  
19 desenvolvimento de ferramentas que utilizam essa quantidade de dados  
20 genéticos que vem se acumulando exponencialmente para busca de novas  
21 moléculas que possam ser utilizadas como novos antimicrobianos.

22 A busca por essas moléculas se dá através de programas que buscam  
23 regiões conservadas do genoma relacionadas a tradução de enzimas ligadas a  
24 síntese de moléculas já conhecidas [49]. Estes genes geralmente estão  
25 organizados em clusters, agrupados em uma região específica do genoma  
26 contendo todos os genes necessários para a síntese da molécula. Os programas  
27 disponíveis para mineração dos genomas em busca de clusters ligados a síntese  
28 de metabólitos utilizam as enzimas conservadas para estimar a posição do  
29 possível cluster e caracterizá-lo de acordo com a classe da molécula que pode  
30 ser produzida.

31 Atualmente programas como o Bayesian Analysis of Gene Essentiality  
32 (BAGEL) [50]–[52] oferecem um serviço de anotação independente que busca  
33 por *Open Reading Frames* (ORFs) conservadas que flanqueiam os genes  
34 ligados a síntese de bacteriocinas. O Antibiotics and Secondary Metabolite

1 Analysis Shell (antiSMASH) (Figura 7) [49] combina diferentes bases de dados  
 2 para fornecer informações sobre BGCs ligadas as metabolismo secundário dos  
 3 genomas analisados. Cluster Scanner (ClusScan) [53], Natural Product Domaine  
 4 Seeker (NapDoS) [54], Secondary Metabolites Unique Region Finder (SMURF)  
 5 são especializados na busca por clusters em genomas de fungos [55] e Natural  
 6 Products Searcher (NP.searcher) [56] que também minera genomas na busca  
 7 por possíveis produtos naturais de interesse no genoma. Todas estas  
 8 ferramentas estão disponíveis para uso na forma de web-server ou programa  
 9 para Desktop e são excelentes ferramentas para auxiliar na busca por BGCs nos  
 10 genomas microbianos sequenciados.  
 11

12  
 13 **Figura 7.** Web Server antiSMASH utilizado para busca de BGCs ligados ao  
 14 metabolismo secundário do genoma desejado.

15

### 16 3.5. BIOFILMES

17 Biofilmes são comunidades bacterianas aderidas a alguma superfície e  
 18 embebidas em uma matriz extracelular. Estes agregados celulares se formam  
 19 nas mais diferentes superfícies, orgânicas ou inorgânicas. Podem ser benéficos  
 20 ou maléficos dependendo do local onde se estabelecem [57].

21 Dentre os diferentes substratos nos quais são possíveis a formação de  
 22 biofilme, alguns chamam a atenção como raízes e solos rizosféricos. Plantas

1 liberam através de sua raiz uma série de compostos de alta diversidade com  
2 diferentes funções tanto para a planta quanto para outros organismos [58], estes  
3 exsudados da planta servem como substrato para o crescimento de bactérias e  
4 outros microrganismos criando um microambiente altamente diverso neste solo.  
5 A formação de biofilme criando um ambiente mais estável para o  
6 estabelecimento das bactérias é de extrema importância e vital para o sucesso  
7 desta relação oferecendo proteção contra predação, troca de metabólitos e  
8 possibilidade de transferência horizontal de genes [59].

9 Esta estratégia de colonização se mostra eficiente uma vez que a  
10 estrutura se estabelece no substrato aderido e se estabiliza, este estado  
11 fisiológico do biofilme passa a ter características únicas que conferem defesa  
12 contra antibióticos [60], a matriz extracelular secretada pelas próprias bactérias  
13 funciona como uma barreira física [57] e de difusão contra pequenas moléculas  
14 e possíveis antimicrobianos.

15 Características importantes para a formação de biofilme e colonização  
16 rizosférica são motilidade (1), capacidade de adesão a diferentes substratos  
17 através de moléculas do tipo adesinas presentes na parte externa da célula (2),  
18 alta produção de EPS (exopolissacarídeo) (3) e produção de moléculas ligadas  
19 com a sinalização intercelular (*quorum-sensing*) como N-Acyl homoserina (4)  
20 [59], [61], [62].

21 A capacidade de motilidade confere ao microrganismo a possibilidade de  
22 buscar por meio da quimiotaxia locais com maior disponibilidade de nutrientes  
23 na raiz visto que a distribuição dos exsudados não ocorre de forma homogênea  
24 [59], além de aumentar a área de distribuição do procaríoto. Esta motilidade pode  
25 ocorrer através do *swarming* ou com o auxílio de flagelos, mutantes sem  
26 capacidade sintetizar flagelos apresentam dificuldade na colonização de raízes  
27 [63].

28 Considerando o grupo das bactérias, as principais estruturas envolvidas  
29 na adesão a um substrato inerte são as fimbrias e os pili, estas estruturas  
30 também possuem papel na agregação celular, característica importante na  
31 formação de biofilme e conseqüentemente colonização de raízes [59], [64]. As  
32 ricadesinas são moléculas essenciais ligadas a adesão celular em *Rhizobium*  
33 *leguminosarum* a substratos inertes [65].

34 A comunicação entre as células se mostra importante pois contribui na

1 coordenação das respostas fisiológicas como um grupo e também modula as  
2 comunicações com outros organismos, importante para o estabelecimento do  
3 biofilme [58], [62].

4 Produção de compostos extracelulares como exopolissacarídeos está  
5 intimamente ligada ao sucesso de crescimento desta comunidade bacteriana,  
6 auxiliando na motilidade devido a liberação de peptídeos como a surfactina que  
7 ajuda na quebra de tensão superficial do meio provendo uma maior capacidade  
8 de movimentação [66]. Devido a essa produção de matriz extracelular, a difusão  
9 de nutrientes através do biofilme é limitada, fazendo com que algumas células  
10 não recebam a mesma quantidade de nutrientes do que outras, criando assim  
11 um ambiente heterogêneo dentro do biofilme com células geneticamente  
12 idênticas mas com expressão genética e atividades metabólicas diferentes [67].  
13 Esta heterogeneidade faz com que algumas células possuam funções  
14 específicas dentro da comunidade, como as células persistentes, que não se  
15 dividem e estão diretamente ligadas a resistência do biofilme a antibióticos.  
16 Algumas células sintetizam flagelos e estão intimamente ligadas a expansão  
17 espacial da matriz extracelular.

18 A formação do biofilme está ligada a respostas fisiológicas das bactérias  
19 que entram em metabolismo secundário, produtos deste metabolismo estão  
20 relacionados com a indução de produção de EPS e diferenciação das células  
21 dentro da matriz.

### 22 23 **3.6. PGPR (PLANT GROWTH-PROMOTING RHIZOBACTERIA)**

24 Ao observar a diversidade microbiana presente nos mais diferentes solos  
25 alguns microrganismos chamam a atenção pela sua capacidade de promover  
26 crescimento, induzir sistemas de defesa e até mesmo defender a planta  
27 diretamente contra fitopatógenos, são os chamados PGPR (Plant Growth  
28 Promoting Rhizobacteria) [22], [68]. Esta comunicação só é possível devido à  
29 alta diversidade de compostos de baixo peso molecular liberados pelas raízes  
30 das plantas criando um ambiente complexo que possibilita interação com vários  
31 outros macro e microrganismos [58]. Esta diversidade microbiana encontrada no  
32 solo rizosférica com tais capacidades é alta porém dominada por bactérias do  
33 gênero *Bacillus* spp. e *Pseudomonas* spp. [69].

1 Os microrganismos podem atuar de diferentes maneiras promovendo o  
2 crescimento da planta de maneira direta, sintetizando compostos que a planta  
3 reconhece como fitormônios que vão atuar nas vias de sinalização de  
4 crescimento da planta, ou de maneira indireta facilitando a absorção de  
5 diferentes compostos que naturalmente não estão disponíveis para a planta ou  
6 se encontram em concentrações muito baixas [7]. Outra atuação destes é a  
7 proteção contra fitopatógenos também de maneira direta ou indireta produzindo  
8 moléculas com potencial antimicrobiano ou nematicida que acabam por inibir o  
9 crescimento destes fitopatógenos nas regiões próximas as raízes, ou pela  
10 ativação de sistemas de defesa da própria planta por sinalização molecular em  
11 casos de presença de patógenos [7].

12 A antibiose se dá devido a síntese de moléculas com atividade pelos  
13 microrganismos presentes nos solos rizosféricos, o grupo *Bacillus subtilis* chama  
14 a atenção por fazer parte da microbiota rizosférica e apresentar atividade  
15 antifúngica já descrita em trabalhos anteriores [5], [70]–[72]. Bactérias deste  
16 gênero apresentam uma elevada diversidade de compostos oriundos de seu  
17 metabolismo primário e secundário com diferentes funções ecológicas [4], [73] e  
18 que já são exploradas como mecanismos de biocontrole [25], [30], [72].

19 Isolados que apresentam estas capacidades (promoção de crescimento e  
20 defesa contra fitopatógenos) são excelentes candidatos a possíveis agentes de  
21 biocontrole já quem possuem mais de uma estratégia para auxiliar no  
22 desenvolvimento da planta oferecendo diferentes alternativas e alvos de atuação  
23 no controle, sendo assim importantes fatores para o desenvolvimento de uma  
24 agricultura sustentável [69].

25

### 26 **3.7. *Bacillus* spp.**

27 O gênero *Bacillus* compreende bactérias aeróbicas Gram-positivas com  
28 formato de bastonetes e capacidade de formar endósporos, foi descrito pelo  
29 primeira vez por Cohn em 1872 [74]. Bactérias deste gênero são encontradas  
30 nos mais diversos ambientes e são exploradas com relação ao seu potencial  
31 biotecnológico a muitos anos. Sua capacidade de produzir enzimas de interesse  
32 biotecnológico em altas concentrações vem sendo utilizadas por diferentes  
33 setores da indústria, sua elevada diversidade de metabólitos e baixos casos de  
34 patogenicidade descritos auxiliam no uso na indústria, alguns exemplos de uso

1 são: síntese de antibióticos, enzimas, potenciadores de sabor na indústria  
2 alimentícia e inseticidas [74].

3 Uma característica que distingue este grupo e o torna interessante para  
4 suas variadas utilizações é a sua capacidade de alterar seu metabolismo e  
5 formar endósporos, formas de resistência que ajudam na permanência da célula  
6 por longos períodos de estresse dos mais variados tipos [75]. Esta estratégia é  
7 uma resposta ao meio onde a bactéria se encontra, em casos de falta de  
8 nutrientes, elevada radiação ou temperatura, alterações bruscas de pH e outros  
9 fatores bióticos ou abióticos a célula sai do seu metabolismo vegetativo e diminui  
10 seu metabolismo a níveis quase indetectáveis [76]. Esta forma de resistência  
11 com baixa atividade metabólica se mostra a forma ideal para propagação  
12 espacial e temporal do microrganismo, uma vez que mesmo com o metabolismo  
13 baixo este é capaz de monitorar o status nutricional nos seus arredores e reativar  
14 seu estado vegetativo em caso de disponibilidade de fontes de macro e  
15 micronutrientes [76].

16 A presença desta plasticidade fenotípica faz com que endósporos sejam  
17 encontrados em virtualmente todos os ambientes terrestres e muitas vezes longe  
18 de seu local de origem, se mostrando uma estratégia de dispersão  
19 extremamente eficiente e de elevada importância tanto ecológica [19], [77]  
20 quanto para conquista de diferentes ambientes ao longo de sua história  
21 evolutiva.

22 Com a advento de novas tecnologias de sequenciamento a quantidade de  
23 dados gerados ajudou a se ter uma ideia da diversidade metabólica presente nos  
24 mais diversos grupo microbianos [13], o gênero *Bacillus* possui uma diversidade  
25 de moléculas sintetizadas que chama atenção em vários setores da indústria e  
26 também no controle de doenças.

27 O grupo *Bacillus subtilis*, que engloba as espécies *B. subtilis*, *B.*  
28 *velezensis*, *B. amyloliquefaciens*, *B. methylotropicus* [...] apresentam uma  
29 grande capacidade de síntese de moléculas com atividade antifúngica que é  
30 explorada na busca por defensivos agrícolas de origem biológica [26], [71], [78],  
31 [79]. Classes importantes de moléculas já foram descritas em bactérias deste  
32 grupo como Fengicina [26], Macrolactina [80], Bacilibactina [81], Iturinas [29],  
33 Dificidina e Bacilisina [82].

34

## 4. RESULTADOS E DISCUSSÃO

### ARTIGO

International Journal of Genomics

## Genome sequencing of *Bacillus velezensis* LABIM22: a plant growth-promoting and antifungal bacterium

Gustavo M. Teixeira,<sup>1</sup> Ulisses de Padua Pereira<sup>2</sup>, Admilton Gonçalves de Oliveira<sup>1\*</sup>

<sup>1</sup> Department of Microbiology, State University of Londrina, Londrina 86057-970, Brazil.

<sup>2</sup> Department of Preventive Veterinary Medicine, State University of Londrina, Londrina 86057-970, Brazil.

\*Correspondente: Admilton Gonçalves de Oliveira; admilton@uel.br

### Abstract

The *Bacillus velezensis* strain LABIM22 (CMRP 4490) obtained from a soil sample in north of Paraná has shown antifungal activity against important phytopathogenic fungi of agronomical interest, in addition, it was identified as belonging to the *Bacillus* genus, a group that stands out due to its capacity to form endospores and host of diverse machinery of secondary metabolites with different roles, showing with that an elevated potential as a biocontrol agent. To a further understanding of the strain and validate its effect as growth-promoting on plants, it was decided to explore its genetic content through the sequencing and data mining of its genome. The strain LABIM22 has 3.996,396 pb with a GC content of 46,4% and presents 4042 coding sequences. In its genome were found 13 biosynthetic gene clusters related to the synthesis of molecules, some already associated with antifungal activity, besides having genes linked to the regulation/formation of biofilms and motility, important characteristics for rhizospheric colonization. The culture of LABIM22, as coating seed of soybean and maize, improved germination and promoted radicular growth. The data here presented showed that *Bacillus velezensis* LABIM22 could be a potential agent to protect soybean and maize plants from infection by phytopathogenic fungi *in vivo* and had properties as a plant-growth promoting rhizobacteria.

**Keywords:** *Bacillus*, genome, antifungal, PGPR

### Introduction

The exploration of free-living microorganisms from rhizospheric soils has been considered an important strategy for disease control in crops and have been gaining strength in the last decade [1]–[5]. Among these soil microorganisms the PGPR (Plant-growth Promoting Rhizobacteria) draws attention for its ability to help on plant growth as well as induction of defense systems from the plant against phytopathogens [6]. The *Bacillus* genus have been on the spotlight on the use as biocontrol agent due to its characteristics like endospore formation capacity, growth-promoting capacities, induction of systemic defenses against crop diseases [6]–[8] and high diversity of synthesized molecules with antimicrobial activity [4, 9–13]. Biofilm formation is also important when using microorganisms as biocontrol agents as showed in other studies for

1 defense against pathogens and growth-promoting activity [14], [15]. The use of these bacteria  
2 as disease control on crops is already explored in other countries, showing that this genus  
3 possesses a high potential yet to be uncovered in its secondary metabolism. These metabolites  
4 present itself in the form of BGCs (Biosynthetic Gene Clusters) and a substantial part of the  
5 genome is dedicated to these metabolites. Large scale studies aiming to comprehend the  
6 diversity of BGCs and its roles on interaction or development of these microorganism suggest  
7 that conserved BGCs might be related with their physiological regulations, while the less  
8 preserved and more diverse clusters are linked with the interaction with other organisms found  
9 on the most diverse environment that this genus was already found [16], [17]. The present study  
10 was aimed at evaluating the biotechnological potential of *Bacillus velezensis* strain LABIM22  
11 by sequencing its genome, exploring its secondary metabolism for evaluated control efficacy  
12 against phytopathogenic fungi, mining it genes linked to biofilm regulation/formation and for  
13 plant growth-promoting activities. The data in this study provided for the possible development  
14 of LABIM22 as a biocontrol agent.

## 15 **Material and Methods**

16 **Identification of the isolated.** LABIM 22 strain was originally isolated in the soil located at  
17 23°19'26.2"S 51°11'50.5"W, in the State University of Londrina, Paraná, Brazil. The isolated  
18 strain was stored in the Laboratory of Microbial Biotechnology of State University of Londrina,  
19 Paraná, Brazil. The strain was deposited under number CMRP4490 at the Coleções  
20 Microbiológicas da Rede Paranaense (CMRP) of Federal University of Paraná, Curitiba,  
21 Brazil.

22 **Complete genome sequencing and assembly.** For complete genome sequencing LABIM22  
23 was cultivated in L.B. broth (Luria Bertani, Acumedia, USA) at 150 RPM at 28 °C. All the  
24 genetic material was obtained with a DNA extraction kit (Gentra Puregene Genomic DNA kit,  
25 Qiagen Brazil), following that the DNA was purified and quantified. Sequencing was carried  
26 on the MiSeq platform, using a MiSeq version 3 reagent kit (600-cycle, Illumina, Brazil) at  
27 Empresa Brasileira de Pesquisa Agropecuária (Embrapa – Soja), Londrina. The reads quality  
28 as well as the trimming parameters were observed and chosen using FastQC, setting a threshold  
29 Phred score of 30 a series of trimming parameters were used to obtain the best data possible  
30 for *de novo* assembly [18], [19]. A series of *de novo* assemblies were carried out with different  
31 software (SPAdes, Velvet, IDBA hybrid and CLC Genomic Workbench 11) [20], [21] trying  
32 different assembly parameters and then compared with QUAST [22]. The main metrics like  
33 total alignment size, number of contigs, biggest contig, N50 values and numbers of genes  
34 according to the reference genome provided on QUAST were used to choose the better  
35 assembly. Using the webserver CONTIGuator contigs from the better assembly were aligned  
36 with a reference genome [23], raw reads from the sequencing were mapped against the  
37 scaffolds generated to check which ones were usable for the next steps, the ones with low read  
38 count to support the sequence were discarded. Gaps within the scaffold were first treated with  
39 the software GapCloser followed by manual curation with read mapping using Bowtie2 and  
40 gap filling using CLC Genomics Workbench 11 GUI [24]. Genome start was determined by  
41 comparison with the reference strain *Bacillus velezensis* Bac57 assuming the gene *dnaA* as the  
42 first gene.

43 **Representation of circularized genome and secondary metabolites cluster prediction.**  
44 Using the BRIG (BLAST Ring Image Generator) software [25] LABIM22 genome was  
45 represented circularly and compared with other reference genomes. The webserver  
46 antiSMASH makes its predictions combining different databases of genetic data, antimicrobial

1 molecules and BGCs to predict position and possible function of the cluster found [26].  
2 Analysis were carried out with the final *fasta* file of LABIM22.

3  
4 **Phylogenomic comparison and tree.** For species determination of the strain ANI (Average  
5 nucleotide Identity) was determined using orthoANI [27] and dDDH (digital DNA-DNA  
6 hybridization) was calculated with GGDC (Genome-to-Genome Distance Calculator) [28].  
7 Using the Gegenees software whole genome comparisons were made with LABIM22 and other  
8 closely related strains [29]. Data from Genenees was exported to SplitsTree to tree confection  
9 using the neighbor joining method [30].

10  
11 **Genome annotation.** After genome finishing genome annotation was realized using the RAST  
12 platform [31] were CDSs were predicted and classified into subsystems. Metabolic routes were  
13 also created using RAST data on the Model SEED server.

14  
15 **Search for biofilm formation related genes.** Following the genome annotation, a local  
16 BLASTN using a custom database containing the LABIM22 genome was made using as query  
17 sequences a list of genes directly or indirectly related to biofilm formation/regulation found on  
18 SubtWiki and other articles [32]–[37] related to biofilms on *Bacillus subtilis*.

19  
20 **Swarming motility and pellicle formation assay.** Knowing that swarming motility is an  
21 important skill to colonize different environments [15], [38] this ability was determined by the  
22 spread of bacteria on TSA agar plates 0,8%, an overnight culture was prepared and then  
23 adjusted to 0,5 on the MacFarland scale ( $1.5 \times 10^8$  CFU/mL), 10  $\mu$ L of this inoculum was  
24 placed in the middle of a 55 mm petri dish. The ability of colonizing the whole surface of the  
25 plate was evaluated. For pellicle formation a 24-well plate with 2 ml of TSB was prepared and  
26 inoculated with 10  $\mu$ L of an overnight culture adjusted to 0,5 on the MacFarland scale. Digital  
27 pictures were taken to evaluate pellicle formation capacity.

28  
29 **Root colonization capacity with scanning electron microscopy (SEM):** To analyze  
30 LABIM22 capacity of colonizing root and seed surfaces maize seeds were first disinfected by  
31 immersion in 95% (v/v) ethanol solution for 30 seconds, followed by soaking in 5% (v/v) H<sub>2</sub>O<sub>2</sub>  
32 solution for 10 min, and then washed six times with sterile deionized water then treated with  
33 lyophilized cell cultures at three different concentrations (doses: 100, 200 and 300g of  
34 lyophilized cell culture respectively for 100 kg of seeds), for a positive control the commercial  
35 product Presence<sup>®</sup> (FMC - *Bacillus subtilis* strain FMCH002 and *Bacillus licheniformis* strain  
36 FMCH001) was used at a dose of 200g for 100 kg of seeds, negative control was disinfected  
37 seeds only. In a beaker the seeds were treated following the recommendation of 600 mL for  
38 100 kg of seeds, after the application the beaker was agitated manually for 1 minute for  
39 homogenization. The seeds stayed incubated for 7 days disposed on a germination paper  
40 moistened with sterilized distilled water and incubated in a growth chamber at 25 °C with 70%  
41 relative humidity. At day 0 four seeds of each treatment were picked for sample preparation  
42 for SEM, after 7 days 4 germinated seed were picked from each treatment for sample  
43 preparation for SEM focusing on the radicular area. Samples were fixed for 24h then  
44 dehydrated on increasing alcohol concentrations for 15 minutes at each concentration (30, 50,  
45 70, 90 and 100%). Following that they were critical-point-dried with CO<sub>2</sub> (BALTEC CPD 030  
46 Critical Point Dryer), attached to a metal stub and covered with gold (BALTEC SDC 050  
47 Sputter Coater) for SEM visualization (FEI Quanta 200, Netherlands).

48  
49 **Germination and growth-promoting on crops.** For germination analysis and growth  
50 promotion on maize and soybean the following treatments were evaluated: T1: negative control

1 (without application of the cells), T2: commercial product Presence<sup>®</sup> (dose: 200g for 100 kg  
2 of seeds) and T3, T4 and T5: *Bacillus velezensis* strain LABIM22 (doses: 100, 200 and 300g  
3 of lyophilized cell culture respectively for 100 kg of seeds). The maize and soybean cultivars  
4 used for the experiments were BALU 388 VIP3 e BMX Potência RR respectively. For  
5 preparation of lyophilized cells LABIM22 was first cultivated for pre-inoculum preparation on  
6 saline solution (0,85% sodium chloride) and adjusted to 0,5 on the MacFarland scale. To  
7 inoculum preparation 30  $\mu$ L of this solution was added to an Erlenmeyer flask of 125 mL  
8 containing 30 mL of LB broth and incubated at 28 °C for 24h at 125 rpm (Orbital shaker Tecnal  
9 - TE 422, Brazil). For fermentation 4  $\mu$ L of the inoculum was added to 400 mL of Formulated  
10 Medium (patent pending) on 1000 mL Erlenmeyer flasks then incubated for 72h at 28 °C and  
11 200 rpm. All the content of the flasks was frozen at -80 °C then lyophilized at -60 °C to obtain  
12 a powder containing LABIM22 metabolites and spores at a concentration of  $10^{10}$  cells per gram.  
13 For the germination and growth promoting experiments the seeds were initially disinfected by  
14 immersion in 95% (v/v) ethanol solution for 30 seconds, followed by soaking in 5% (v/v) H<sub>2</sub>O<sub>2</sub>  
15 solution for 10 min, and then washed six times with sterile deionized water. Following that was  
16 realized the inoculation of the seeds, in a beaker the seeds were treated with 3 different  
17 concentrations of lyophilized cells following the recommended of 600 mL for 100 kg of seeds,  
18 after the application the beaker was agitated manually for 1 minute for homogenization. For  
19 germination evaluation, 50 seeds per treatment were placed on germination paper moistened  
20 with sterilized distilled water and incubated in a growth chamber at 25  $\pm$ 2 °C and 70% relative  
21 humidity for 7 days. For the experiment on germination paper, a completely randomized design  
22 with four replications was used. For growth promoting evaluation the treated seeds were spread  
23 on 900 mL Styrofoam pots with sand, leaving three plants for each pot. Fifteen days after  
24 sowing the seedlings were scanned at 300 dpi and the images treated and analysed with the  
25 software GiA Roots [39]. The total root surface (cm<sup>3</sup>) and total root length (mm) were  
26 evaluated. After, the shoot part and root system of the seedlings were oven-dried separately  
27 under forced ventilation at 60 °C during 72h to determine shoot dry mass (g) and root dry mass  
28 (g), For the experiment on growth-promoting, a completely randomized design with six  
29 replications was used. All data was submitted to variance analysis and afterwards to means  
30 comparison with Tukey test. All analysis were realized using R (<http://www.r-project.org>)  
31 with the *agricolae* package [40].

32

## 33 **Results and Discussion**

34 **Genome assembly and annotation.** Following the proposed strategy CLC Genomics  
35 Workbench 11 and IDBA Hybrid have showed the better and similar results concerning  
36 assembly metrics (Figure 1).

1

Estatísticas dos genomas	IDBA k-mer 100	IDBA k-mer 80	Velvet k-mer 45	CLC	SPAdes k-mer 127	SPAdes k-mer 99
# Genes, transcritos e CDS	4035 + 57 part	3933 + 63 part	3829 + 95 part	3986 + 65 part	3189 + 45 part	2846 + 34 part
# contigs (>= 500 bp)	98	94	133	100	55	56
Maior contig	452642	452642	187930	254617	319575	319575
Tamanho total alinhado	3993000	3977919	3929659	3985628	3106362	2740950
N50	147973	129667	59664	84162	133096	130665

2

3

4

**Figure 1:** QUAST report showing the main metrics chosen to evaluate the better assembly among the ones made, better results are highlighted in blue and worst results in red.

5

6

7

8

9

10

11

12

13

14

**Table 1:** Main genomic data from LABIM22 genome

Genome size	3.996,396 pb
CDS	4042
GC%	46.4
tRNA	85
Read alignment	89.5%
Access nº GenBank	CP045993

15

16

17

18

19

20

21

22

23

24

25

26

27

28

**Strain identification and phylogenomic analysis.** The ANI and dDDH results indicate that when compared to the three main species from the *Bacillus subtilis* group LABIM22 showed more similarity with *Bacillus velezensis* Bac57 (99,26% for ANI and 94,30% for dDDH) (Table 2). Another strain that showed high similarity with LABIM22 was *B. velezensis* S141 (98,24% ANI, 84,90% dDDH), the isolated S141 as well as the reference *B. velezensis* FZB42 shows high capacity of growth-promoting activity mainly on soybean [2], [3]. The strain *B. velezensis* NKG-1 (98,19% ANI, 84,70% dDDH) showed on *in vitro* assays a wide range of antifungal activity, showing antagonism against 13 phytopathogenic fungi [5], [42], its high similarity with LABIM22 draws attention to a possible wider range of antifungal activity against fungi not tested yet. When compared to *B. amyloliquefaciens* DSM7T, the values of 93,67% for ANI and 56% for dDDH have showed that LABIM22 and DSM7T belongs to different species, this classification is important for these species due to the inconsistent delimitation for this group over the years [43]–[45].

**Table 2:** Genome comparison of LABIM22 with other *Bacillus* species.

Strains	Accession N° on GenBank	dDDH %	ANI (%)	GC (%)	Pb (Gb)	Source
<i>B. velezensis</i>						
<b>LABIM22</b>	<b>CP045993</b>	<b>100</b>	<b>100</b>	<b>46.40</b>	<b>3.99</b>	<b>Solo</b>
S141	AP018402.1	84.90	98.24	46.50	4.03	Soybean Rhizosphere [3]
NKG-1	CP024203.1	84.70	98.19	46.30	4.19	Volcanic soils at the Changbai mountains [42]
Bac57	CP033054.1	94.30	99.26	45.89	4.23	Red Sea lagoon
LABIM40	CP025079.1	88.10	98.22	46.50	3.97	Contaminant [46]
FZB42	CP000560.2	85.70	98.23	46.50	3.91	Soil [2]
QST713	CP025079	85.10	98,30	45.90	4.23	Commercial product Bayer® [47]
<i>B. amyloliquefaciens</i>						
Y2	NC_017912.1	92.90	99.15	45.90	4.23	Wheat Rhizosphere [48]
DSM7T	NC_014551.1	56.00	93.67	46.10	3.98	Non specified [4]
MT45	CP011252.1	56.00	94,13	46.10	3.89	Non specified
<i>B. subtilis</i>						
D125	CP020102.1	20.80	80.36	43.60	4.14	Soil [49]
168	NC_000964.3	20.90	80.36	43.50	4.21	Model microorganism [50]
<i>B. cereus</i>						
FRI-35	CP003747.1	39.30	68.10	35.45	5.38	Non specified

dDDH: digital DNA-DNA hybridization; ANI: Average Nucleotide Identity, GC: Content of guanine and cytosine on the genome, Pb: genome size in gigabases ( $10^6$ )

1  
2 The less similar species that appeared on the comparisons were *B. subtilis* D12-5 and *B. subtilis*  
3 168 (80,36%/80,36% for ANI and 20,80%/20,90% for dDDH respectively). *B. subtilis* 168 is  
4 a model microorganism studied mainly for its capacity of forming biofilm [51]–[53],  
5 suggesting that LABIM22 might have some potential in this important activity as a biocontrol  
6 agent.  
7 Due to the comparisons made with Gegenees and orthoANI LABIM22 is located within a  
8 cluster containing most of the *B. velezensis* species suggesting that it belongs to this group  
9 (Figure 2). Comparisons with strains like *B. amyloliquefaciens* Y2 suggest that some  
10 identifications are yet to be updated in the GenBank database. This analysis brings more robust  
11 conclusions concerning the phylogenetic position of the strain LABIM22 because it uses the  
12 whole genome as data for comparison instead of only the 16S rDNA, commonly used for other  
13 less complex bacterial groups [54]  
14

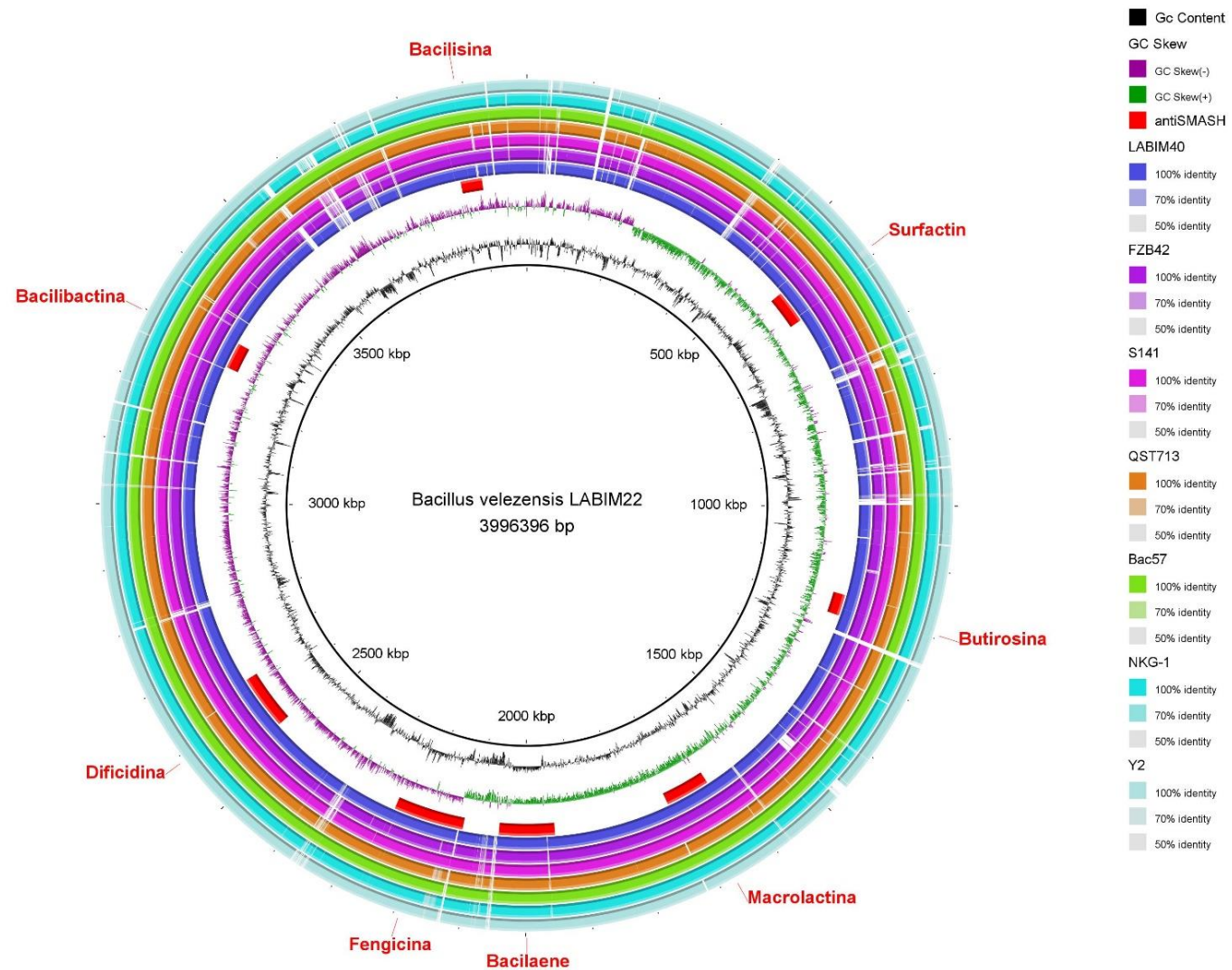


**Figure 2.** Phylogenomic tree representing the similarity between strains from the *Bacillus subtilis* group. The matrix was generated by Gegenees and exported to SplitsTree. Highlighted in green are the most similar strains to LABIM22 belonging to the *Bacillus velezensis* species. In yellow are the *B. amyloliquefaciens* species, less similar to LABIM22 and highlighted in red are the *B. subtilis* species with a *B. cereus* FRI35 as an extern group.

**antiSMASH analysis of secondary metabolism.** The webserver antiSMASH 5.1.0 found 13 BGCs on LABIM22 genome presented at Table 3. Clusters are also represented at the genome circular image (Figure 3). Among the 13, 8 have showed similarity with BGCs already found and described on MiBiG [55], those are linked respectively to the synthesis of Surfactin, Butirosin A/B, Macrolactin H, Bacillaen, Fengycin, Difficidin, Bacillibactin and Bacilysin. Surfactins are lipopeptides with great surfactant and antimicrobial properties as described on the literature [56], this group of molecules are also linked to the root colonization capacity, cell motility is directly associated with the presence of Surfactin on the extracellular matrix [14], [53] due to its capacity of decreasing the superficial tension of this matrix, also Surfactin is related to the induction of systemic resistance by the plant [57]. Macrolactins are a group of macrolides found at first in deep sea bacteria and have showed antimicrobial activity [58], [59] and growth inhibition of melanomas [60]. The 7<sup>th</sup> cluster found on the genome contains all the genes related to the biosynthesis of Fengycin, a lipopeptide with antifungal activity that alters the membrane permeability of the target cell causing its death [61], [62], as well as Surfactin it has the ability of inducing systemic resistance by the plant against potential phytopathogens [57]. This cluster is important because it might be directly related with the antifungal activity shown by LABIM22 on the *in vitro* assays against phytopathogenic fungi, possibly being the target of future *in silico* studies aiming its regulatory pathways to the synthesis of Fengycin. Studies made with the reference strain *Bacillus velezensis* FZB42 associated the antagonistic activity against *Erwinia amylovora*, a bacterial phytopathogen responsible for diseases on pomaceous fruits from the Rosaceae family [63], with the production of Difficidin and Bacilysin [64], 2 clusters found on LABIM22 genome with 100% similarity to the database MiBiG are related to the synthesis of these molecules, showing a potential against bacterial diseases not explored yet. The 13<sup>th</sup> cluster showed 100% similarity with a cluster of Bacillibactin production, an important siderophore [65], [66]. *Bacillus* species that are able to produce siderophores already have studies linking them with antifungal activity [67], this cluster also draws attention to a potential of ion solubilization present on the strain. 5 of the 13 clusters didn't showed any similarity clusters found in the database. Therefore, its products are yet to be identified and described, deeper studies are necessary to identify the product issue from it and associate it with the database, reinforcing the importance of the unexplored metabolic potential of the strain. All this data supports the findings of HIGASHI, 2020 [68]

1 that shows LABIM22 *in vitro* antifungal activity against *Sclerotinia sclerotiorum*,  
2 *Macrophomina phaseolina*, *Rhizoctonia solani*, *Botrytis cinerea*, *Fusarium solani* and  
3 *Fusarium oxysporum*. All of these phytopathogenic fungi are of agricultural importance  
4 causing damage to the production of various agricultural crop productions. The wide range of  
5 activity against phytopathogenic fungi is consistent with what is found on *Bacillus velezensis*  
6 NKG-1, that shows activity against 13 soil phytopathogens [42].  
7

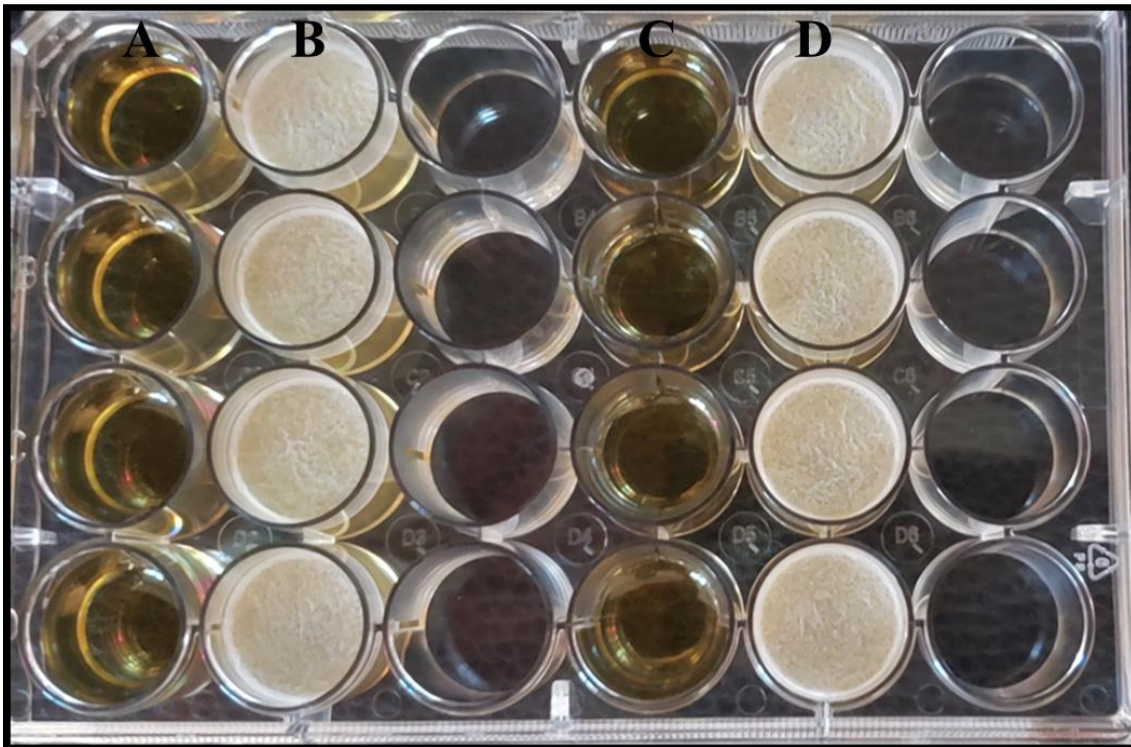
8 **Genes related to the biofilm formation/regulation and *in vitro* assays.** A series of genes  
9 related directly or indirectly to the biofilm production and root colonization were found on  
10 LABIM22 genome (Table 4). Among these a few are highlighted like *abrB*, a gene repressor  
11 that induces the endospore formation [32], [36], The gene *abrB* interacts with *spoOA*, another  
12 one found within LABIM22 genome and that its product have showed affinity for the  
13 promoting region of *abrB* [37], showing then to be a possible regulator of its expression and  
14 target of possible regulation strategies. *degQ* is a gene that have its expression associated with  
15 the antifungal activity of the strains studied. Mutants of this gene show low production of  
16 lipopeptides, main group of molecules linked to the antifungal activity on *Bacillus* spp. [69].  
17 Mutants of the gene *degU*, gene also found on LABIM22 genome, presented low biofilm  
18 formation capacity when compared to the wild strain [35], suggesting then to have a role on  
19 biofilm formation or regulation. Another repressor found on the genome was *sinR*, this one  
20 shows affinity for the promoting region of the operon *epsA-O* [70], this operon is related to the  
21 synthesis of exopolysaccharides, important for biofilm settling on the substrate [52]. All genes  
22 from this operon are present on LABIM22 genome. Another important operon found is *swrA-*  
23 *C*, this one is responsible for the swarm that helps with the bacterial motility, a relevant  
24 characteristic for the capacity of root colonization. Surfactins support the swarming decreasing  
25 the superficial tension on extracellular matrix from the biofilm allowing a larger mobility to  
26 the cells composing the biofilm [71]. Figure 7 represents the interaction of the genes  
27 highlighted and their role, which interacts directly or indirectly with the root colonization and  
28 biofilm formation, those relations are important for future studies aiming to genetically regulate  
29 LABIM22 root colonization capacity through genes associated with this pathway. All this  
30 machinery proves to be functional when LABIM22 is tested *in vitro* for its mobility capacity  
31 and pellicle formation (Figures 4 and 5). From the cell inoculum in the middle of the Petri dish,  
32 LABIM22 is able to grow over the whole surface of the TSA 0,8% medium, suggesting that its  
33 motility genes are active giving the strain capacity of moving through the dish, important  
34 ability for colonizing roots [53], [72], [73].



1

2 **Figure 3:** Circular representation of LABIM22 genome using the program BRIG. From the inside to the outside the legends are the following: (1)  
 3 GC Content, (2) GC Skew, (3) BGCs position on the genome indicated by antiSMASH, (4-10) respectively LABIM40, FZB42, S141, QST713,  
 4 Bac57, NKG-1 and Y2.

1



2

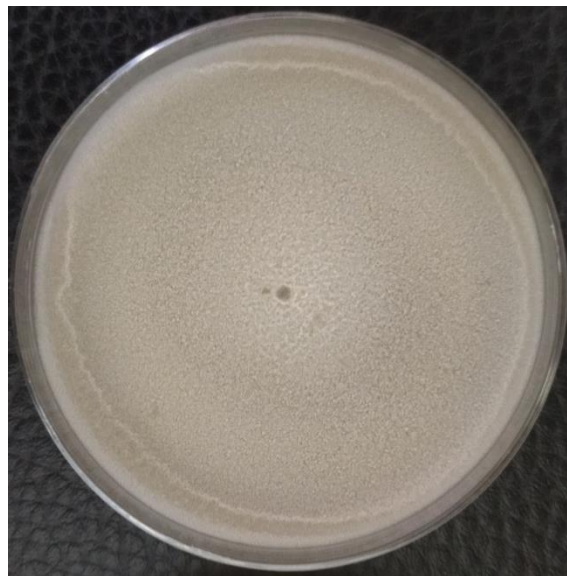
3

4

5

6

**Figure 4:** Pellicle formation on liquid-air interface on static liquid medium. Columns A and C: Medium Control; Columns B and D: inoculated with 10  $\mu$ L of solution containing LABIM22 cells adjusted to the tube 0.5 of MacFarland scale.



7

8

9

10

11

12

**Figure 5:** *In vitro* tests made with LABIM22, its motility capacity was tested on a TSA 0,8%. A cell inoculum was put in the middle of the dish and its growth was evaluated after 24h. LABIM22 was able to colonize the whole surface of the medium because of its ability to move thanks to flagella that its genome showed that has the capacity to synthesize.

1 **Table 3.** Biosynthetic Gene Clusters (BGCs) found within LABIM22 genome using the webserver antiSMASH 5.1.0.

CLUSTER	TYPE	FROM (PB)	TO (PB)	MOST SIMILAR KNOW CLUSTER		SIMILARITY
<b>CLUSTER 1</b>	NRPS	556,245	620,583	surfactin	NRP:Lipopeptide	82%
<b>CLUSTER 2</b>	Phosphonate	855,967	896,851			
<b>CLUSTER 3</b>	PKS-like	1,173,665	1,214,909	butirosin A / butirosin B	Saccharide	7%
<b>CLUSTER 4</b>	terpene	1,307,518	1,324,794			
<b>CLUSTER 5</b>	transAT-PKS	1,631,362	1,717,721	macrolactin H	Polyketide	100%
<b>CLUSTER 6</b>	transAT-PKS, T3PKS, transAT- PKS-like,NRPS	1,943,834	2,052,707	bacillaene	Polyketide + NRP	100%
<b>CLUSTER 7</b>	NRPS,transAT- PKS,betalactone	2,123,165	2,259,899	fengycin	NRP	100%
<b>CLUSTER 8</b>	terpene	2,283,957	2,305,840			
<b>CLUSTER 9</b>	T3PKS	2,381,461	2,422,561			
<b>CLUSTER 10</b>	transAT-PKS- like,transAT-PKS	2,537,196	2,643,371	difficidin	Polyketide + NRP	100%
<b>CLUSTER 11</b>	NRPS, bacteriocin	3,272,003	3,322,513	bacillibactin	NRP	100%
<b>CLUSTER 12</b>	NRPS	3,599,816	3,665,158			
<b>CLUSTER 13</b>	other	3,867,578	3,908,996	bacilysin	Other	100%

NRPS: NON-RIBOSOMAL PEPTIDE SYNTHETASE. NRP: NON-RIBOSOMAL PEPTIDE. PKS: POLYKETIDE SYNTHETASE. AT: ACETYL TRANSFERASE. T3PKS: TYPE 3 PKS.

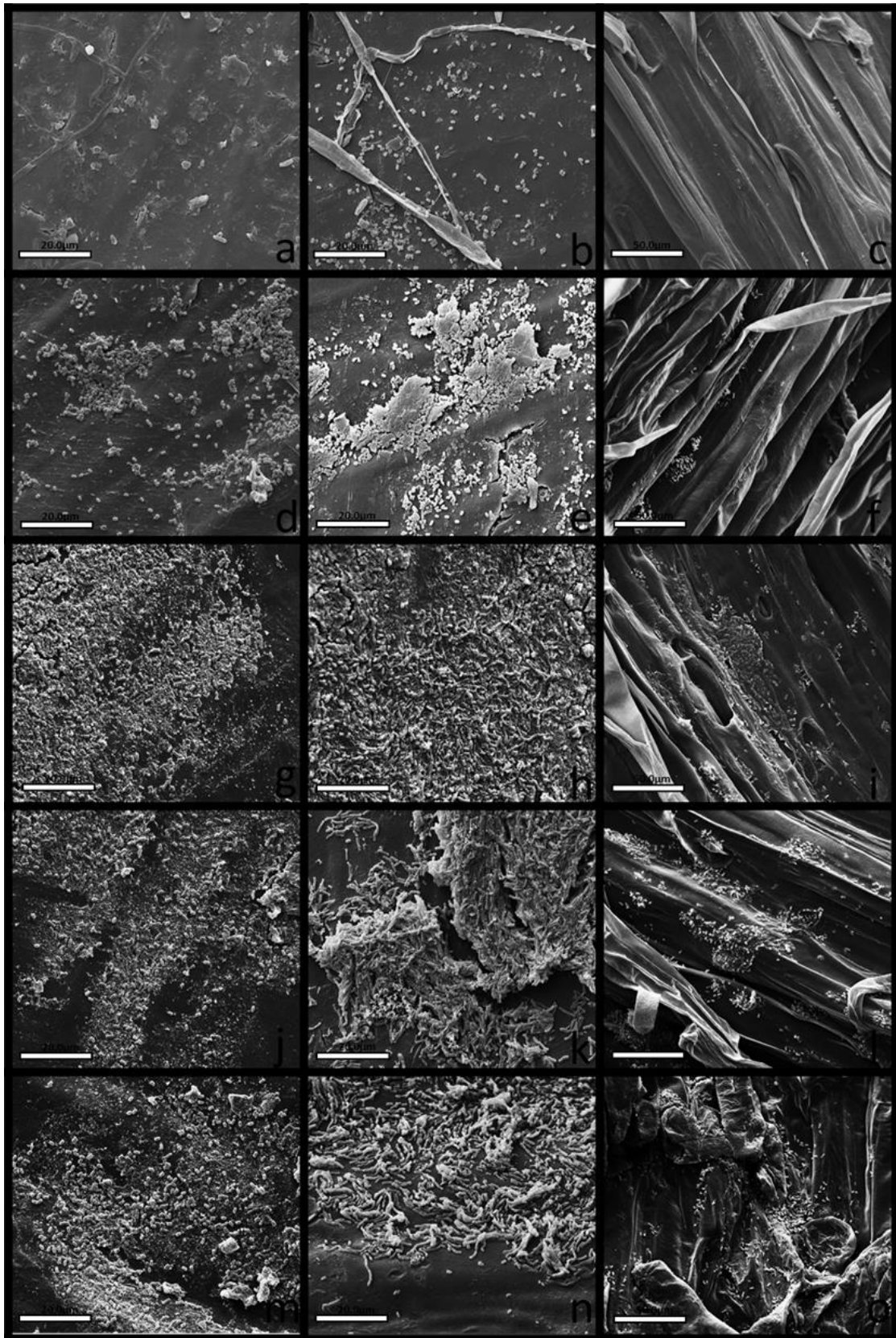
1 **Table 4.** Genes related to biofilm formation/regulation on LABIM22.

<b>GENES</b>	<b>POSITION ON LABIM22 (PB)</b>	<b>% OF IDENTITY WITH REFERENCE</b>	<b>DESCRIBED FUNTION ON SUBTWIKI [74]</b>
<i>ABRB</i>	252940 a 252651	91 %	Repressor of genes that induces sporulation
<i>DEGQ</i>	3241639 a 3241504	87%	Stimulates the production of degradative enzymes and extracelular poli-gama-glutamate
<i>DEGU</i>	3647286 a 3646597	86%	Capsule biosynthesis (along with SwrA), DegU non phosphorylated is necessary for motility on swarming
<i>EPSA</i>	3512467 a 3511760	99%	Synthesis of extracellular polysaccharides
<i>EPSB</i>	3511754 a 3511074	99%	
<i>EPSC</i>	3510827 a 3509034	99%	
<i>EPSD</i>	3509018 a 3507879	98%	
<i>EPSE</i>	3507882 a 3507043	98%	Motility and glicosiltransferase inhibitor needed for EPS biosynthesis
<i>EPSF</i>	3507047 a 3505911	99%	
<i>EPSG</i>	3505907 a 3504804	99%	
<i>EPSH</i>	3504785 a 3503748	99%	
<i>EPSI</i>	3503743 a 3502667	99%	Synthesis of poli-N-acetilglucosamine
<i>EPSJ</i>	3502670 a 3501636	99%	
<i>EPSK</i>	3501639 a 3500122	99%	
<i>EPSL</i>	3500125 a 3499517	99%	
<i>EPSM</i>	3499520 a 3498873	99%	
<i>EPSN</i>	3498868 a 3497696	99%	
<i>EPSO</i>	3497717 a 3496752	99%	
<i>GALEI</i>	1392187 a 1391195	99%	Galactose utilization
<i>KINA</i>	1573488 a 1575308	78%	
<i>MOTA</i>	1536204 a 1535389	79%	
<i>MOTB</i>	1535417 a 1534675	76%	
<i>REMA</i>	1789679 a 1789948	97%	Transcriptional regulator of genes from the extracellular matrix, acts

			alongside with SinR, AbrB and DegU
<b>SIGH</b>	324256 a 324893	88%	
<b>SIGW</b>	415480 a 416043	83%	
<b>SINI</b>	2676152 a 2676325	80%	Biofilm formation control
<b>SINR</b>	2676359 a 2676694	97%	Transcriptional regulator of post-exponential-phase responses genes
<b>SIPW</b>	2678176 a 2677592	99%	Bifunctional signal peptidase I that controls surface-adhered biofilm formation and processes TasA and TapA
<b>SPOOA</b>	2642989 a 2642189	99%	Regulator of the sporulation initiation
<b>SRFAA</b>	571928 a 573918	79%	Surfactin synthesis
<b>SRFAB</b>	564380 a 575138	75%	
<b>SRFAC</b>	575178 a 578974	87%	
<b>SWRA</b>	3622716 a 3622380	84%	Important activator of the flagellar biosynthesis, controls the number of flagellar basal bodies, controls degU activity
<b>SWRB</b>	1865316 a 1865494	82%	Controls SigD activity. Activates type 3 flagelar secretion by the protein membrane FliP
<b>SWRC</b>	895855 a 898998	80%	
<b>TASA</b>	2677527 a 2676742	99%	Major component of the biofilm matrix, forms amiloids fibers
<b>YHXB</b>	1141855 a 1143597	99%	
<b>YQXM</b>	2678819 a 2678148	99%	

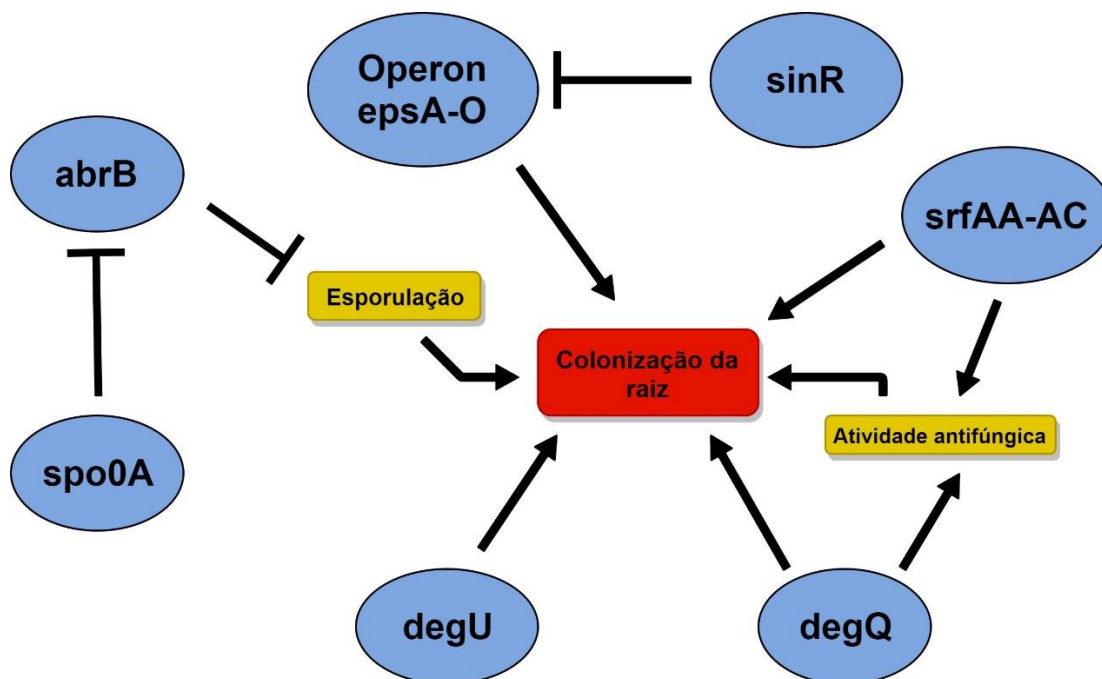
1  
2 **Root colonization on *in vivo* assays.** At Figure 6 it can be seen that LABIM22 is able to  
3 attach to the maize seed surface and multiplies itself, which is important for introduction  
4 of the bacteria into the rhizosphere [73] as well as being able to colonize the initial root  
5 formation on the seedlings. Each line represents each treatment, the first column shows  
6 the seed surface treated at day 0 at 3.000x, the second column shows the seed surface  
7 after 5 days of incubation at 3.000x and the third column displays the root surface at  
8 2.000x. With the three concentrations tested LABIM22 colonized the root at various parts  
9 of the tissue and formed cell aggregates that are able to produce biofilm and consequently  
10 produce the antifungal compounds [8] issue of their secondary metabolism. A primary  
11 biofilm formation can be spotted at item I and L, consistent with the data found on its  
12 genome and the *in vitro* assays proving its mobility and biofilm formation capacity.  
13 Results suggest that *Bacillus velezensis* LABIM22 can be used also as seed defensive as  
14 it its capable of colonizing the seed surface after treatment and is able to continue its life

1 cycle until initial root formation, which the strain is also capable of colonizing and form  
 2 biofilms [15], [75], [76].  
 3



4  
 5 **Figure 6:** LABIM22 root colonization. Set of 15 images showing LABIM22 *in vivo*  
 6 colonization capacity. First 2 columns are Seed images on the first day of treatment and

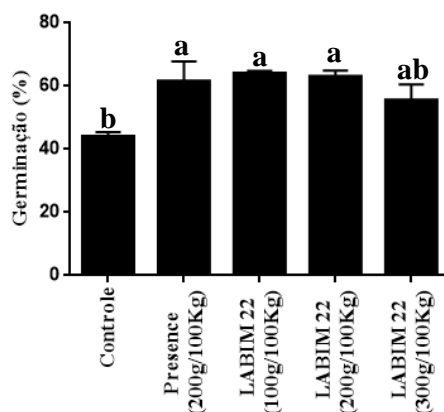
1 7<sup>th</sup> day respectively. Third column shows the initial root formation of the seedlings. Each  
 2 line from top to bottom correspond to T1 (negative control); T2 (positive control)  
 3 treatment with the commercial biological product Presence; T3 treatment with LABIM22  
 4 at a concentration of 100g per 100 kg of seeds; T4 treatment with LABIM22 at a  
 5 concentration of 200g per 100 kg of seeds; T5 treatment with LABIM22 at a  
 6 concentration of 300g per 100 kg of seeds.



9  
 10 **Figure 7:** Representation of the main genes found within the LABIM22 genome related  
 11 with the antifungal activity and sporulation inducing the biofilm formation/regulation  
 12 aiming the root colonization. Arrows indicate positive regulation, “T” bars indicate  
 13 repression of the gene expression.

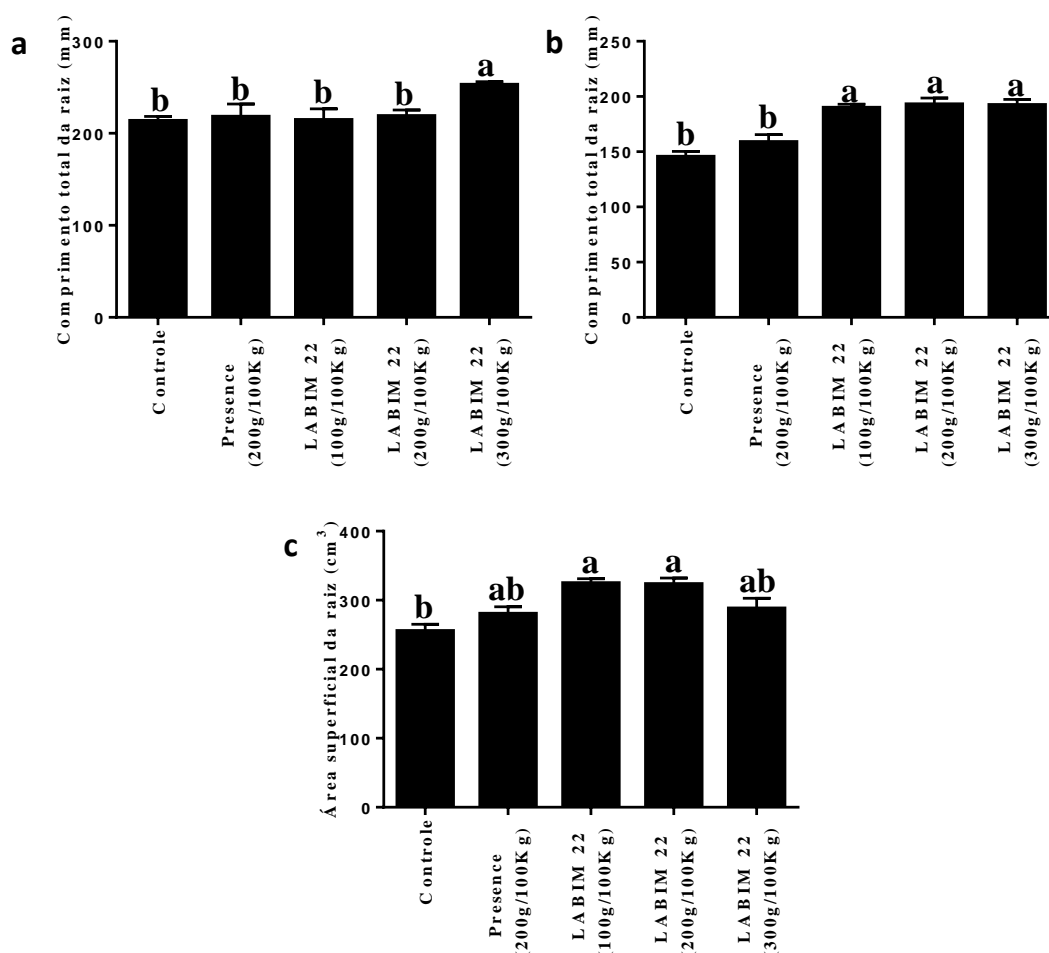
14  
 15 **Germination and Growth Promoting ability.** *In vivo* tests with maize and soybean  
 16 showed that there was no significant difference in germination rates within the treatments  
 17 with the maize. The amplitude of germination ranged from 96% and 99% for this cultivar  
 18 in all treatments. Soybean germination rates increased with cell treatments (from 55,5%  
 19 to 64%) when compared to the control (44%) (Figure 7), indicating that LABIM22 and  
 20 biological based products in general may have positive effects on germination rates for  
 21 this kind of crop. Further analysis is needed to better comprehend the positive biological  
 22 effect of LABIM22 on germination rates to stablish its potential for germination  
 23 promotion Concerning the root and aerial part dry mass there was no significant  
 24 difference treatments. However, for the radicular properties the biological treatments  
 25 induced a difference in the total root length on the maize and soybean plants and total  
 26 root surface area for the soybean plants (Figure 8). Total root length in soybean was  
 27 increased in all treatments with LABIM22 when compared to both controls (Figure 8b),  
 28 maize plants increased total root length only with the highest of the three cell  
 29 concentrations used with LABIM22 treatments (Figure 8a), indicating a larger product  
 30 usage for this cultivar to obtain larger roots is needed, for soybeans all concentrations  
 31 induced an increase in total root length, total surface area increased on all biological

1 treatments but didn't showed any difference between the LABIM22 treatments and the  
2 positive control (Figure 8c). Evidencing LABIM22 potential as a PGPR.



3

4 **Figure 7:** Inoculation effect of bacteria on the germination percentage (%) on soybean  
5 culture. The bars are averages of 4 repetitions; each repetition contained 50 seed per  
6 treatment. Bars with different letters above differs significantly between them at  $p < 0,05$ ,  
7 according with the Tukey test.



8

9 **Figure 8:** Effect of bacteria inoculation on total root length and root surface area of maize  
10 (a) and soybean (b and c) roots. The bars are averages of 4 repetitions; each repetition

1 contained 50 seed per treatment. Bars with different letters above differs significantly  
2 between them at  $p < 0.05$ , according with the Tukey test.

### 3 **Conclusions**

4 With the presented work is possible to affirm that *Bacillus velezensis* LABIM22 harbors  
5 a biotechnological potential due to its high genetic similarity with other strains from the  
6 *Bacillus* genus already used as biocontrol agents and established on the market. This  
7 potential is reinforced by the presence of genes found in the genome that are directly and  
8 indirectly associated with biofilm formation/regulation and by the exhibited capacity of  
9 colonizing roots *in vivo*. The study of the genetic machinery linked with the secondary  
10 metabolism also strengthens the biotechnological potential of the strain by finding BGCs  
11 that are related with antimicrobial activity already described on the literature, in addition  
12 to that, BGCs that were found that didn't had any similarity with the ones described on  
13 the database can be targets of *in silico* studies focused on the diversity of molecules that  
14 can be synthesized by LABIM22 during its secondary metabolism. Also, its growth-  
15 promoting capacity described in this work classifies LABIM22 as a PGPR, important  
16 characteristic of a bacteria that is aimed to be used as a crop defensive.

### 17 **Conflict of interest**

18 The authors declare that they have no conflict of interests.

### 20 **Acknowledgments**

21 We acknowledge the financial support from the Brazilian National Council of Scientific  
22 and Technological Development (CNPq) and Fundação Araucária de Apoio ao  
23 Desenvolvimento Científico e Tecnológico do Estado do Paraná (FAPPR). A.G. de  
24 Oliveira was supported by a research fellowship from CNPq No. 315593/2018-0. The  
25 authors also acknowledge Laboratório de Microscopia Eletrônica e Microanálises,  
26 Central Multiusuárias de Laboratórios de Pesquisa (CMLP), Universidade Estadual de  
27 Londrina, Paraná, Brazil.

### 30 **REFERÊNCIAS**

- 31
- 32 [1] N. Mizuhara, M. Kuroda, A. Ogita, T. Tanaka, Y. Usuki, and K. I. Fujita,  
33 “Antifungal thiopeptide cyclothiazomycin B1 exhibits growth inhibition  
34 accompanying morphological changes via binding to fungal cell wall chitin,”  
35 *Bioorganic Med. Chem.*, vol. 19, no. 18, pp. 5300–5310, 2011.
- 36 [2] X. H. Chen *et al.*, “Genome analysis of *Bacillus amyloliquefaciens* FZB42  
37 reveals its potential for biocontrol of plant pathogens,” *J. Biotechnol.*, vol. 140,  
38 no. 1, pp. 27–37, Mar. 2009.
- 39 [3] S. Sibponkrung *et al.*, “Genome sequence of *Bacillus velezensis* S141, a new  
40 strain of plant growthpromoting rhizobacterium isolated from soybean  
41 rhizosphere,” *Genome Announc.*, vol. 5, no. 48, pp. 7–8, 2017.
- 42 [4] C. Rückert, J. Blom, X. H. Chen, O. Reva, and R. Borriss, “Genome sequence of  
43 *B. amyloliquefaciens* type strain DSM7T reveals differences to plant-associated  
44 *B. amyloliquefaciens* FZB42,” *J. Biotechnol.*, vol. 155, no. 1, pp. 78–85, 2011.
- 45 [5] B. Liu, B. Ge, N. Azhar, W. Zhao, H. Cui, and K. Zhang, “Complete genome  
46 sequence of *Bacillus methylotrophicus* strain NKG-1, isolated from the Changbai  
47 Mountains, China,” *Genome Announc.*, vol. 6, no. 3, pp. 1–2, 2018.

- 1 [6] A. Beneduzi, A. Ambrosini, and L. M. P. Passaglia, "Plant growth-promoting  
2 rhizobacteria (PGPR): Their potential as antagonists and biocontrol agents,"  
3 *Genetics and Molecular Biology*, vol. 35, no. 4 SUPPL. pp. 1044–1051, 2012.
- 4 [7] G. V. Bloemberg and B. J. J. Lugtenberg, "Molecular basis of plant growth  
5 promotion and biocontrol by rhizobacteria," *Current Opinion in Plant Biology*,  
6 vol. 4, no. 4. Elsevier Ltd, pp. 343–350, 2001.
- 7 [8] N. Zhang *et al.*, "Whole transcriptomic analysis of the plant-beneficial  
8 rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm  
9 formation regulated by maize root exudates," *BMC Genomics*, vol. 16, no. 1, pp.  
10 1–20, 2015.
- 11 [9] T. Xu, T. Zhu, and S. Li, "B-1,3-1,4-Glucanase Gene From *Bacillus velezensis*  
12 Zj20 Exerts Antifungal Effect on Plant Pathogenic Fungi," *World J. Microbiol.*  
13 *Biotechnol.*, vol. 32, no. 2, pp. 1–9, 2016.
- 14 [10] S. H. Ji, N. C. Paul, J. X. Deng, Y. S. Kim, B. S. Yun, and S. H. Yu, "Biocontrol  
15 activity of *Bacillus amyloliquefaciens* CNU114001 against fungal plant  
16 diseases," *Mycobiology*, vol. 41, no. 4, pp. 234–242, 2013.
- 17 [11] S. M. Lim *et al.*, "Diffusible and volatile antifungal compounds produced by an  
18 antagonistic *Bacillus velezensis* G341 against various phytopathogenic fungi,"  
19 *Plant Pathol. J.*, vol. 33, no. 5, pp. 488–498, 2017.
- 20 [12] R. Bleich, J. D. Watrous, P. C. Dorrestein, A. A. Bowers, and E. A. Shank,  
21 "Thiopeptide antibiotics stimulate biofilm formation in *Bacillus subtilis*," *Proc.*  
22 *Natl. Acad. Sci. U. S. A.*, vol. 112, no. 10, pp. 3086–3091, 2015.
- 23 [13] J.-S. Yoo, C.-J. Zheng, S. Lee, J.-H. Kwak, and W.-G. Kim, "Macrolactin N, a  
24 new peptide deformylase inhibitor produced by *Bacillus subtilis*," *Bioorg. Med.*  
25 *Chem. Lett.*, vol. 16, no. 18, pp. 4889–4892, Sep. 2006.
- 26 [14] H. P. Bais, R. Fall, and J. M. Vivanco, "Biocontrol of *Bacillus subtilis* against  
27 infection of *Arabidopsis* roots by *Pseudomonas syringae* is facilitated by  
28 biofilm formation and surfactin production," *Plant Physiol.*, vol. 134, no. 1, pp.  
29 307–319, Jan. 2004.
- 30 [15] B. Fan, X. H. Chen, A. Budiharjo, W. Bleiss, J. Vater, and R. Borriss, "Efficient  
31 colonization of plant roots by the plant growth promoting bacterium *Bacillus*  
32 *amyloliquefaciens* FZB42, engineered to express green fluorescent protein," *J.*  
33 *Biotechnol.*, vol. 151, no. 4, pp. 303–311, Feb. 2011.
- 34 [16] K. J. Grubbs *et al.*, "Large-Scale Bioinformatics Analysis of *Bacillus* Genomes  
35 Uncovers Conserved Roles of Natural Products in Bacterial Physiology.,"  
36 *mSystems*, vol. 2, no. 6, 2017.
- 37 [17] P. Cimermancic *et al.*, "Insights into secondary metabolism from a global  
38 analysis of prokaryotic biosynthetic gene clusters," *Cell*, vol. 158, no. 2, pp. 412–  
39 421, Jul. 2014.
- 40 [18] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: a flexible trimmer for  
41 Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, Aug.  
42 2014.
- 43 [19] R. M. Leggett, R. H. Ramirez-Gonzalez, B. J. Clavijo, D. Waite, and R. P.  
44 Davey, "Sequencing quality assessment tools to enable data-driven informatics  
45 for high throughput genomics," *Front. Genet.*, vol. 4, 2013.
- 46 [20] D. R. Zerbino and E. Birney, "Velvet: algorithms for de novo short read  
47 assembly using de Bruijn graphs.," *Genome Res.*, vol. 18, no. 5, pp. 821–9, May  
48 2008.
- 49 [21] Y. Peng, H. C. M. Leung, S. M. Yiu, and F. Y. L. Chin, "IDBA - A practical  
50 iterative De Bruijn graph De Novo assembler," in *Lecture Notes in Computer*

- 1 *Science (including subseries Lecture Notes in Artificial Intelligence and Lecture*  
2 *Notes in Bioinformatics)*, 2010, vol. 6044 LNBI, pp. 426–440.
- 3 [22] A. Gurevich, V. Saveliev, N. Vyahhi, and G. Tesler, “QUAST: Quality  
4 assessment tool for genome assemblies,” *Bioinformatics*, vol. 29, no. 8, pp.  
5 1072–1075, Apr. 2013.
- 6 [23] M. Galardini, E. G. Biondi, M. Bazzicalupo, and A. Mengoni, “CONTIGuator: a  
7 bacterial genomes finishing tool for structural insights on draft genomes.,”  
8 *Source Code Biol. Med.*, vol. 6, p. 11, Jun. 2011.
- 9 [24] I. J. Tsai, T. D. Otto, and M. Berriman, “Improving draft assemblies by iterative  
10 mapping and assembly of short reads to eliminate gaps,” *Genome Biol.*, vol. 11,  
11 no. 4, p. R41, 2010.
- 12 [25] N.-F. F. Alikhan, N. K. Petty, N. L. Ben Zakour, and S. A. Beatson, “BLAST  
13 Ring Image Generator (BRIG): Simple prokaryote genome comparisons,” *BMC*  
14 *Genomics*, vol. 12, p. 402, Aug. 2011.
- 15 [26] T. Weber *et al.*, “antiSMASH 3.0—a comprehensive resource for the genome  
16 mining of biosynthetic gene clusters,” *Nucleic Acids Research*, vol. 43, no. Web  
17 Server issue. pp. W237-43, Jul-2015.
- 18 [27] M. Richter and R. Rossello-Mora, “Shifting the genomic gold standard for the  
19 prokaryotic species definition,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 45, pp.  
20 19126–19131, Nov. 2009.
- 21 [28] J. P. Meier-Kolthoff, A. F. Auch, H.-P. P. Klenk, and M. Göker, “Genome  
22 sequence-based species delimitation with confidence intervals and improved  
23 distance functions,” *BMC Bioinformatics*, vol. 14, no. 1, p. 60, 2013.
- 24 [29] J. Agren *et al.*, “Gegenees: Fragmented alignment of multiple genomes for  
25 determining phylogenomic distances and genetic signatures unique for specified  
26 target groups,” *PLoS One*, vol. 7, no. 6, p. e39107, Jun. 2012.
- 27 [30] D. H. Huson, “SplitsTree: Analyzing and visualizing evolutionary data,”  
28 *Bioinformatics*, vol. 14, no. 1, pp. 68–73, Feb. 1998.
- 29 [31] R. K. Aziz *et al.*, “The RAST Server: Rapid annotations using subsystems  
30 technology,” *BMC Genomics*, vol. 9, p. 75, Feb. 2008.
- 31 [32] J. B. Robertson, M. Gocht, M. A. Marahiel, and P. Zuber, “AbrB, a regulator of  
32 gene expression in Bacillus, interacts with the transcription initiation regions of a  
33 sporulation gene and an antibiotic biosynthesis gene,” *Proc. Natl. Acad. Sci. U. S.*  
34 *A.*, vol. 86, no. 21, pp. 8457–8461, Nov. 1989.
- 35 [33] F. Chu, D. B. Kearns, A. McLoon, Y. Chai, R. Kolter, and R. Losick, “A novel  
36 regulatory protein governing biofilm formation in Bacillus subtilis.,” *Mol.*  
37 *Microbiol.*, vol. 68, no. 5, pp. 1117–27, Jun. 2008.
- 38 [34] T. Msadek, F. Kunst, A. Klier, and G. Rapoport, “DegS-DegU and ComP-ComA  
39 modulator-effector pairs control expression of the Bacillus subtilis pleiotropic  
40 regulatory gene degQ.,” *J. Bacteriol.*, vol. 173, no. 7, pp. 2366–77, Apr. 1991.
- 41 [35] K. Kobayashi, “Gradual activation of the response regulator DegU controls serial  
42 expression of genes for flagellum formation and biofilm formation in Bacillus  
43 subtilis,” *Mol. Microbiol.*, vol. 66, no. 2, pp. 395–409, Oct. 2007.
- 44 [36] M. A. Strauch, B. G. Bobay, J. Cavanagh, F. Yao, A. Wilson, and Y. Le Breton,  
45 “Abh and AbrB control of Bacillus subtilis antimicrobial gene expression,” *J.*  
46 *Bacteriol.*, vol. 189, no. 21, pp. 7720–7732, Nov. 2007.
- 47 [37] M. Strauch, V. Webb, G. Spiegelman, and J. A. Hoch, “The SpoOA protein of  
48 Bacillus subtilis is a repressor of the abrB gene.,” *Proc. Natl. Acad. Sci. U. S. A.*,  
49 vol. 87, no. 5, p. 1801, 1990.
- 50 [38] S. Mordini *et al.*, “The Role of SwrA, DegU and PD3 in fla/che Expression in B.

- 1 subtilis,” *PLoS One*, vol. 8, no. 12, p. e85065, Dec. 2013.
- 2 [39] T. Galkovskiy *et al.*, “GiA Roots: Software for the high throughput analysis of  
3 plant root system architecture,” *BMC Plant Biol.*, vol. 12, no. 1, p. 116, Jul. 2012.
- 4 [40] F. De Mendiburu and R. Simon, “Agricolae - Ten years of an open source  
5 statistical tool for experiments in breeding, agriculture and biology,” Sep. 2015.
- 6 [41] F. S. Kremer, A. J. A. McBride, and L. da S. Pinto, “Approaches for in silico  
7 finishing of microbial genome sequences,” *Genetics and Molecular Biology*, vol.  
8 40, no. 3. Brazilian Journal of Genetics, pp. 553–576, 2017.
- 9 [42] B. Ge, B. Liu, T. T. Nwet, W. Zhao, L. Shi, and K. Zhang, “Bacillus  
10 methylotrophicus strain NKG-1, Isolated from changbai mountain, China, Has  
11 potential applications as a biofertilizer or biocontrol agent,” *PLoS One*, vol. 11,  
12 no. 11, pp. 1–13, 2016.
- 13 [43] S.-J. Kim, S.-W. Kwon, C. A. Dunlap, and A. P. Rooney, “Phylogenomic  
14 analysis shows that *Bacillus amyloliquefaciens* subsp. *plantarum* is a later  
15 heterotypic synonym of *Bacillus methylotrophicus*,” *Int. J. Syst. Evol. Microbiol.*,  
16 vol. 65, no. 7, pp. 2104–2109, Jul. 2015.
- 17 [44] C. A. Dunlap, S.-J. Kim, S.-W. Kwon, and A. P. Rooney, “*Bacillus velezensis* is  
18 not a later heterotypic synonym of *Bacillus amyloliquefaciens*; *Bacillus*  
19 *methylotrophicus*, *Bacillus amyloliquefaciens* subsp. *plantarum* and ‘*Bacillus*  
20 *oryzicola*’ are later heterotypic synonyms of *Bacillus velezensis* based on  
21 phylogenomics,” *Int. J. Syst. Evol. Microbiol.*, vol. 66, no. 3, pp. 1212–1217,  
22 Mar. 2016.
- 23 [45] B. Fan, J. Blom, H. P. Klenk, and R. Borriss, “*Bacillus amyloliquefaciens*,  
24 *Bacillus velezensis*, and *Bacillus siamensis* Form an ‘Operational Group B.  
25 *amyloliquefaciens*’ within the *B. subtilis* species complex,” *Front. Microbiol.*,  
26 vol. 8, no. JAN, 2017.
- 27 [46] J. P. Baptista *et al.*, “Complete genome sequence of *Bacillus velezensis*  
28 LABIM40, an effective antagonist of fungal plant pathogens,” *Genome*  
29 *Announc.*, vol. 6, no. 25, pp. 1–2, Jun. 2018.
- 30 [47] C. Pandin *et al.*, “Complete genome sequence of *Bacillus velezensis* QST713: A  
31 biocontrol agent that protects *Agaricus bisporus* crops against the green mould  
32 disease,” *J. Biotechnol.*, vol. 278, no. 2010, pp. 10–19, Jul. 2018.
- 33 [48] K. Hao *et al.*, “The Genome of plant growth-promoting *Bacillus*  
34 *amyloliquefaciens* subsp. *Plantarum* Strain YAU B9601-Y2 contains a gene  
35 cluster for mersacidin synthesis,” *J. Bacteriol.*, vol. 194, no. 12, pp. 3264–3265,  
36 2012.
- 37 [49] A. Niazi, S. Manzoor, S. Asari, S. Bejai, J. Meijer, and E. Bongcam-Rudloff,  
38 “Genome Analysis of *Bacillus amyloliquefaciens* Subsp. *plantarum* UCMB5113:  
39 A Rhizobacterium That Improves Plant Growth and Stress Management,” *PLoS*  
40 *One*, vol. 9, no. 8, p. 15, Aug. 2014.
- 41 [50] S. Borchert *et al.*, “The complete genome sequence of the gram-positive  
42 bacterium *Bacillus subtilis*,” *Nature*, vol. 390, no. 6657, p. 249–256, 1997.
- 43 [51] R. Gallegos-Monterrosa, E. Mhatre, and Á. T. Kovács, “Specific *Bacillus subtilis*  
44 168 variants form biofilms on nutrient-rich medium,” vol. 162, no. 11, pp. 1922–  
45 1932, Nov. 2016.
- 46 [52] H. Vlamakis, Y. Chai, P. Beauregard, R. Losick, and R. Kolter, “Sticking  
47 together: building a biofilm the *Bacillus subtilis* way,” *Nat. Rev. Microbiol.*, vol.  
48 11, no. 3, pp. 157–168, Mar. 2013.
- 49 [53] R. F. Kinsinger, M. C. Shirk, and R. Fall, “Rapid surface motility in *Bacillus*  
50 *subtilis* is dependent on extracellular surfactin and potassium ion,” *J. Bacteriol.*,

- 1 vol. 185, no. 18, pp. 5627–5631, Sep. 2003.
- 2 [54] P. C. Y. Woo, S. K. P. Lau, J. L. L. Teng, H. Tse, and K.-Y. Yuen, “Then and  
3 now: use of 16S rDNA gene sequencing for bacterial identification and discovery  
4 of novel bacteria in clinical microbiology laboratories.,” *Clin. Microbiol. Infect.*,  
5 vol. 14, no. 10, pp. 908–34, Oct. 2008.
- 6 [55] M. H. Medema *et al.*, “Minimum Information about a Biosynthetic Gene  
7 cluster,” *Nat. Chem. Biol.*, vol. 11, no. 9, pp. 625–631, Sep. 2015.
- 8 [56] B. Kluge, J. Vater, J. Salnikow, and K. Eckart, “Studies on the biosynthesis of  
9 surfactin, a lipopeptide antibiotic from *Bacillus subtilis* ATCC 21332,” *FEBS*  
10 *Lett.*, vol. 231, no. 1, pp. 107–110, Apr. 1988.
- 11 [57] M. Ongena *et al.*, “Surfactin and fengycin lipopeptides of *Bacillus subtilis* as  
12 elicitors of induced systemic resistance in plants,” *Environ. Microbiol.*, vol. 9,  
13 no. 4, pp. 1084–1090, Apr. 2007.
- 14 [58] M.-J. Sohn, C.-J. Zheng, and W.-G. Kim, “Macrolactin S, a New Antibacterial  
15 Agent with Fab G-inhibitory Activity from *Bacillus* sp. AT28,” *J. Antibiot.*  
16 (*Tokyo*), vol. 61, no. 11, pp. 687–691, Nov. 2008.
- 17 [59] X. L. Lu *et al.*, “Macrolactin S, a novel macrolactin antibiotic from marine  
18 *Bacillus* sp.,” *Nat. Prod. Res.*, vol. 22, no. 4, pp. 342–347, Mar. 2008.
- 19 [60] K. Gustafson, M. Roman, and W. Fenical, “The Macrolactins, a Novel Class of  
20 Antiviral and Cytotoxic Macrolides from a Deep-Sea Marine Bacterium,” *J. Am.*  
21 *Chem. Soc.*, vol. 111, no. 19, pp. 7519–7524, Sep. 1989.
- 22 [61] M. Deleu, M. Paquot, and T. Nylander, “Effect of fengycin, a lipopeptide  
23 produced by *Bacillus subtilis*, on model biomembranes,” *Biophys. J.*, vol. 94, no.  
24 7, pp. 2667–2679, 2008.
- 25 [62] N. Vanittanakom, W. Loeffler, U. Koch, and G. Jung, “Fengycin-A Novel  
26 Antifungal Lipopeptide Antibiotic Produced by *Bacillus Subtilis* F-29-3,” *J.*  
27 *Antibiot. (Tokyo)*, vol. 39, no. 7, pp. 888–901, 1986.
- 28 [63] I.-S. Myung *et al.*, “Fire Blight of Apple, Caused by *Erwinia amylovora*, a New  
29 Disease in Korea,” *Plant Dis.*, vol. 100, no. 8, pp. 1774–1774, Aug. 2016.
- 30 [64] X. H. Chen *et al.*, “Difficidin and bacilysin produced by plant-associated *Bacillus*  
31 *amyloliquefaciens* are efficient in controlling fire blight disease,” *J. Biotechnol.*,  
32 vol. 140, no. 1–2, pp. 38–44, Mar. 2009.
- 33 [65] M. Miethke, O. Klotz, U. Linne, J. J. May, C. L. Beckering, and M. A. Marahiel,  
34 “Ferri-bacillibactin uptake and hydrolysis in *Bacillus subtilis*,” *Mol. Microbiol.*,  
35 vol. 61, no. 6, pp. 1413–1427, Sep. 2006.
- 36 [66] E. A. Dertz, J. Xu, A. Stintzi, and K. N. Raymond, “Bacillibactin-mediated iron  
37 transport in *Bacillus subtilis*,” *J. Am. Chem. Soc.*, vol. 128, no. 1, pp. 22–23, Jan.  
38 2006.
- 39 [67] X. Yu, C. Ai, L. Xin, and G. Zhou, “The siderophore-producing bacterium,  
40 *Bacillus subtilis* CAS15, has a biocontrol effect on *Fusarium* wilt and promotes  
41 the growth of pepper,” *Eur. J. Soil Biol.*, vol. 47, no. 2, pp. 138–145, Mar. 2011.
- 42 [68] A. Higashi, “Antagonismo da cepa LABIM22 de *Bacillus* sp. frente a fungos  
43 fitopatógenos e biocontrole do mofo branco em sementes de soja,” Universidade  
44 Estadual de Londrina, 2020.
- 45 [69] M. Ongena and P. Jacques, *Bacillus lipopeptides: versatile weapons for plant*  
46 *disease biocontrol*, vol. 16, no. 3. 2008, pp. 115–125.
- 47 [70] D. B. Kearns, F. Chu, S. S. Branda, R. Kolter, and R. Losick, “A master regulator  
48 for biofilm formation by *Bacillus subtilis*,” *Mol. Microbiol.*, vol. 55, no. 3, pp.  
49 739–749, Dec. 2004.
- 50 [71] D. B. Kearns and R. Losick, “Swarming motility in undomesticated *Bacillus*

- 1 subtilis,” *Mol. Microbiol.*, vol. 49, no. 3, pp. 581–590, Jan. 2004.
- 2 [72] S. De Weert *et al.*, “Flagella-driven chemotaxis towards exudate components is  
3 an important trait for tomato root colonization by *Pseudomonas fluorescens*,”  
4 *Mol. Plant-Microbe Interact.*, vol. 15, no. 11, pp. 1173–1180, Nov. 2002.
- 5 [73] T. Danhorn and C. Fuqua, “Biofilm Formation by Plant-Associated Bacteria,”  
6 *Annu. Rev. Microbiol.*, vol. 61, no. 1, pp. 401–422, Oct. 2007.
- 7 [74] B. Zhu and J. Stülke, “SubtiWiki in 2018: from genes and proteins to functional  
8 network annotation of the model organism *Bacillus subtilis*,” *Nucleic Acids Res.*,  
9 vol. 46, no. D1, pp. D743–D748, Jan. 2018.
- 10 [75] J. W. Kloepper, R. Lifshitz, and R. M. Zablotowicz, “Free-living bacterial  
11 inocula for enhancing crop productivity,” *Trends in Biotechnology*, vol. 7, no. 2.  
12 Elsevier Current Trends, pp. 39–44, 01-Feb-1989.
- 13 [76] M. N. Schroth and J. G. Hancock, “Disease-suppressive soil and root-colonizing  
14 bacteria,” *Science (80-. )*, vol. 216, no. 4553, pp. 1376–1381, Jun. 1982.
- 15

## 1 **CONCLUSÃO**

2 Os estudos *in silico* realizados com a cepa *Bacillus velezensis* LABIM22  
3 mostraram que este isolado possui um elevado potencial biotecnológico a ser  
4 explorado, reforçado pelas análises de similaridade realizadas com cepas já  
5 estabelecidas como eficientes agentes de biocontrole com uma diversa  
6 maquinaria molecular oriunda do metabolismo secundário, esta diversa  
7 maquinaria também está presente no genoma de LABIM22 codificando  
8 importantes grupos moleculares com diversas atividades, algumas já exploradas  
9 em experimentos *in vitro* com atividade antifúngica, outras ainda não exploradas  
10 mas com potencial uso biotecnológico descrito em literatura. Por se tratar de um  
11 possível agente de biocontrole, os genes relacionados a formação de biofilme  
12 encontrados no genoma reforçam ainda mais o seu potencial, uma vez que a  
13 colonização rizosférica é de extrema importância para o estabelecimento do  
14 microrganismo e conseqüentemente a produção dos metabólitos que auxiliam a  
15 planta, seja por indução de defesa quanto por promoção de crescimento ou  
16 atividade antimicrobiana.

17 Mais estudos são necessários para elucidar de maneira mais clara os  
18 mecanismos envolvidos no biocontrole e compreender melhor os processos de  
19 biossíntese das moléculas de interesse, podendo auxiliar nos processos  
20 fermentativos para uma melhor extração e purificação de compostos e suas  
21 respectivas caracterizações.

1   **REFERÊNCIAS**

- 2
- 3   [1]   G. V. Bloemberg and B. J. J. Lugtenberg, "Molecular basis of plant growth  
4   promotion and biocontrol by rhizobacteria," *Current Opinion in Plant  
5   Biology*, vol. 4, no. 4. Elsevier Ltd, pp. 343–350, 2001.
- 6   [2]   A. Beneduzi, A. Ambrosini, and L. M. P. Passaglia, "Plant growth-  
7   promoting rhizobacteria (PGPR): Their potential as antagonists and  
8   biocontrol agents," *Genetics and Molecular Biology*, vol. 35, no. 4 SUPPL.  
9   pp. 1044–1051, 2012.
- 10   [3]   M. Ongena and P. Jacques, *Bacillus lipopeptides: versatile weapons for  
11   plant disease biocontrol*, vol. 16, no. 3. 2008, pp. 115–125.
- 12   [4]   E. Sansinenea and A. Ortiz, "Secondary metabolites of soil *Bacillus* spp.,"  
13   *Biotechnol. Lett.*, vol. 33, no. 8, pp. 1523–1538, 2011.
- 14   [5]   T. Stein, "Bacillus subtilis antibiotics: structures, syntheses and specific  
15   functions," *Mol. Microbiol.*, vol. 56, no. 4, pp. 845–857, Mar. 2005.
- 16   [6]   X. H. Chen *et al.*, "Genome analysis of *Bacillus amyloliquefaciens* FZB42  
17   reveals its potential for biocontrol of plant pathogens," *J. Biotechnol.*, vol.  
18   140, no. 1, pp. 27–37, Mar. 2009.
- 19   [7]   B. R. Glick, "The enhancement of plant growth by free-living bacteria,"  
20   *Canadian Journal of Microbiology*, vol. 41, no. 2. Canadian Science  
21   Publishing, pp. 109–117, 1995.
- 22   [8]   C. A. Dunlap, S.-J. Kim, S.-W. Kwon, and A. P. Rooney, "*Bacillus*  
23   velezensis is not a later heterotypic synonym of *Bacillus*  
24   amyloliquefaciens; *Bacillus methylotrophicus*, *Bacillus amyloliquefaciens*  
25   subsp. plantarum and 'Bacillus oryzicola' are later heterotypic synonyms  
26   of *Bacillus velezensis* based on phylogenomics," *Int. J. Syst. Evol.*  
27   *Microbiol.*, vol. 66, no. 3, pp. 1212–1217, Mar. 2016.
- 28   [9]   S.-J. Kim, S.-W. Kwon, C. A. Dunlap, and A. P. Rooney, "Phylogenomic  
29   analysis shows that *Bacillus amyloliquefaciens* subsp. plantarum is a later  
30   heterotypic synonym of *Bacillus methylotrophicus*," *Int. J. Syst. Evol.*  
31   *Microbiol.*, vol. 65, no. 7, pp. 2104–2109, Jul. 2015.
- 32   [10]   B. Fan, J. Blom, H. P. Klenk, and R. Borriss, "*Bacillus amyloliquefaciens*,  
33   *Bacillus velezensis*, and *Bacillus siamensis* Form an 'Operational Group  
34   *B. amyloliquefaciens*' within the *B. subtilis* species complex," *Front.*  
35   *Microbiol.*, vol. 8, no. JAN, 2017.
- 36   [11]   D. Fritze, "Taxonomy of the genus *Bacillus* and related genera: The  
37   aerobic endospore-forming bacteria," *Phytopathology*, vol. 94, no. 11, pp.  
38   1245–1248, 2004.
- 39   [12]   M. Richter and R. Rossello-Mora, "Shifting the genomic gold standard for  
40   the prokaryotic species definition," *Proc. Natl. Acad. Sci.*, vol. 106, no. 45,  
41   pp. 19126–19131, Nov. 2009.
- 42   [13]   O. Genilloud, "Mining Actinomycetes for Novel Antibiotics in the Omics  
43   Era: Are We Ready to Exploit This New Paradigm?," *Antibiot. (Basel,  
44   Switzerland)*, vol. 7, no. 4, Sep. 2018.
- 45   [14]   J. Zhang, R. Chiodini, A. Badr, and G. Zhang, "The impact of next-  
46   generation sequencing on genomics," *Journal of genetics and genomics =  
47   Yi chuan xue bao*, vol. 38, no. 3. pp. 95–109, Mar-2011.
- 48   [15]   K. V. Voelkerding, S. A. Dames, and J. D. Durtschi, "Next-Generation  
49   Sequencing: From Basic Research to Diagnostics," *Clin. Chem.*, vol. 55,  
50   no. 4, pp. 641–658, Apr. 2009.

- 1 [16] L. Acuña-Amador, A. Primot, E. Cadieu, A. Roulet, and F. Barloy-Hubler,  
2 "Genomic repeats, misassembly and reannotation: A case study with  
3 long-read resequencing of *Porphyromonas gingivalis* reference strains,"  
4 *BMC Genomics*, vol. 19, no. 1, Jan. 2018.
- 5 [17] S. Ciufu *et al.*, "Using average nucleotide identity to improve taxonomic  
6 assignments in prokaryotic genomes at the NCBI," *Int. J. Syst. Evol.*  
7 *Microbiol.*, vol. 68, no. 7, 2018.
- 8 [18] N. Nagarajan *et al.*, "Finishing genomes with limited resources: Lessons  
9 from an ensemble of microbial genomes," *BMC Genomics*, vol. 11, no. 1,  
10 Apr. 2010.
- 11 [19] K. J. Grubbs *et al.*, "Large-Scale Bioinformatics Analysis of *Bacillus*  
12 Genomes Uncovers Conserved Roles of Natural Products in Bacterial  
13 Physiology," *mSystems*, vol. 2, no. 6, 2017.
- 14 [20] T. Weber, "In silico tools for the analysis of antibiotic biosynthetic  
15 pathways," *International Journal of Medical Microbiology*, vol. 304, no. 3–  
16 4. pp. 230–235, May-2014.
- 17 [21] K. J. Perez *et al.*, "*Bacillus* spp. isolated from puba as a source of  
18 biosurfactants and antimicrobial lipopeptides," *Front. Microbiol.*, vol. 8, no.  
19 JAN, pp. 1–14, 2017.
- 20 [22] H. P. Bais, R. Fall, and J. M. Vivanco, "Biocontrol of *Bacillus subtilis*  
21 against Infection of *Arabidopsis* Roots by *Pseudomonas syringae* Is  
22 Facilitated by Biofilm Formation and Surfactin Production," *Plant Physiol.*,  
23 vol. 134, no. 1, pp. 307–319, Jan. 2004.
- 24 [23] B. Kluge, J. Vater, J. Salnikow, and K. Eckart, "Studies on the  
25 biosynthesis of surfactin, a lipopeptide antibiotic from *Bacillus subtilis*  
26 ATCC 21332," *FEBS Lett.*, vol. 231, no. 1, pp. 107–110, Apr. 1988.
- 27 [24] M. Ongena *et al.*, "Surfactin and fengycin lipopeptides of *Bacillus subtilis*  
28 as elicitors of induced systemic resistance in plants," *Environ. Microbiol.*,  
29 vol. 9, no. 4, pp. 1084–1090, Apr. 2007.
- 30 [25] S. M. Lim *et al.*, "Diffusible and volatile antifungal compounds produced  
31 by an antagonistic *Bacillus velezensis* G341 against various  
32 phytopathogenic fungi," *Plant Pathol. J.*, vol. 33, no. 5, pp. 488–498,  
33 2017.
- 34 [26] N. Vanittanakom, W. Loeffler, U. Koch, and G. Jung, "Fengycin-A Novel  
35 Antifungal Lipopeptide Antibiotic Produced by *Bacillus Subtilis* F-29-3," *J.*  
36 *Antibiot. (Tokyo)*, vol. 39, no. 7, pp. 888–901, 1986.
- 37 [27] M. Ongena, P. Jacques, Y. Touré, J. Destain, A. Jabrane, and P. Thonart,  
38 "Involvement of fengycin-type lipopeptides in the multifaceted biocontrol  
39 potential of *Bacillus subtilis*," *Appl. Microbiol. Biotechnol.*, vol. 69, no. 1,  
40 pp. 29–38, Nov. 2005.
- 41 [28] M. Deleu, M. Paquot, and T. Nylander, "Effect of fengycin, a lipopeptide  
42 produced by *Bacillus subtilis*, on model biomembranes," *Biophys. J.*, vol.  
43 94, no. 7, pp. 2667–2679, 2008.
- 44 [29] G. Y. Yu, J. B. Sinclair, G. L. Hartman, and B. L. Bertagnolli, "Production  
45 of iturin A by *Bacillus amyloliquefaciens* suppressing *Rhizoctonia solani*,"  
46 *Soil Biol. Biochem.*, vol. 34, no. 7, pp. 955–963, 2002.
- 47 [30] H. Abriouel, C. M. A. P. Franz, N. Ben Omar, and A. Galvez, "Diversity  
48 and applications of *Bacillus bacteriocins*," *FEMS Microbiol. Rev.*, vol. 35,  
49 no. 1, pp. 201–232, Jan. 2011.
- 50 [31] A. M. Maxam and W. Gilbert, "A new method for sequencing DNA," *Proc.*

- 1 *Natl. Acad. Sci. U. S. A.*, vol. 74, no. 2, pp. 560–564, Feb. 1977.
- 2 [32] J. M. Heather and B. Chain, “The sequence of sequencers: The history of  
3 sequencing DNA.,” *Genomics*, vol. 107, no. 1, pp. 1–8, Jan. 2016.
- 4 [33] W. Ansorge, B. S. Sproat, J. Stegemann, and C. Schwager, “A non-  
5 radioactive automated method for DNA sequence determination,” *J.*  
6 *Biochem. Biophys. Methods*, vol. 13, no. 6, pp. 315–323, Dec. 1986.
- 7 [34] J. A. Luckey *et al.*, “High speed DNA sequencing by capillary  
8 electrophoresis,” *Nucleic Acids Res.*, vol. 18, no. 15, pp. 4417–4421, Aug.  
9 1990.
- 10 [35] H. Kambara, T. Nishikawa, Y. Katayama, and T. Yamaguchi,  
11 “Optimization of parameters in a DNA sequenator using fluorescence  
12 detection,” *Bio/Technology*, vol. 6, no. 7, pp. 816–821, Jul. 1988.
- 13 [36] P. Nyérén and A. Lundin, “Enzymatic method for continuous monitoring of  
14 inorganic pyrophosphate synthesis,” *Anal. Biochem.*, vol. 151, no. 2, pp.  
15 504–509, Dec. 1985.
- 16 [37] J. M. Rothberg *et al.*, “An integrated semiconductor device enabling non-  
17 optical genome sequencing,” *Nature*, vol. 475, no. 7356, pp. 348–352,  
18 Jul. 2011.
- 19 [38] A. Rhoads and K. F. Au, “PacBio Sequencing and Its Applications,”  
20 *Genomics, Proteomics and Bioinformatics*, vol. 13, no. 5. Beijing  
21 Genomics Institute, pp. 278–289, 2015.
- 22 [39] S. Koren and A. M. Phillippy, “One chromosome, one contig: Complete  
23 microbial genomes from long-read sequencing and assembly,” *Current*  
24 *Opinion in Microbiology*, vol. 23. Elsevier Ltd, pp. 110–120, 01-Feb-2015.
- 25 [40] P. Zhu and H. G. Craighead, “Zero-Mode Waveguides for Single-  
26 Molecule Analysis,” *Annu. Rev. Biophys.*, vol. 41, no. 1, pp. 269–293,  
27 Jun. 2012.
- 28 [41] M. Jain, H. E. Olsen, B. Paten, and M. Akeson, “The Oxford Nanopore  
29 MinION: delivery of nanopore sequencing to the genomics community,”  
30 *Genome Biol.*, vol. 17, no. 1, pp. 1–11, Dec. 2016.
- 31 [42] M. Loose, S. Malla, and M. Stout, “Real-time selective sequencing using  
32 nanopore technology,” *Nat. Methods*, vol. 13, no. 9, pp. 751–754, Aug.  
33 2016.
- 34 [43] A. L. Greninger *et al.*, “Rapid metagenomic identification of viral  
35 pathogens in clinical samples by real-time nanopore sequencing  
36 analysis,” *Genome Med.*, vol. 7, no. 1, pp. 1–13, Sep. 2015.
- 37 [44] D. A. Benson, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L.  
38 Wheeler, “GenBank.,” *Nucleic Acids Res.*, vol. 35, no. Database issue,  
39 pp. D21–5, Jan. 2007.
- 40 [45] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic  
41 local alignment search tool,” *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–410,  
42 Oct. 1990.
- 43 [46] G. M. Garrity *et al.*, “Report of the ad hoc committee for the re-evaluation  
44 of the species definition in bacteriology.,” *Int. J. Syst. Evol. Microbiol.*, vol.  
45 52, no. 3, pp. 1043–1047, May 2002.
- 46 [47] P. C. Y. Woo, S. K. P. Lau, J. L. L. Teng, H. Tse, and K.-Y. Yuen, “Then  
47 and now: use of 16S rDNA gene sequencing for bacterial identification  
48 and discovery of novel bacteria in clinical microbiology laboratories.,” *Clin.*  
49 *Microbiol. Infect.*, vol. 14, no. 10, pp. 908–34, Oct. 2008.
- 50 [48] C. R. Woese, O. Kandler, and M. L. Wheelis, “Towards a natural system

- 1 of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.,"  
2 *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87, no. 12, pp. 4576–9, Jun. 1990.
- 3 [49] T. Weber *et al.*, "antiSMASH 3.0—a comprehensive resource for the  
4 genome mining of biosynthetic gene clusters," *Nucleic Acids Research*,  
5 vol. 43, no. Web Server issue. pp. W237-43, Jul-2015.
- 6 [50] A. J. van Heel, A. de Jong, M. Montalbán-López, J. Kok, and O. P.  
7 Kuipers, "BAGEL3: automated identification of genes encoding  
8 bacteriocins and (non-)bactericidal posttranslationally modified peptides,"  
9 *Nucleic Acids Res.*, vol. 41, no. W1, pp. W448–W453, Jul. 2013.
- 10 [51] A. de Jong, S. A. F. T. van Hijum, J. J. E. Bijlsma, J. Kok, and O. P.  
11 Kuipers, "BAGEL: a web-based bacteriocin genome mining tool," *Nucleic*  
12 *Acids Res.*, vol. 34, no. Web Server, pp. W273–W279, Jul. 2006.
- 13 [52] A. de Jong, A. J. van Heel, J. Kok, and O. P. Kuipers, "BAGEL2: mining  
14 for bacteriocins in genomic data," *Nucleic Acids Res.*, vol. 38, no.  
15 suppl\_2, pp. W647–W651, Jul. 2010.
- 16 [53] A. Starcevic, J. Zucko, J. Simunkovic, P. F. Long, J. Cullum, and D.  
17 Hranueli, "ClustScan : an integrated program package for the semi-  
18 automatic annotation of modular biosynthetic gene clusters and in silico  
19 prediction of novel chemical structures," *Nucleic Acids Res.*, vol. 36, no.  
20 21, pp. 6882–6892, Dec. 2008.
- 21 [54] N. Ziemert, S. Podell, K. Penn, J. H. Badger, E. Allen, and P. R. Jensen,  
22 "The Natural Product Domain Seeker NaPDoS: A Phylogeny Based  
23 Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity,"  
24 *PLoS One*, vol. 7, no. 3, p. e34064, Mar. 2012.
- 25 [55] N. Khaldi *et al.*, "SMURF: Genomic mapping of fungal secondary  
26 metabolite clusters," *Fungal Genet. Biol.*, vol. 47, no. 9, pp. 736–741,  
27 Sep. 2010.
- 28 [56] M. H. Li, P. M. Ung, J. Zajkowski, S. Garneau-Tsodikova, and D. H.  
29 Sherman, "Automated genome mining for natural products," *BMC*  
30 *Bioinformatics*, vol. 10, no. 1, p. 185, Jun. 2009.
- 31 [57] D. López, H. Vlamakis, and R. Kolter, "Biofilms.," *Cold Spring Harbor*  
32 *perspectives in biology*, vol. 2, no. 7. 01-Jul-2010.
- 33 [58] H. P. Bais, T. L. Weir, L. G. Perry, S. Gilroy, and J. M. Vivanco, "The role  
34 of root exudates in rhizosphere interactions with plants and another  
35 organisms," *Annu. Rev. Plant Biol.*, vol. 57, no. 1, pp. 233–266, Jun.  
36 2006.
- 37 [59] T. Danhorn and C. Fuqua, "Biofilm Formation by Plant-Associated  
38 Bacteria," *Annu. Rev. Microbiol.*, vol. 61, no. 1, pp. 401–422, Oct. 2007.
- 39 [60] K. Lewis, "Persister cells and the riddle of biofilm survival," *Biochemistry*  
40 *(Moscow)*, vol. 70, no. 2. pp. 267–274, Feb-2005.
- 41 [61] C. Fuqua and E. P. Greenberg, "Listening in on bacteria: Acyl-homoserine  
42 lactone signalling," *Nature Reviews Molecular Cell Biology*, vol. 3, no. 9.  
43 Nature Publishing Group, pp. 685–695, 2002.
- 44 [62] H. L. Wei and L. Q. Zhang, "Quorum-sensing system influences root  
45 colonization and biological control ability in *Pseudomonas fluorescens*  
46 2P24," *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.*, vol. 89, no.  
47 2, pp. 267–280, Feb. 2006.
- 48 [63] S. De Weert *et al.*, "Flagella-driven chemotaxis towards exudate  
49 components is an important trait for tomato root colonization by  
50 *Pseudomonas fluorescens*," *Mol. Plant-Microbe Interact.*, vol. 15, no. 11,

- 1 pp. 1173–1180, Nov. 2002.
- 2 [64] B. J. J. Lugtenberg, T. F. C. Chin-A-Woeng, and G. V. Bloemberg,  
3 “Microbe-plant interactions: Principles and mechanisms,” *Antonie van*  
4 *Leeuwenhoek, Int. J. Gen. Mol. Microbiol.*, vol. 81, no. 1–4, pp. 373–383,  
5 2002.
- 6 [65] N. Ausmees, K. Jacobsson, and M. Lindberg, “A unipolarly located, cell-  
7 surface-associated agglutinin, RapA, belongs to a family of Rhizobium-  
8 adhering proteins (Rap) in *Rhizobium leguminosarum* bv. *trifolii*,”  
9 *Microbiology*, vol. 147, no. 3, pp. 549–559, Mar. 2001.
- 10 [66] R. F. Kinsinger, M. C. Shirk, and R. Fall, “Rapid surface motility in *Bacillus*  
11 *subtilis* is dependent on extracellular surfactin and potassium ion,” *J.*  
12 *Bacteriol.*, vol. 185, no. 18, pp. 5627–5631, Sep. 2003.
- 13 [67] R. M. Donlan, “Biofilms: Microbial life on surfaces,” *Emerging Infectious*  
14 *Diseases*, vol. 8, no. 9. Centers for Disease Control and Prevention  
15 (CDC), pp. 881–890, 2002.
- 16 [68] J. W. Kloepper, R. Lifshitz, and R. M. Zablotowicz, “Free-living bacterial  
17 inocula for enhancing crop productivity,” *Trends in Biotechnology*, vol. 7,  
18 no. 2. Elsevier Current Trends, pp. 39–44, 01-Feb-1989.
- 19 [69] A. R. Podile and G. K. Kishore, “Microbe–plant interactions: principles and  
20 mechanisms,” in *Plant-Associated Bacteria*, Springer Netherlands, 2006,  
21 pp. 195–230.
- 22 [70] M. J. Torres, C. P. Brandan, G. Petroselli, R. Erra-Balsells, and M. C.  
23 Audisio, “Antagonistic effects of *Bacillus subtilis* subsp. *subtilis* and *B.*  
24 *amyloliquefaciens* against *Macrophomina phaseolina*: SEM study of  
25 fungal changes and UV-MALDI-TOF MS analysis of their bioactive  
26 compounds,” *Microbiol. Res.*, vol. 182, pp. 31–39, 2016.
- 27 [71] C. Rückert, J. Blom, X. H. Chen, O. Reva, and R. Borriss, “Genome  
28 sequence of *B. amyloliquefaciens* type strain DSM7T reveals differences  
29 to plant-associated *B. amyloliquefaciens* FZB42,” *J. Biotechnol.*, vol. 155,  
30 no. 1, pp. 78–85, 2011.
- 31 [72] J. J. Arroyave-Toro, S. Mosquera, and V. Villegas-Escobar, “Biocontrol  
32 activity of *Bacillus subtilis* EA-CB0015 cells and lipopeptides against  
33 postharvest fungal pathogens,” *Biol. Control*, vol. 114, pp. 195–200, 2017.
- 34 [73] L. Belbahri *et al.*, “Comparative Genomics of *Bacillus amyloliquefaciens*  
35 Strains Reveals a Core Genome with Traits for Habitat Adaptation and a  
36 Secondary Metabolites Rich Accessory Genome.,” *Front. Microbiol.*, vol.  
37 8, p. 1438, 2017.
- 38 [74] C. R. Harwood, “Introduction to the Biotechnology of *Bacillus*,” in *Bacillus*,  
39 Springer US, 1989, pp. 1–4.
- 40 [75] F. Kunst *et al.*, “The complete genome sequence of the gram-positive  
41 bacterium *Bacillus subtilis*,” *Nature*, vol. 390, no. 6657, pp. 249–256, Nov.  
42 1997.
- 43 [76] W. L. Nicholson, N. Munakata, G. Horneck, H. J. Melosh, and P. Setlow,  
44 “Resistance of *Bacillus* Endospores to Extreme Terrestrial and  
45 Extraterrestrial Environments,” *Microbiol. Mol. Biol. Rev.*, vol. 64, no. 3,  
46 pp. 548–572, Sep. 2000.
- 47 [77] W. L. Nicholson, “Roles of *Bacillus* endospores in the environment,”  
48 *Cellular and Molecular Life Sciences*, vol. 59, no. 3. Springer, pp. 410–  
49 416, 2002.
- 50 [78] C. Pandin *et al.*, “Complete genome sequence of *Bacillus velezensis*

- 1 QST713: A biocontrol agent that protects *Agaricus bisporus* crops against  
2 the green mould disease,” *J. Biotechnol.*, vol. 278, no. 2010, pp. 10–19,  
3 Jul. 2018.
- 4 [79] I. Zouari, L. Jlaiel, S. Tounsi, and M. Trigui, “Biocontrol activity of the  
5 endophytic *Bacillus amyloliquefaciens* strain CEIZ-11 against *Pythium*  
6 *aphanidermatum* and purification of its bioactive compounds,” *Biol.*  
7 *Control*, vol. 100, pp. 54–62, 2016.
- 8 [80] X. L. Lu *et al.*, “Macrolactin S, a novel macrolactin antibiotic from marine  
9 *Bacillus* sp.,” *Nat. Prod. Res.*, vol. 22, no. 4, pp. 342–347, Mar. 2008.
- 10 [81] E. A. Dertz, J. Xu, A. Stintzi, and K. N. Raymond, “Bacillibactin-mediated  
11 iron transport in *Bacillus subtilis*,” *J. Am. Chem. Soc.*, vol. 128, no. 1, pp.  
12 22–23, Jan. 2006.
- 13 [82] X. H. Chen *et al.*, “Difficidin and bacilysin produced by plant-associated  
14 *Bacillus amyloliquefaciens* are efficient in controlling fire blight disease,” *J.*  
15 *Biotechnol.*, vol. 140, no. 1–2, pp. 38–44, Mar. 2009.
- 16