

GENÔMICA DE BACTÉRIAS PROMOTORAS DO CRESCIMENTO VEGETAL
ISOLADAS DE VERMICOMPOSTOS

FILIFE PEREIRA MATTEOLI

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO -
UENF

CAMPOS DOS GOYTACAZES - RJ

MARÇO – 2018

GENÔMICA DE BACTÉRIAS PROMOTORAS DO CRESCIMENTO VEGETAL
ISOLADAS DE VERMICOMPOSTOS

FILIPE PEREIRA MATTEOLI

Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Biociências e Biotecnologia.

Orientador: Prof. Dr. Thiago Motta Venancio

Co-orientador: Prof. Dr. Fabio Lopes Olivares

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO –

UENF

CAMPOS DOS GOYTACAZES - RJ

MARÇO – 2018

GENÔMICA DE BACTÉRIAS PROMOTORAS DO CRESCIMENTO VEGETAL
ISOLADAS DE VERMICOMPOSTOS

FILIPE PEREIRA MATTEOLI

Tese apresentada ao Centro de
Biotecnologia e Biociências da
Universidade Estadual do Norte
Fluminense Darcy Ribeiro, como parte
das exigências para obtenção do título
de Doutor em Biociências e
Biotecnologia.

Orientador: Prof. Dr. Thiago Motta
Venancio

Co-orientador: Prof. Dr. Fabio Lopes
Olivares

Aprovado em 26/03/2018.

Comissão Examinadora:

Prof. Dr. Thiago Motta Venancio - UENF

Prof. Dr. Fabio Lopes Olivares - UENF

Prof. Dr. Renato Augusto daMatta - UENF

Profa. Dra. Marília Amorim Berbert de Molina - UENF

Prof. Dr. José Roberto da Silva - UFRJ

CAMPOS DOS GOYTACAZES - RJ

MARÇO – 2018

AGRADECIMENTOS

À minha mãe Rosa, exemplo de luta e raça no cotidiano. Deus foi generoso demais comigo ao me conceber seu filho. Aos meus familiares pelo apoio e carinho. À minha namorada Lorena, pessoa especial que ilumina minha vida.

Aos colegas de laboratório Daniel, Lupis, Rajesh, Kanhu, Gustavo, Francisnei, Hemanoel, Fabrício pelas inúmeras contribuições ao longo dessa jornada. Agradeço também aos colegas de outros grupos como Régis, Kamila e Pollyanna do NUDIBA.

Aos professores Thiago Motta Venancio e Fabio Lopes Olivares por terem me aceitado no programa e dedicado um pouco de suas vidas à minha formação. Agradeço também a tantos outros professores com quem colaborei nessa caminhada: Clícia Grativol, Gonçalo Apolinário, Jorge Hernandez, Valdirene Gomes entre outros.

A todos os funcionários da UENF, que insistem no sonho de manter a chama da educação superior pública e de qualidade na planície campista a despeito das circunstâncias.

Ao Prof. Arnaldo Façanha e ao corpo administrativo do Programa de Pós-Graduação em Biociências e Biotecnologia.

Às agências de fomento FAPERJ, CNPq e CAPES.

SUMÁRIO

1. INTRODUÇÃO	1
1.1 Panorama agrícola internacional	1
1.2 A interação planta-microrganismo	2
1.3 Mecanismos de promoção direta de crescimento.....	4
1.3.1 Solubilização de fosfato e zinco.....	4
1.3.2 A solubilização de fosfato inorgânico.....	5
1.3.3 A solubilização de fosfato orgânico	5
1.3.4 Fixação de nitrogênio.....	7
1.3.5 Produção de hormônios vegetais	7
1.4 Mecanismos de promoção indireta de crescimento.....	9
1.4.1 Biocontrole.....	9
1.4.2 Quitinases.....	11
1.4.3 Resistência a metais pesados	12
1.5 Biotecnologia de inoculantes	12
1.5.1 Formulação de inoculantes microbianos.....	14
1.6 Tecnologia de vermicompostagem	16
1.6.1 As substâncias húmicas	17
1.8 Sequenciamento de nova geração	20
1.9 Montagem de genomas	21
1.9.1 Algoritmos <i>greedy</i>	23
1.9.2 Overlap layout consensus (OLC).....	23
1.9.3 Grafos de Bruijn.....	24
3. OBJETIVO GERAL	27
3.1 OBJETIVOS ESPECÍFICOS.....	27
4. METODOLOGIA.....	28
4.1 Isolamento das bactérias	28
4.2 Seleção de bactérias enviadas para o sequenciamento.....	28
4.3 Extração do DNA	29
4.4 Sequenciamento dos genomas	29
4.5 Qualidade e pré-processamento dos <i>reads</i>	29
4.6 Priorização de genomas	30
4.7 Montagem dos genomas	30
4.8 Anotação dos genomas	31
4.8.1 Anotação de operons.....	32
4.8.2 Anotação de vias metabólicas	33

4.9	Visualização e integração	33
4.10	Análise filogenéticas	33
4.10.1	Obtenção das sequências de RNAr 16s diretamente dos <i>reads</i>	34
4.10.2	Reconstrução filogenética.....	34
4.10.3	Análise taxonômica a partir dos <i>reads</i>	34
4.11	Pan-genoma	35
4.12	Inspeção manual de regiões genômicas.....	36
4.13	Testes <i>in vivo</i>	36
4.14	Testes <i>in vitro</i>	37
4.14.1	Solubilização de P e Zn	37
4.14.2	Produção de compostos indólicos	38
4.14.3	Inibição do crescimento de fungos fitopatogênicos.....	38
5.	RESULTADOS E DISCUSSÃO.....	39
5.1	Testes preditivos <i>in vitro</i>	39
5.2	Sequenciamento	40
5.3	Análises da qualidade dos <i>reads</i> sequenciados.	40
5.4	Inferências filogenéticas obtidas a partir dos <i>reads</i> - 16s rRNA	41
5.5	Inferências filogenéticas obtidas diretamente a partir dos <i>reads</i>	43
5.6	Montagem dos genomas	44
5.6.1	Presença de genes ortólogos conservados	45
5.7	Resultados da cepa <i>Serratia marcescens</i> UENF-22GI.....	47
5.8	Caracterização <i>in vitro</i> do potencial promotor do crescimento vegetal de SMU.....	48
5.10	Reconstrução do pan-genoma de SMU.....	53
5.11	Análise filogenética de SMU	55
5.12	Correlação dos achados <i>in vitro</i> com genes presentes em SMU	56
5.12.1	Genes relacionados a solubilização de P e Zn	58
5.12.2	Tolerância a metais tóxicos	59
5.12.3	IAA e espermidina.....	60
5.12.4	Formação de biofilme	60
5.12.5	Biocontrole e competição.....	61
5.13	Análise de transferência horizontal.....	62
5.14	Inferência taxonômica por ANI em <i>Enterobacter</i> sp.	64
5.15	Testes <i>in vivo</i>	65
6.	CONSIDERAÇÕES FINAIS	67
7.	CONCLUSÕES	68
8.	REFERÊNCIAS.....	69

ÍNDICE DE FIGURAS E TABELAS

Figura 1. Mapa de calor mostrando a tendência de consumo de fertilizantes (a), inseticidas (b), herbicidas (c) e fungicidas e bactericidas (d) em todo o mundo, tendo como base dados de 1961 a 2010. S – sem dados; A – aumentando; A+E – aumentou e estagnou; N – neutro; A+D – aumentou e diminuiu; D – diminuiu. Adaptado de Liu et al., 2015.....	1
Figura 2. Principais vias de síntese de auxina descritas em bactérias. Rotas metabólicas descritas para a síntese de auxina em bactérias, destaca-se a relevância do Triptofano como precursor da biossíntese. Adaptado de Spaepen et al., (2007a).....	8
Figura 3. Mecanismos principais envolvidos no biocontrole de patógenos e indução da resposta imune sistêmica (RIS) do vegetal. Adaptado de Haas e Défago (2005).	10
Figura 4. Mecanismos moleculares provavelmente envolvidos na colonização e promoção do crescimento de plantas identificados no genoma de <i>H. seropedicae</i> SmR1 Pedrosa et al., 2011.	19
Figura 5. Representação esquemática do sequenciamento por síntese empregado pelos equipaentos Illumina. a) Preparo e amplificação das amostras. b) Sequenciamento por síntese. c) Fluorescência gerada em cores de acordo com o nucleotídeo adicionado à fita. Adaptado de Metzker 2010.	21
Figura 6. Grafos de sobreposição não orientados (a). Um subgrafo orientado (b). Adaptado de Kececioğlu e Myers 1995.....	24
Figura 7. Representação esquemática dos Grafos de Brujin. a) DNA a ser sequenciado (10 pb). b) Reads curtos ilustrativos obtidos pelo sequenciamento (7 pb) e a esquematização da problemática da reconstrução da sequência original. c) Grafo obtido ao se utilizar um k-mer = 3. d) Ordenação dos k-mers reconstruindo a sequência original. Adaptado de Compeau et al., 2001.	26
Figura 8. Componentes principais do PGAP. Adaptado de Tatusova et al., 2016.	32
Figura 9. Boxplot com a distribuição de qualidade de uma das bibliotecas sequenciadas (22GI forward reads) ao longo das coordenadas dentro dos reads. Esta figura é representativa de todas as bibliotecas sequenciadas no projeto.	41
Figura 10. Representação circular da filiação taxonômica dos reads totais computados com KAIJU (Menzel et al., 2016) e visualizados através do KRONA (Ondov et al., 2011).....	44
Figura 11. Presença de genes ortólogos únicos da ordem <i>enterobacteriales</i> nos isolados 6GII, 14GII, 21GI, 22GI, 23GI, 24GI, 25GI e <i>bacilales</i> em 25GII.....	46
Figura 12. Visualização circular do genoma de <i>S. marcescens</i> UENF-22GI. O conteúdo de GC ao longo de todo o genoma (círculo azul mais externo). O N50, que coincide com o maior scaffold (em vermelho) e o N90 em laranja.	47
Figura 13. a) Solubilização de fosfato de cálcio ($\text{Ca}_3(\text{PO}_4)_2$). b) Solubilização de óxido de zinco (ZNO). c) Quantificação da solubilização de fosfato de cálcio e da rocha de fluorapatita (rock P). d) Variação do pH após o tratamento de fosfato de cálcio e rocha de Araxá com SMU. As barras pretas são os sistemas controle, as cinzas os inoculados na presença de SMU.	50
Figura 14. a) Vias de produção de auxina conhecidas em bactérias, em vermelho o gene presente em SMU. b) Produção de compostos indólicos por SMU com	

suplementação de Trp ou não.....	51
Figura 15. Ensaio de inibição de duas espécies de fungos fitopatogênicos por SMU. Sistemas controle foram compostos somente pelos fungos <i>F. oxysporum</i> e <i>F. solani</i> (a, d, respectivamente. No sistema de co-inoculação SMU foi inserida em 4 pontos equidistantes de <i>F. oxysporum</i> e <i>F. solani</i> (b, e, respectivamente). A aderência da bactéria às hifas fúngicas foi demonstrada através de microscopia óptica em <i>F. oxysporum</i> e <i>F. solani</i> (c, d, respectivamente)	52
Figura 16. Efeito da inoculação de SMU no crescimento de plântulas de milho. Plântulas germinadas (com comprimento radicular de 2 a 2,5 cm) foram transferidas para tubos de vidro contendo vermiculita estéril (uma plântula por tubo). A inoculação foi feita através da aplicação de 1 mL de suspensão bacteriana (10^8 células·mL ⁻¹) sobre as plântulas. Sistemas inoculados com 1 mL de meio Dygs estéril foram usados como controle negativos. As seguintes medidas foram aferidas após 10 dias transcorridos: Massa fresca da raiz (MFR), massa fresca da parte aérea (MFPA), massa seca da raiz (MSR), massa seca parte aérea (MSPA), comprimento radicular (CR) e altura (A). Todos os experimentos foram conduzidos em triplicata e diferenças significativas foram aferidas pelo teste de Tukey ($p < 0.05$).	53
Figura 17. Pan-genoma de <i>Serratia marcescens</i> . Em verde isolados pertencentes a coleções de bactérias, em laranja cepas clínicas e em roxo isolados não clínicos, inclusive SMU grafada em vermelho possuindo 57 genes únicos.....	55
Figura 18. Árvore filogenética multilocus computada por máxima verossimilhança dos 35 isolados de <i>S. marcescens</i> . As sequências das 10 proteínas foram obtidas e alinhadas separadamente, posteriormente os alinhamentos foram concatenados e a filogenia inferida utilizando-se RAxML (Stamatakis, 2014) com suporte estatístico de bootstrap de 1000 vezes representado em porcentagem. Em laranja isolados clínicos, em verde, pertencentes à coleções e em roxo não clínicos.	56
Figura 19. Operons correlacionados com funções PGPR presentes no genoma de SMU. a) Biossíntese de PQQ. b) Transportador Pst específico para fosfato. c) e d) Produção de biofilme, respectivamente poli-N-acetilglicosamina e celulose. e) Prodigiosina. f) Sistema de secreção tipo 6.	58
Figura 20. Análise de transferência horizontal. a) Alinhamento do genoma de SMU com genomas de referência. b) Sintenia da região do Gap1 mostrando os genes associados à produção do peptídeo não ribossomal e seus respectivos domínios.....	64
Figura 21. Matriz de ANI representativa. Ilustra a metodologia aplicada para obtenção de isolados filogeneticamente mais próximos das amostras pré-classificadas como <i>Enterobacter</i> spp.. Os 1.357 genomas depositados foram divididos em 6 devido a questões computacionais. a) Output de 235 genomas, os círculos em amarelo destacam clusters nos quais nossos isolados estão inseridos. b) Cluster de 21GII. c) Cluster de 6GII, 14GII, 23GI, 24GI e 25GI. ...	65
Figura 22. Promoção do crescimento de plântulas de milho pelos isolados 6GII, 14GII, 21GII, 23GI, 24GI e 25GI. Plântulas germinadas (com comprimento radicular de 2 a 2,5 cm) foram transferidas para tubos de vidro contendo vermiculita estéril (uma plântula por tubo). A inoculação foi feita através da aplicação de 1 mL de suspensão bacteriana (10^8 células·mL ⁻¹) sobre as plântulas. a) Altura. b) Comprimento radicular. c) Massa fresca da raiz. d) Massa	

fresca parte aérea. e) Matéria seca da raiz. f) Massa seca parte aérea. Sistemas inoculados com 1 mL de meio Dygs estéril foram usados como controle negativos, as medições foram feitas após 10 dias. Os experimentos foram realizados em triplicata..... 66

TABELAS:

Tabela 1. O score de qualidade phred está logaritmicamente correlacionado a probabilidade de erro.	30
Tabela 2. Perfil bioquímico da promoção do crescimento em testes <i>in vitro</i>	39
Tabela 3. Quantidade e qualidade dos reads obtidos pelo sequenciamento. ..	40
Tabela 4. Tamanho das sequências do gene do rRNA 16s obtidas a partir dos reads, hit mais provável com respectivo NR e grau de identidade.	42
Tabela 5. Estatísticas das montagens obtidas para os isolados selecionados	45
Tabela 6. Comparativo do genoma de SMU e outras cepas de <i>S. marcescens</i> filogeneticamente próximas.	48
Tabela 7. Genomas públicos utilizados na construção do pan-genoma em associação com SMU.	54
Tabela 8. Genes promotores do crescimento vegetal em SMU.	57

LISTA DE ABREVIATURAS

AIA – ácido indol acético

IAM – via do indol-3-acetamida

AIP – via do Ácido indol pirúvico

dNTPs - desoxinucleotídeos trifosfatados

FBN – fixação biológica de nitrogênio

gdh – glicose desidrogenase (gene)

IAN – via do indol-3-acetonitrila

ipdc – indol-3-piruvato-descarboxilase

MP – metais pesados

NAG - n-acetilglicosamina

pb - pares de bases

PGAP - *prokaryotic Genome Annotation Pipeline* (pipeline para anotação de genomas procarióticos)

PGPR – *plant growth promoting rhizobacteria* (bactéria promotora do crescimento de plantas)

ORF - *open reading frames* (fase aberta de leitura)

pqq – pirroloquinolina-quinona (gene)

Pqq – pirroloquinolina-quinona (proteína)

NGS – *next generation sequencing* (sequenciamento de nova geração)

SHs - substâncias húmicas

SMU – *Serratia marcescens* UENF-22GI

TAM - via da triptamina

RESUMO

O crescimento populacional contínuo gera crescente demanda por alimentos, tornando imperativa a busca por alternativas que melhorem a produtividade agrícola de forma sustentável. Neste sentido, uma das tecnologias que vem sendo usada com sucesso no Brasil é o uso de inoculantes vegetais baseados em microrganismos com potencial de biocontrole, biofertilizante e bioestimulante. Apesar de haver extensas coleções de microrganismos com estas potencialidades, pouco se sabe sobre seu metabolismo e conteúdo genômico, limitando sua aplicação na agricultura na forma de bioinoculantes. Neste trabalho realizamos a seleção e sequenciamento genômico de bactérias promotoras do crescimento vegetal isoladas de vermicompostos. Montamos e anotamos *in silico* um total de 7 genomas, incluindo 5 isolados do gênero *Enterobacter*, e aprofundamos os estudos no isolado *Serratia marcescens* UENF-22GI. Através de análises filogenéticas, de pan-genoma e alinhamento genômico, fomos capazes de reconstruir a árvore filogenética dessa bactéria, analisar seu conteúdo genômico único e sua relação com cepas de próximas. Também caracterizamos a presença de uma região oriunda de transferência horizontal e relacionamos seu principal produto, um peptídeo não-ribossomal, com o modo de vida do microrganismo. No tocante aos genes de promoção do crescimento vegetal, anotamos os operons *bcs*, *pqq*, *pst*, *pig* e *spe*; além de quitinases e genes de resistência a micotoxinas. Nossos resultados foram corroborados com testes *in vitro* de solubilização de P e Zn, produção de auxina, biocontrole de fungos fitopatogênicos e promoção do crescimento de plântulas de milho. Dessa forma, reportamos neste trabalho os genes responsáveis por diversos processos biológicos envolvidos na promoção do crescimento vegetal e os correlacionamos com evidências experimentais, suportando o uso de *Serratia marcescens* UENF-22GI em formulações de bioinoculantes.

Palavras chave: PGPR. Bioinoculante. Genômica.

ABSTRACT

Continuous population growth creates an increasing demand for food, making imperative the search for new alternatives to improve agriculture productivity in a sustainable manner. In this direction, the use of microbial inoculants for plants with biocontrol, biofertilizers and biostimulants properties constitute a technology that has been successfully used in Brazil. Despite the extensive collections of microorganisms bearing these abilities, little is known about their genomic content, limiting their application in agriculture in the form of bioinoculants. In this work we selected and sequenced plant growth promoting bacteria isolated from vermicompost. We assembled and annotated *in silico* a total of 7 genomes, including 5 isolates of *Enterobacter* genera, and further explored the strain *Serratia marcescens* UENF-22GI in more detail. Through phylogenetic reconstructions, pan-genome analysis and genomic alignments, we were able to study the phylogeny of this bacterium, access its unique gene content and relationship with closely related strains. We also characterized the presence of a horizontally-transferred region and correlate its main product, a non-ribosomal peptide, with the organism lifestyle. Regarding plant growth-promoting genes, we annotated the operons *bcs*, *pqq*, *pst*, *pig* e *spe*; chitinases and resistance to mycotoxins. Our results were corroborated by *in vitro* tests of P and Zn solubilization, auxin production, phytopathogenic fungi biocontrol and growth promotion of maize seedlings. Hence, we report the genes that are likely responsible for various biological processes related with plant growth-promotion and correlated these results with experimental evidence supporting the biotechnological application of *Serratia marcescens* UENF-22GI in bioinoculants formulations.

Keywords: PGPR. Bioinoculants. Genomics.

1. INTRODUÇÃO

1.1 Panorama agrícola internacional

A partir da década de 1950 a agricultura passou a progredir no sentido industrial e globalizado, em virtude do modelo produtivo em larga-escala advindo da revolução verde, esse concentra-se na monocultura com intenso uso de maquinário, combustível, água, pesticidas sintéticos e fertilizantes (McKenzie e Williams, 2015). Esse processo triplicou a produção de grãos, ocasionando a redução dos preços (Pingali, 2012). Além disso, durante a segunda metade do século 20, a área irrigada cresceu 97 %, na mesma medida o uso de fertilizantes nitrogenados aumentou em 500 %, e de pesticidas em 850 % (Figura 1) (Liu *et al.*, 2015).

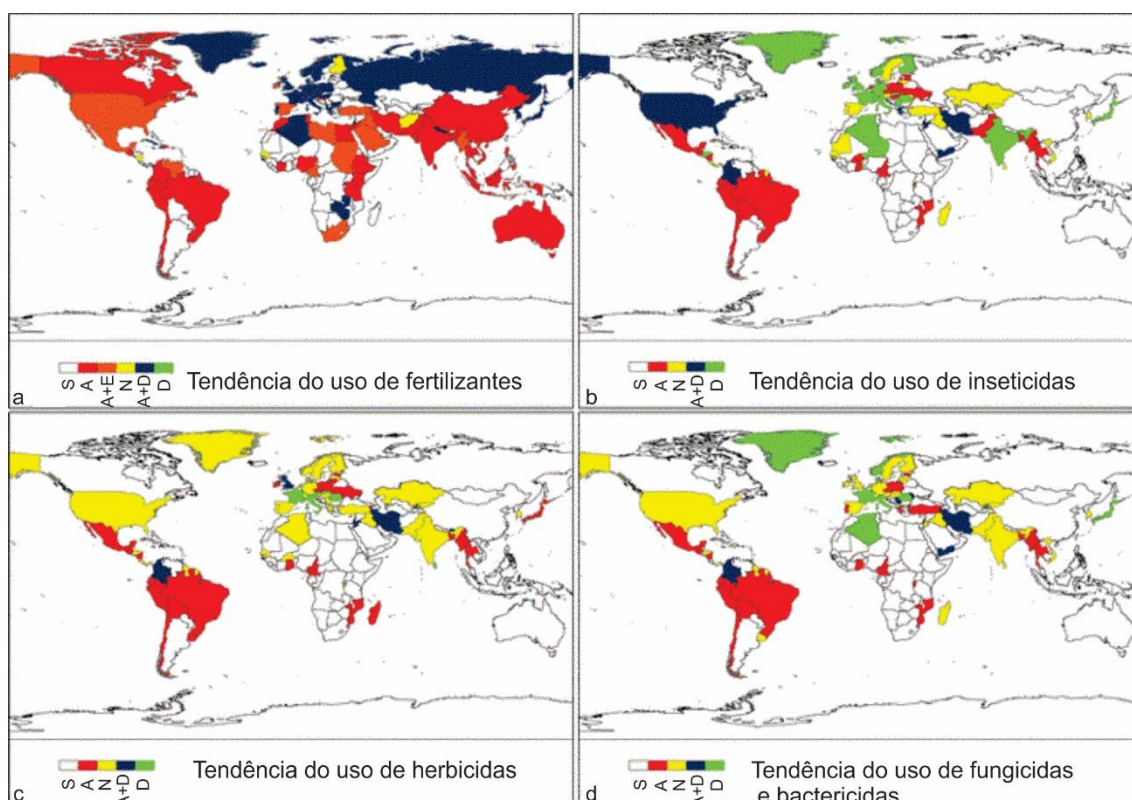


Figura 1. Mapa de calor mostrando a tendência de consumo de fertilizantes (a), inseticidas (b), herbicidas (c) e fungicidas e bactericidas (d) em todo o mundo, tendo como base dados de 1961 a 2010. S – sem dados; A – aumentando; A+E – aumentou e estagnou; N – neutro; A+D – aumentou e diminuiu; D – diminuiu. Adaptado de Liu *et al.*, 2015.

As projeções indicam que até 2050 a população global será 50% maior que a atual e a demanda por grãos duplicará (Pradhan *et al.*, 2015). O tamanho

populacional e o consumo per capita são considerados os dois principais fatores responsáveis pelas mudanças no ecossistema global (Tilman *et al.*, 2011); sendo assim, o aumento da produtividade agrícola é essencial para manter a estabilidade e equidade social (Tilman *et al.*, 2002), suprir a crescente pressão por alimentos com foco na responsabilidade ambiental e na saúde pública, em tempos de mudanças climáticas intensas, tornam o desafio ainda maior (Bodirsky *et al.*, 2015).

Uma das alternativas mais viáveis para a implantação de um modelo de produção agrícola sustentável é a utilização de microrganismos (Dutta *et al.*, 2015). Microrganismos benéficos podem promover o crescimento das plantas através de efeitos biofertilizantes (e.g. fixação biológica de nitrogênio, solubilização de fosfato e zinco, biodisponibilização de enxofre), bioestimulantes (e.g. produção de fitormônios) e de biocontrole (proteção contra microrganismos fitopatogênicos) (Leveau, 2007; Berg *et al.*, 2013). Assim, bioinoculantes são microrganismos conhecidos, isolados ou em consórcio, capazes de promover o crescimento vegetal que são adicionados diretamente no solo ou na semente (Owen *et al.*, 2015).

Avanços relevantes ocorreram em decorrência do acúmulo de conhecimento sobre ecologia e diversidade microbiana, mecanismos de interação microrganismo-planta e pela habilidade de manter, manipular e modificar populações benéficas sob condições de campo (Berg *et al.*, 2014). Os resultados econômicos da fixação biológica de nitrogênio (FBN) no Brasil são excelentes, na safra de 2012/2013 cerca de 25 milhões de hectares de soja foram plantados, e o uso de bioinoculantes gerou uma economia de 15 bilhões de dólares em fertilizantes nitrogenados (Hungria e Mendes, 2015). A prática foi também adotada com sucesso na Argentina, Bolívia e Uruguai (Izaguirre-Mayoral *et al.*, 2007), ilustrando a versatilidade e aplicabilidade dos insumos microbiológicos em variados tipos de solos, microclimas e diferentes perfis de microbiota residente na rizosfera (Dobbelaere *et al.*, 2003).

1.2 A interação planta-microrganismo

As plantas interagem com uma complexa comunidade de microrganismos, denominada microbioma vegetal, que possui caráter variável de acordo com o micro-habitat (Imam, 2016). No que tange o interesse

agronômico, um dos sítios mais relevantes de microbioma vegetal é a rizosfera, área do solo influenciada pelas raízes e dotada de população microbiana singular se comparada ao restante do solo (Hartmann *et al.*, 2009), especialmente em decorrência da atividade exercida pelo sistema radicular das plantas (Fibach-Paldi *et al.*, 2012). De forma geral denomina-se rizobactérias qualquer bactéria presente na rizosfera (Nanjareddy *et al.*, 2014), esses microrganismos desempenham papéis relevantes na fisiologia dos vegetais (Hayashi *et al.*, 2014).

A rizosfera representa a interface solo-planta, enquanto a filosfera consitui-se da região de contato entre o ar e o vegetal, sendo este micro-habitat também de especial interesse devido à grande exposição a patógenos aéreos (Lindow e Brandl, 2003; Vorholt, 2012). Além da rizosfera e filosfera, os vegetais possuem outros micro-ambientes, como por exemplo, antosfera (flor), espermosfera (sementes) e carposfera (frutos) (Nelson, 2004). Além destes, diferenciam-se a endosfera (tecidos internos) e a ectosfera (superfícies externas) (Berg *et al.*, 2014).

Algumas rizobactérias são capazes de promover o crescimento vegetal diretamente, produzindo e secretando estimuladores químicos (e.g. fitormônios) (Spaepen *et al.*, 2009). Além disso, a participação bacteriana em outros processos pode beneficiar fisiologicamente a planta, como na fixação biológica de nitrogênio ou na solubilização de fósforo no solo (Pankiewicz *et al.*, 2015). Outras rizobactérias beneficiam o crescimento vegetal por meio de efeitos indiretos, principalmente relacionados com a redução dos danos causados por patógenos (Verbon e Liberman, 2016). A essas bactérias capazes de modular positivamente o crescimento vegetal designou-se o termo PGPR, do inglês *plant growth-promoting rhizobacteria*.

A microbiota vegetal é variável nos diferentes tecidos; algumas bactérias colonizam o interior dos vegetais sem causar danos substanciais ao seu hospedeiro (endofíticos) (Santoyo *et al.*, 2016). Há uma grande diversidade de rizobactérias endofíticas culturáveis, sendo a maioria de gêneros comumente encontrados no solo, como *Enterobacter*, *Pseudomonas*, *Burkholderia*, *Bacillus* e *Azospirillum* (Lodewyckx *et al.*, 2002; Mastretta *et al.*, 2006). As rizobactérias são capazes de produzir fitormônios como ácido indol-acético (AIA) (Khan *et al.*, 2016) e fixar nitrogênio (diazotrofia) (Prakamhang *et al.*, 2009). Além disso, os

endófitos conferem às plantas proteção contra patógenos por meio da indução de mecanismos de defesa, produção de substâncias antagonistas ou exercendo competição por sítios de colonização e nutrientes (Weyens *et al.*, 2009b; Compant *et al.*, 2010).

Em solos desgastados ou contaminados, as bactérias endofíticas podem auxiliar a planta hospedeira a sobrepujar os efeitos fitotóxicos causados pela contaminação ambiental (Taghavi *et al.*, 2005). Dessa forma, essa interação biológica constitui-se de uma valiosa ferramenta no manejo de resíduos e no controle de poluição (Weyens *et al.*, 2009a).

1.3 Mecanismos de promoção direta de crescimento

Mecanismos diretos de indução do crescimento vegetal são aqueles envolvidos na facilitação da aquisição de nutrientes essenciais (e.g. P, Zn ou N), ou a modulação hormonal direta da planta (e.g. produção de fitormônios) (Gupta *et al.*, 2015).

1.3.1 Solubilização de fosfato e zinco

Considerado elemento essencial na biologia vegetal, o fósforo (P) participa do metabolismo energético e da biossíntese de ácidos nucleicos e biomembranas (Goswami *et al.*, 2016). Tanta relevância caracteriza esse elemento como macro-nutriente limitante em diversas culturas e tipos de solo (Mehra *et al.*, 2017). Para contornar essa limitação, produtores se valem do uso de fertilizantes fosfatados, obtidos a partir de: rocha fosfatada, ácido sulfúrico, e potássio em alta temperatura (Cordell *et al.*, 2009). Entretanto, estimativas apontam que as reservas de rochas fosfatadas podem se exaurir no próximo século (Cordell e White, 2015). Ademais, grande parte do P solúvel aplicado é rapidamente precipitado em compostos de baixa solubilidade como complexos cálcio (Ca), ferro (Fe) ou alumínio (Al), formas indisponíveis para o metabolismo vegetal (Randriamanantsoa *et al.*, 2013) e prejudiciais ao solo e aos recursos hídricos (Sharma *et al.*, 2013). A fração de P precipitada é lixiviada, se acumulando em ecossistemas lacustres, acarretando eutrofização desses ecossistemas (Schindler *et al.*, 2008).

Encontrar soluções para que a produção agrícola disponha de P suficiente para atender às demandas nutricionais da sociedade é um desafio global (MacDonald *et al.*, 2011). Nesse sentido, bactérias do solo são capazes

de modular a quantidade de P disponível para uso vegetal, solubilizando o P inorgânico (Arcand e Schneider, 2006) e hidrolisando fosfatos oriundos de fontes orgânicas (Shropshire e Bordenstein, 2016).

Necessário em quantidades pequenas pelas plantas ($20 \text{ mg} \cdot \text{Kg}^{-1}$), o zinco (Zn) é considerado um micro-nutriente para os vegetais (Barker e Pilbeam, 2015). Apesar disso, a deficiência de Zn é a mais comum dentre os micro-nutrientes para diversas culturas (Hafeez *et al.*, 2013). Na fisiologia vegetal o Zn está envolvido na expressão de genes de tolerância a estresse ambiental, um exemplo é a proteína *zinc finger* Zat12 foi descrita como tendo sua expressão aumentada em diferentes situações de estresse em *Arabidopsis thaliana*, como danos físicos, infecções, presença de metais pesados (MPs), seca, entre outros (Davletova *et al.*, 2005). Também em *A. thaliana* a proteína *zinc finger* BIRD participa da distribuição da auxina ao longo do eixo longitudinal (Long *et al.*, 2015).

1.3.2 A solubilização de fosfato inorgânico

No solo, a maior fração de P encontra-se insolúvel em compostos inorgânicos, pouco acessíveis às plantas (An e Moe, 2016), panorama que é agravado em solos degradados (DeLonge *et al.*, 2013). Na vizinhança das raízes, o fosfato se move principalmente através de difusão em solução a partir da qual é absorvida pelas plantas (Santner *et al.*, 2015). Alguns microrganismos são capazes de converter essas formas minerais de P em compostos iônicos solúveis, uma etapa essencial no ciclo do P e que garante a biodisponibilidade desse elemento no solo (Sharma *et al.*, 2013). A principal rota de solubilização é a produção e secreção ânions orgânicos, como ácidos capazes de reduzir o pH do meio nas adjacências das células microbianas, aumentando a solubilidade do fosfato inorgânico (Zaidi *et al.*, 2009). Alguns genes responsáveis por participar da solubilização de P já foram elucidados, como glicose desidrogenase (*gdh*) e pirronoquinolina-quinona (*pqq*) em *Gluconacetobacter diazotrophicus* (Intorne *et al.*, 2009), e a expressão destes parece ser afetada por condições ambientais; por exemplo, o aumento na relação C/P acarreta aumento na síntese de ácidos orgânicos (Wang *et al.*, 2004).

1.3.3 A solubilização de fosfato orgânico

O fosfato orgânico constitui-se de uma fração relevante do total de P no

solo e desempenha um importante papel no ciclo do fósforo nos solos agrícolas (Tarafdar e Gharu, 2006). A forma predominante de fosfato orgânico é o fitato ($C_6H_{18}O_{24}P_6$), que constitui cerca de 60% de todo fosfato orgânico presente no solo (Gerke, 2015). Contudo, o fitato encontra-se fortemente ligado à fase sólida do solo (Celi e Barberis, 2007), sendo, portanto, de difícil assimilação pelas plantas (Richardson *et al.*, 2000). Além do fitato, os demais compostos orgânicos fosfatados presentes no solo apresentam, tipicamente, alta massa molecular, e estão associados às substâncias húmicas (SHs) (Gerke, 2010). O fitato também atua como molécula de armazenamento de P em plantas, encontrando-se em concentração variada nos diferentes tecidos vegetais. Vários estudos indicam que o fitato é abundante e particularmente importante em sementes (Raboy, 2009; Alkarawi e Zotz, 2014).

Na maioria dos solos, a absorção de P pelas raízes se dá nas formas $H_2PO_4^-$ e HPO_4^{2-} (Raghothama e Karthikeyan, 2005). Enzimas fosfatases atuam na clivagem do ânion fosfato em uma etapa preliminar e necessária à absorção (Zeller *et al.*, 2015). No caso do fitato, as fitases são as enzimas responsáveis por hidrolisar o anel inositol; são conhecidos três tipos de fitases: 3-fitases (EC 3.1.3.8), 4-/6-fitases (EC 3.1.3.26) e 5-fitases (EC 3.1.3.72); uma classificação que se refere à posição inicial de ação no anel inositol na qual a desfosforilação se inicia (Zeller *et al.*, 2015). Outra classificação baseada na estrutura proteica e propriedades catalíticas alberga 4 famílias distintas de fitases: fosfatases histidínicas ácidas, fitases beta-hélice, fosfatases ácidas roxas e fitases tirosina fosfatases-like (Lei *et al.*, 2013).

Apesar de diversas plantas produtoras de fitase serem conhecidas (Brinch-Pedersen *et al.*, 2002), estudos apontam que a atividade dessas enzimas nas raízes é insuficiente para a efetiva utilização do fitato como fonte de P (Richardson *et al.*, 2000). As fitases também já foram caracterizadas em fungos, leveduras e bactérias (Lei *et al.*, 2007). Trabalhos de inserção genética têm sido realizados em *Arabidopsis thaliana* gerando plantas transgênicas, com o intuito de obter um vegetal produtor de fitase microbiana de alto desempenho (Mudge *et al.*, 2003; Belgaroui *et al.*, 2016).

1.3.4 Fixação de nitrogênio

A FBN foi o primeiro mecanismo de promoção de crescimento vegetal mediado por bactérias a ser descrito (Elmerich, 2015). As bactérias fixadoras de nitrogênio (diazotróficas) podem se estabelecer no interior dos tecidos vegetais, regiões externas do vegetal ou no solo, em vida livre ou na rizosfera. Tendo em vista o fato do nitrogênio ser um dos principais macro-nutrientes limitantes do crescimento vegetal, a FBN apresenta grande relevância econômica. No Brasil, o exemplo mais emblemático é a cultura da soja, que com o uso de bioinoculantes baseados em *Bradyrhizobium* gerou significativa economia com fertilizantes nitrogenados industrializados (Souza *et al.*, 2015).

Outro aspecto importante é o baixo custo ambiental associado a FBN, uma vez que os fertilizantes nitrogenados são produzidos pela síntese de Haber-Bosch, um processo que demanda alta temperatura e pressão para reduzir o dinitrogênio atmosférico em amônia, empregando recursos não renováveis (Vicente e Dean, 2017) e aumentando a emissão de carbono. Além disso, o uso de fertilizantes nitrogenados pode ainda aumentar a produção de óxido nitroso (Ravishankara *et al.*, 2009) e lixiviação de N para corpos de água, contribuindo para eutrofização (Bodirsky *et al.*, 2014).

1.3.5 Produção de hormônios vegetais

Sabe-se que algumas espécies de bactérias são capazes de produzir hormônios vegetais, como AIA, giberelinas, citocininas e ácido abscísico (Bashan *et al.*, 2014). A grande maioria dos trabalhos na área, evidencia a capacidade de PGPRs sintetizarem AIA (Baca e Elmerich, 2007), que frequentemente é capaz de modular o desenvolvimento e crescimento vegetal (Kochar *et al.*, 2011; Spaepen e Vanderleyden, 2011), participando de processos chave como a divisão celular, alongamento, desenvolvimento do fruto e senescência (Grossmann, 2010; Premachandra *et al.*, 2016). Além disso, a AIA bacteriana também atua na rizosfera através da estimulação radicular causando aumento das raízes laterais e primárias (Patten e Glick, 2002). A síntese de AIA também é notória em microrganismos fitopatogênicos como *Agrobacterium* spp. e *Pseudomonas savastanoi* pv. *savastanoi*, causando tumores e galhas respectivamente (Jameson, 2000; Mole *et al.*, 2007).

A síntese de AIA dá-se pela conversão do Triptofano por 4 vias

metabólicas principais (Figura 2): a via do ácido indol pirúvico (AIP), a via da indol-3-acetamida (IAM), a via da triptamina (TAM) e a via do indol-3-acetonitrila (IAN) (Spaepen *et al.*, 2007b), sendo a via do AIP a melhor caracterizada; nesta via, o Triptofano é transaminado a AIP por aminotransferases aromáticas, subsequentemente descarboxilado a indol acetaldéido e, finalmente, oxidado a AIA (Jijon-Moreno *et al.*, 2015). Os genes *hisC1* e *hisC2* (Castro-Guerrero *et al.*, 2012) codificam as aminotransferases aromáticas 1 e 2 respectivamente (Pedraza *et al.*, 2004). Nessa via, a enzima considerada central é a indol piruvato descarboxilase, codificada pelo gene *ipdC* (Romasi e Lee, 2013).

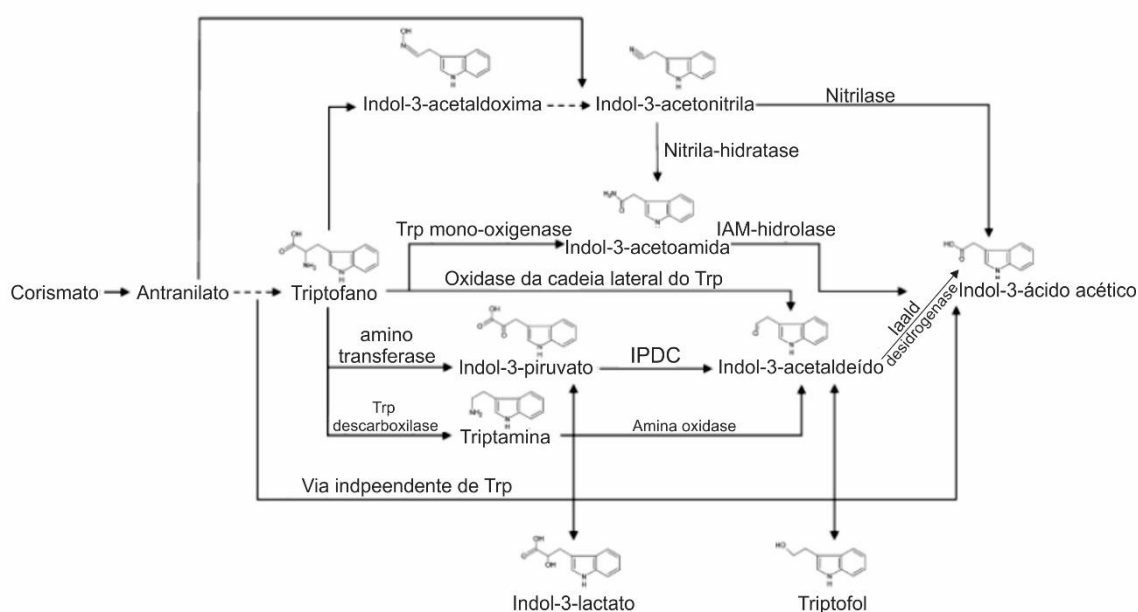


Figura 2. Principais vias de síntese de auxina descritas em bactérias. Rotas metabólicas descritas para a síntese de auxina em bactérias, destaca-se a relevância do Triptofano como precursor da biossíntese. Adaptado de Spaepen *et al.*, (2007a).

A via IAM consiste de duas etapas realizadas pelas enzimas Trp-monooxigenase-2 (*Iaam*) e indol-3-acetamida-hidrolase (*IaaH*) (Duca *et al.*, 2014). Ambos os genes são bem documentados em bactérias patogênicas e PGPRs (Spaepen *et al.*, 2007a). Em *A. tumefaciens*, estes genes estão presentes no plasmídeo pTi (*tumor inducing plasmid*), juntamente com o gene *ipt* envolvido na produção de citocinina (Ueda *et al.*, 2012). *A. tumefaciens* é capaz de infectar dicotiledôneas transferindo material genético para o interior de células vegetais, causando um desbalanço hormonal que resulta na formação de galhas (Păcurar *et al.*, 2011). Interessantemente, em *P. savastanoi pv. savastanoi*, os genes *Iaam* e *IaaH* também são plasmidiais (Kahlon, 2016) e a eliminação dos

plasmídios pIAA1 elimina a virulência do microrganismo (Mount, 2012).

A via da TAM foi sugerida para *Bacillus cereus* e *Azospirillum spp.*, através da identificação de uma Triptofano-descarboxilase (Perley e Stowe, 1966). Porém os genes envolvidos nesse processo ainda permanecem elusivos (Shao *et al.*, 2015). Outra via que ainda carece de confirmação em bactérias é a IAN, a qual foi postulada após a identificação de uma nitrilase capaz de converter IAN a AIA em *A. thaliana* (Bartling *et al.*, 1992). Em *Alcaligenes faecalis*, nitrilases com especificidade para IAN foram detectadas (Kobayashi *et al.*, 1995), embora os genes envolvidos permaneçam desconhecidos.

1.4 Mecanismos de promoção indireta de crescimento

Além dos efeitos benéficos diretos, as bactérias PGPR também são capazes de atuar de forma indireta na promoção do crescimento vegetal (Son *et al.*, 2014). Os mecanismos indiretos são aqueles que atenuam situações de estresse para as plantas como biocontrole de fitopatógenos, indução a resposta imune sistêmica da planta (Vejan *et al.*, 2016), resistência a estresse osmótico e salinidade, remoção de MPs e conversão de substratos (e.g quitina) (Datta *et al.*, 2015).

1.4.1 Biocontrole

Bactérias exercem o biocontrole através de mecanismos distintos que incluem a secreção de sideróforos, moléculas quelantes de ferro (Ahmed e Holmström, 2014), a produção de enzimas, antibióticos e uma ampla gama de moléculas capazes de inibir o crescimento de fitopatógenos (Hakeem e Akhtar, 2016), ou ainda induzindo a resposta imune sistêmica do vegetal (Venturi e Keel, 2016) (Figura 3). O ferro é considerado um composto limitante no solo e essencial para a vida. O íon Fe^{3+} é constituinte do citocromo e de proteínas que possuem o grupamento heme, além de atuar como co-fator em um vasto grupo de enzimas (Shaikh e Sayyed, 2015). Sideróforos bacterianos podem ser divididos em quatro classes principais: catecolatos, hidroximatos, carboxilatos e pioverdinas (Crowley, 2006). Assim, a produção de moléculas quelantes de ferro por agentes de biocontrole em quantidade suficiente pode limitar o Fe^{3+} disponível para o patógeno, limitando seu crescimento e favorecendo a resistência da planta (Bakker *et al.*, 2007).

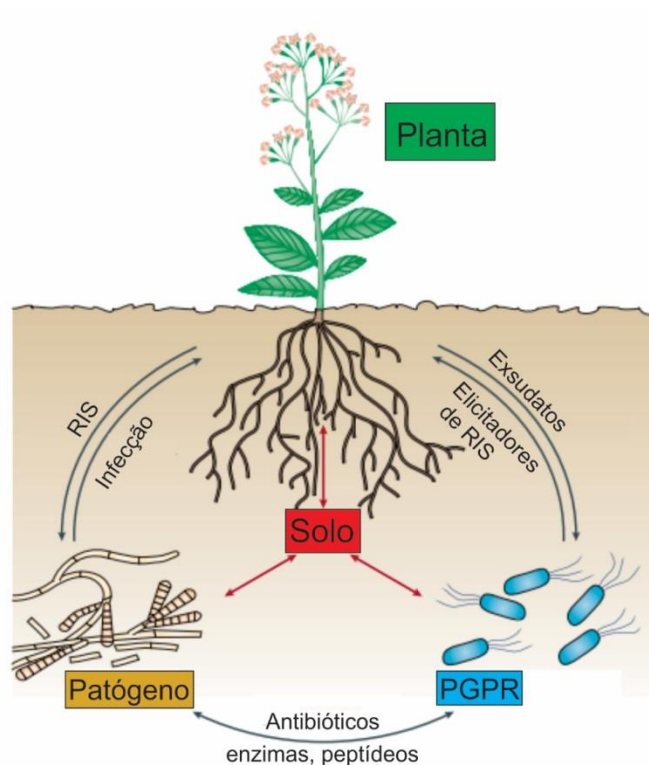


Figura 3. Mecanismos principais envolvidos no biocontrole de patógenos e indução da resposta imune sistêmica (RIS) do vegetal. Adaptado de Haas e Défago (2005).

A antibiose é o mecanismo melhor caracterizado de biocontrole e refere-se à inibição do fitopatógeno por produtos do metabolismo secundário do agente efetor (Anwer, 2017). Esses produtos incluem compostos voláteis, tóxicos ou antibióticos que são deletérios para as atividades metabólicas de outros microrganismos no meio (Saraf *et al.*, 2014). Dentre tais compostos, destacam-se: floroglucinol (Combes-Meynet *et al.*, 2011), fenazinas e pirrolnitrinas (Wang *et al.*, 2015), lipopeptídeos cíclicos (Ongena *et al.*, 2010) e enzimas líticas e quitinases (Gupta *et al.*, 2015). Isolados pertencentes aos gêneros *Bacillus*, *Burkholderia*, *Enterobacter*, *Herbaspirillum*, *Ochrobactrum*, *Pseudomonas*, *Serratia* e *Stenotrophomonas* já foram descritos como bactérias capazes de atuar inibindo o crescimento de fitopatógenos no solo (Tariq *et al.*, 2017).

Outra estratégia empregada pelas bactérias antagonistas é a produção de antibióticos, que geralmente são compostos de baixo peso molecular que inibem atividades metabólicas essenciais em seus alvos (Haas e Défago, 2005). Alguns dos antibióticos mais comuns produzidos por PGPRs são: kanosamina, 2, 4-diacetylphloroglucinol (2, 4-DAPG), oligomicina A, gramicidina, xantobaccina, fenazina, pirrolnitrina e zwittermicina A (Ulloa-Ogaz *et al.*, 2015). Outros antibióticos são tipicamente produzidos por determinados microrganismos, como

por exemplo a prodigiosina, comum em diversos isolados de *S. marcescens*. Trabalhos relataram que a prodigiosina possui atividade antibacteriana (Gulani *et al.*, 2012), antifúngica (Madison *et al.*, 2017), antiparasítica e larvicida (Suryawanshi *et al.*, 2015a; Suryawanshi *et al.*, 2015b). Tendo em vista potenciais usos terapêuticos da prodigiosina, um trabalho recente avaliou o potencial toxigênico desse antibiótico frente a *Caenorhabditis elegans* e concluiu que o composto não participa do processo de infecção ou morte do nematódeo, não exibindo, portanto, toxicidade à eucariotos complexos (Seah *et al.*, 2016).

1.4.2 Quitinases

A desconstrução enzimática da biomassa de carboidratos é vital para que ocorra o ciclo do carbono. A quitina $(C_8H_{13}O_5N)_n$ é um polissacarídeo complexo de difícil biodegradação, composto por unidades de N-acetilglicosamina (NAG) conectadas por ligações glicosídicas β -1,4 que são quebradas pelas enzimas quitinolíticas (Hamid *et al.*, 2013). Abundantemente distribuída no planeta, constitui-se do segundo polissacarídeo mais abundante na terra. Serve como elemento estrutural na constituição de muitos organismos como fungos, algas, insetos e crustáceos (Beier e Bertilsson, 2013). A quitina pode ser categorizada em três tipos: α -quitina, β -quitina e γ -quitina. A α -quitina é mais abundante e forma estruturas cristalinas e compactas como a encontrada nas carapaças de insetos e na parede celular de fungos, ao passo que β -quitina apresenta ligações mais fracas entre as moléculas, formando assim compostos menos estáveis (Deeba *et al.*, 2016).

Existem dois principais tipos de quitinases, exo-quitinases e endo-quitinases, que diferem em seu modo de ação. Endo-quitinases são capazes de quebrar as ligações glicosídicas em pontos internos aleatórios da macromolécula, gerando unidades multiméricas menores compostas de NAG (e.g. quitotriose). As exo-quitinases dividem-se ainda em duas subcategorias que atacam as terminações da quitina: quitobiosidases e β -(1,4)-N-acetilglucosaminidases (Bibra *et al.*, 2017). Com base na sequência de aminoácidos, as enzimas quitinolíticas são ainda classificadas em 3 famílias principais de glicosil-hidrolases (18, 19 e 20), sendo as quitinases pertencentes somente às famílias 18 e 19 (Stoykov *et al.*, 2015). As quitinases bacterianas melhor caracterizadas pertencem à família glicosil-hidrolase 18, classificadas nas

subfamílias A, B, C e D, as quais são determinadas com base nas diferenças de domínios protéicos conservados (Vaikuntapu *et al.*, 2016).

1.4.3 Resistência a metais pesados

A contaminação do solo com MPs se tornou um sério problema ambiental em todo o mundo devido às atividades geológicas e antropogênicas (Chibuike e Obiora, 2014). Os principais MPs são, principalmente, alumínio (Al), zinco (Zn), manganês (Mn), cromo (Cr), cobre (Cu), cádmio (Cd), chumbo (Pb), mercúrio (Hg) e arsênico (As). Esses compostos não são degradáveis e, portanto, persistem indefinidamente no ecossistema (Emamverdian *et al.*, 2015).

Uma alternativa aos métodos físicos e químicos de remoção desses poluentes é o uso de microrganismos benéficos, que podem atuar de diferentes formas na biorremediação (e.g. composição de biofiltros, compostagem, e fitorremediação) (Ullah *et al.*, 2015). Esta última é uma técnica biológica que utiliza plantas para remover poluentes do solo, o que pode ser potencializado com o uso de PGPRs (Chang *et al.*, 2014).

O pré-requisito inicial para que uma bactéria possa participar da fitorremediação é sobreviver à presença de MPs, processo que pode envolver diferentes estratégias, por exemplo: evitar o influxo de MPs na célula, transporte ativos transmembrana, volatilização, detoxificação e biocomplexação (Pavel *et al.*, 2013). A biocomplexação é um mecanismo promissor que envolve o sequestro de MPs pelos exo-polissacarídeos constituintes do biofilme, protegendo a comunidade bacteriana e a planta dos efeitos nocivos dos MPs (Hou *et al.*, 2013).

Muitas espécies de plantas já foram caracterizadas como capazes de acumular MPs eficientemente, entretanto, muitos desses vegetais não produzem biomassa suficiente que os tornem capazes de biorremediar poluentes em condições de campo. Nesses casos PGPRs podem promover o crescimento vegetal, incrementando a biomassa e, majorando a eficiência da fitorremediação (Ullah *et al.*, 2015). Dentre os organismos já descritos como capazes de atuar nesse sentido estão: *Enterobacter* sp., *Pseudomonas tolasii* e *Microbacterium arborescens* (Ashraf *et al.*, 2017).

1.5 Biotecnologia de inoculantes

Apesar de produtos microbianos serem utilizados comercialmente na

agricultura por mais de 120 anos (Nobbe e Hiltner, 1896; Deaker *et al.*, 2004), estes têm recebido cada vez mais atenção. Atualmente existem mais de 150 cepas microbianas registradas como produtos agrícolas nos EUA, sendo que estimativas recentes dão conta que o segmento de biofertilizantes movimentará US\$ 1.6 bilhões em 2022 (Timmusk *et al.*, 2017).

Inoculantes contendo rizóbio são produzidos no mundo inteiro, sendo a iniciativa mais emblemática na área de insumos biológicos para promoção do crescimento vegetal e incremento da produtividade agrícola pela Sociedade Americana de Microbiologia (Reid e Greene, 2013). Além disso, a inoculação com microrganismos (e.g. bactérias rizosféricas do gênero *Azotobacter*) foi amplamente utilizada na Rússia entre 1930 e 1940, mas foi abandonada por ter apresentado resultados não conclusivos (Rubenchik e Starkey, 1964). Ainda em 1930, ensaios de inoculação utilizando *Bacillus megaterium* como agente solubilizador de fosfato foram conduzidos no sudoeste Europeu, aparentemente sem sucesso (Bashan e Levanony, 1988). Entretanto, os entusiastas das tecnologias de inoculação de plantas presenciaram um período muito fértil de pesquisas ao final da década de 1970, com o uso de *Azospirillum* como agente de promoção do crescimento em plantas não-leguminosas (Döbereiner e Day, 1976; Döbereiner *et al.*, 1995), além de *Pseudomonas fluorescens* e *P. putida* em práticas de biocontrole (Kloepper e Schroth, 1981).

Um inoculante microbiano é uma formulação que contém uma ou mais estirpes (ou espécies) bacterianas em um veículo de fácil utilização, econômico, de origem orgânica ou inorgânica, ou sintetizada a partir de moléculas definidas. O inoculante é o meio pelo qual o microrganismo é transportado da fábrica até a planta (Trabelsi e Mhamdi, 2013).

A intensificação dos estudos focados na biotecnologia de inoculantes é componente estratégico para o incremento do uso de insumos biológicos no ambiente agrícola (Morrissey *et al.*, 2002). Uma análise do tema na agricultura contemporânea revela iniciativas discretas e de baixo alcance, com pouca adoção dos agricultores (Olivares, 2009). Inúmeros aspectos contribuem para este quadro e envolvem questões culturais, econômicas, agronômicas, científicas e tecnológicas. Inoculantes microbianos podem desempenhar um papel cada vez mais significativo na agricultura dos países em desenvolvimento, apesar de importantes aspectos envolvendo os procedimentos de campo que

devem ser trabalhados cuidadosamente (Timmusk *et al.*, 2017). A resposta imediata à inoculação com bactérias promotoras do crescimento vegetal varia consideravelmente dependendo da bactéria, planta, tipo de solo, densidade do inóculo e das condições do ambiente (Conn e Franco, 2004) .

Em geral, a população de bactérias declina progressivamente após sua introdução no solo (Goswami *et al.*, 2016). Este fenômeno, juntamente com o status fisiológico da bactéria no inoculante, pode impedir o estabelecimento de uma população suficientemente na rizosfera para obtenção de uma resposta desejável. Ademais, a heterogeneidade do solo é outro obstáculo (Nesme *et al.*, 2016), a bactéria inoculada muitas vezes não encontra um nicho na matriz do solo e, após introduzida, deve competir com a biota nativa adaptada e ser confrontada com predadores existentes na rizosfera (e.g. protozoários) (Li *et al.*, 2012a).

O primeiro objetivo quando consideramos a prática de inoculação com bactérias benéficas é encontrar a melhor bactéria disponível e a formulação específica a ser utilizada, o que em última instância determinará o potencial de sucesso do inoculante (Kumar *et al.*, 2018). Muitas das potencialidades de bactérias reportadas na literatura não resultaram em produtos comerciais, parcialmente em virtude de formulações inadequadas.

1.5.1 Formulação de inoculantes microbianos

Muitas etapas estão envolvidas no desenvolvimento de um fertilizante microbiano, sendo a sobrevivência e manutenção da atividade microbiológica sob condições de campo uma das etapas mais críticas. Para facilitar a introdução de um elevado número de células no solo, as preparações devem ser baseadas em carreadores e estar com a formulação correta (Sahai e Kumar, 2017). Simultaneamente, a viabilidade dessas preparações em condições de estocagem também é fundamental para garantir a comercialização (Vyas *et al.*, 2010).

A maioria das formulações para as bactérias promotoras do crescimento vegetal têm utilizado carvão ou outro material inerte como carreador, porém recentemente formulações utilizando polímeros como carreadores têm sido desenvolvidas (Arora *et al.*, 2017). Esse material imobiliza as células vivas, protegendo os microrganismos contra estresses ambientais (Pandey, 2009) e

liberando-os no solo gradualmente, porém em grande quantidade (Trivedi *et al.*, 2005).

Outros aspectos importantes dos carreadores são (Bashan, 1998): (a) Propriedades químicas e físicas: deve ser estéril ou facilmente esterilizável e as características físico-químicas as mais uniformes possíveis; devem ser de qualidade consistente, alta capacidade de retenção de água e adequado para muitos microrganismos; (b) Qualidade de manufatura: manufaturado pela indústria, permitindo a adição de nutrientes, ter pH facilmente ajustável e matéria-prima acessível; (c) Qualidade no manuseio na fazenda: fácil manuseio, liberação rápida e controlada da bactéria no solo e viabilidade de aplicação com máquinas agrícolas convencionais; (d) Características ambientais: biodegradável e não poluente. Baixo risco de dispersão na atmosfera e lençol freático; (e) Qualidade de estocagem: longa vida de prateleira em temperatura ambiente (1 a 2 anos).

Estudos apontam tendências promissoras no campo da tecnologia de inoculação pelo uso de inoculantes mistos (combinação de microrganismos) que interagem sinergicamente (Baldotto e Olivares, 2008). Testes de compatibilidade *in vitro* indicam que algumas combinações permitem a bactéria interagir com outras de forma sinérgica, fornecendo nutrientes e removendo produtos inibitórios (Santiago *et al.*, 2017). Estes processos podem potencializar aspectos benéficos de sua fisiologia, como a fixação biológica de nitrogênio. Combinações metabólicas onde os produtos da degradação e fermentação de um dado microrganismo poderiam ser utilizados por bactérias diazotróficas em troca de nitrogênio fixado, poderiam sustentar a atividade do inoculante (Tabassum *et al.*, 2017). De forma análoga, associações entre bactérias diazotróficas e microrganismos capazes de metabolizar carboidratos complexos podem ser mediados pela adição de resíduos da atividade agrícola (Shen *et al.*, 2016).

Trabalhos com foco nos efeitos dos inoculantes mistos nas plantas, mostram benefícios no crescimento e produtividade quando comparados com inoculação simples, são descritos aumentos da absorção de N, P e outros minerais (Amalraj *et al.*, 2015). A resposta à inoculação pode ser incrementada pela dupla introdução de *Azospirillum* e bactérias solubilizadoras de fosfato. Apesar do progresso nas pesquisas com inoculantes mistos, ainda não há

produção comercial destas combinações.

1.6 Tecnologia de vermicompostagem

A estratégia predominante para a prospecção de bactérias promotoras do crescimento vegetal usualmente é isolá-las do solo ou de tecidos vegetais (Castro *et al.*, 2014). Entretanto nosso grupo tem estudado diferentes tipos de vermicompostos, visto que estes constituem um microambiente em ótimas condições físicas e químicas, contendo ampla diversidade de microrganismos a serem prospectados (Aguiar, 2012). O termo compostagem define processos de tratamento dos resíduos orgânicos, sejam estes de origem urbana, industrial, agrícola e florestal (Hanc e Dreslova, 2016). Trata-se de um processo aeróbio controlado, desenvolvido por uma população diversificada de microrganismos, efetuada em duas fases: a primeira quando ocorrem as reações bioquímicas predominantemente termofílicas; a segunda, ou fase de maturação, quando ocorre o processo de humificação (Duan *et al.*, 2016).

As matérias-primas utilizadas para a compostagem podem ser divididos em duas classes: a dos materiais ricos em carbono (C) e a dos ricos em nitrogênio (N). Os materiais ricos em C fornecem a matéria orgânica e energia para a compostagem, enquanto que os materiais nitrogenados aceleram a compostagem (Aguiar, 2012). Durante a compostagem pode-se identificar quatro fases distintas quanto à temperatura (Trautmann e Olynciw, 2000): Fase mesofílica, com temperaturas moderadas, até cerca de 40 °C, que tem duração média de dois a cinco dias; fase termofílica, quando o material atinge sua temperatura máxima acima de 40 °C e é degradado mais rapidamente. Esta fase pode durar de poucos dias a vários meses, dependendo das características do material compostado); fase de resfriamento, que é marcada pela queda da temperatura para valores de temperatura ambiente e, finalmente; a fase da maturação, que é um período de estabilização também em temperatura ambiente, e que resulta em um composto maturado, altamente estabilizado e humificado, livre de toxicidade (Lim *et al.*, 2015; Olivares *et al.*, 2015).

A vermicompostagem se inicia após a maturação do composto com participação das minhocas, acelerando a decomposição e produzindo um composto de melhor qualidade (Huang *et al.*, 2013). A matéria orgânica é rapidamente ingerida pelos anelídeos e decomposta em um substrato rico em N,

P, Ca, Mg, S, elementos essenciais para o desenvolvimento das plantas (Song *et al.*, 2014). *Eisenia foetida* é a espécie de minhoca mais utilizada no processo de vermicompostagem, em especial por ser muito disseminada e ter uma ampla faixa de tolerância à temperatura (Aguiar *et al.*, 2012). A vermicompostagem acelera a produção da matéria orgânica humificada devido à ação de enzimas e microrganismos presentes no tubo digestivo das minhocas, além de alterar a abundância relativa dos microrganismos presentes no substrato compostado (Chen *et al.*, 2015). No interior do trato digestivo, o material orgânico é clivado, gerando nutrientes que podem ser posteriormente assimilados pelos vegetais (Mahmood *et al.*, 2013), influenciando positivamente a acumulação de nutrientes e o crescimento vegetal (Paradelo *et al.*, 2012).

1.6.1 As substâncias húmicas

As substâncias húmicas (SHs) são o principal componente orgânico do solo, sendo consideradas componentes cruciais para esse ecossistema (Kelley e Stevenson, 1995). Essas substâncias não podem ser facilmente incorporadas pelos seres vivos, por causa de suas íntimas interações com a porção mineral do solo, além de serem muito complexas quimicamente para a utilização por microrganismos (Trevisan *et al.*, 2010).

No solo, uma das características mais marcantes das SH é sua capacidade de interagir com íons metálicos, óxidos, hidróxidos, compostos minerais e orgânicos (Albers *et al.*, 2008), incluindo toxinas poluentes (Luo e Gu, 2008; Cattani *et al.*, 2009; Wang e Mulligan, 2009), formando complexos solúveis ou insolúveis. Quando em solução, as SHs proporcionam o aumento da absorção de nutrientes pelos vegetais, devido à ativação da H⁺ATPase presente na membrana plasmática das plantas (Canellas *et al.*, 2002). Estas enzimas são capazes de hidrolisar ATP, gerando energia e gradiente eletroquímico que são essenciais para sistemas secundários de translocação de íons fundamentais, absorção de nutrientes e aumento da plasticidade da parede celular para possibilitar o processo de crescimento e divisão celular (Trevisan *et al.*, 2010).

Os ácidos húmicos promovem o crescimento vegetal induzindo a germinação e a emergência de raízes laterais, o que pode aumentar o número de pontos de colonização para as bactérias (Piccolo *et al.*, 1993). Dessa forma, ácidos húmicos podem promover a população de bactérias benéficas no interior

da planta, amplificando os efeitos das PGPR (da Conceição *et al.*, 2008).

1.7 Genômica de bactérias promotoras do crescimento vegetal

Os mecanismos moleculares do reconhecimento, colonização e troca de nutrientes entre plantas e bactérias endofíticas ainda estão longe de ser totalmente conhecidos. Entretanto os recentes projetos de sequenciamento de bactérias capazes de promover o crescimento vegetal possibilitaram avanços na compreensão desses sistemas (Donoso *et al.*, 2017). Um dos primeiros genomas de bactérias endofíticas sequenciados foi o de *Azoarcus* sp., onde foi demonstrado que esta bactéria endofítica de gramíneas possui genoma de 4.376.040 pares de bases (pb), com 3.992 sequências codificantes (Krause *et al.*, 2006). Ao contrário de boa parte das bactérias patogênicas de plantas, *Azoarcus* sp. não possui os sistemas de secreção tipo III e IV, além de codificar poucas enzimas de degradação de parede celular. Também é notável a baixa prevalência de elementos de transposição, o que indica uma baixa frequência de transferência gênica, possivelmente devido à adaptação a um microambiente estável (Krause *et al.*, 2006). A bactéria apresenta 38 enzimas e transportadores envolvidos no metabolismo de nitrato (NO_3^-) (e.g nasFED, nirC, nasA, nasC e nirBD). Além disso, o genoma possui 22 genes que codificam proteínas relacionadas ao transporte de Fe e 41 genes envolvidos na montagem e regulação do pili tipo IV, que é um fator indicativo da capacidade de estabelecimento da interação com a planta (Dörr *et al.*, 1998).

Em 2009, Bertalan e colaboradores publicaram o sequenciamento do genoma de *Gluconacetobacter diazotrophicus* Pal5 (Bertalan *et al.*, 2009), uma bactéria endofítica diazotrófica que vive em associação com a cana-de-açúcar. Os resultados mostraram um cromossoma circular com 3.944.163 pb e a presença de dois plasmídeos de 38.818 e 16.610 pb. No cromossomo circular, há um total de 3,864 regiões codificadoras. Dentre os genes preditos 2,861 foram designados a uma função putativa e 1077 codificam proteínas hipotéticas. De acordo com os autores, o genoma de *G. diazotrophicus* apresenta um elevado número de elementos de transposição quando comparado com outros genomas de bactérias endofíticas disponíveis, possivelmente devido ao relaxamento da pressão seletiva.

Posteriormente, o genoma do endófito *Herbaspirillum seropedicae* cepa

SmR1 foi sequenciado e analisado por pesquisadores brasileiros (Pedrosa *et al.*, 2011). Capaz de colonizar gramíneas como arroz e cana-de-açúcar, *H. seropedicae* possui um genoma composto por um cromossoma circular de 5.513.887 pb contendo um total de 4.804 genes. A bactéria é capaz de crescer aerobicamente utilizando o NO_3^- como única fonte de nitrogênio, apesar de ser incapaz de realizar a desnitrificação anaerobicamente (Baldani *et al.*, 1986). Os genes para assimilação de nitratos estão localizados em duas regiões. A primeira destas contém genes transportadores de nitrato do tipo ABC (*nasFED*), enquanto a segunda alberga o gene *narK*, que codifica um transportador de nitrato, *nirBD* codificando a enzima nitrito redutase assimilatória. Entretanto a relação bactéria-planta envolve diversos eventos moleculares que vão além da capacidade de fixar nitrogênio, alguns dos quais ilustrados na Figura 4.

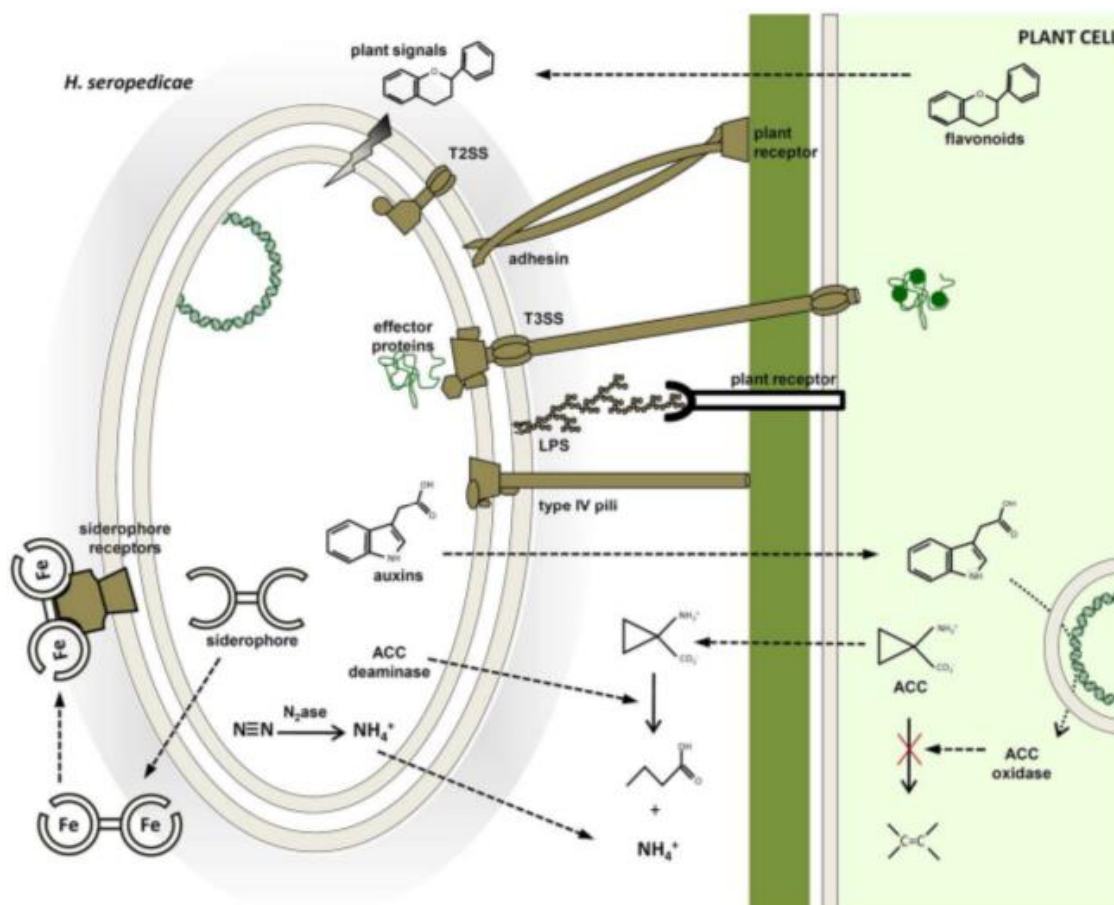


Figura 4. Mecanismos moleculares provavelmente envolvidos na colonização e promoção do crescimento de plantas identificados no genoma de *H. seropedicae* SmR1 (Pedrosa *et al.*, 2011).

1.8 Sequenciamento de nova geração

Desde o final do século 20 até 2005 o mercado de sequenciamento genômico era dominado por instrumentos que utilizavam o método de Sanger automatizado (Liu *et al.*, 2012). Esses equipamentos possibilitaram um avanço notável inclusive com a publicação da 1^o versão do genoma humano (Venter *et al.*, 2001), e a essas máquinas designou-se o termo de sequenciadores de primeira geração. Criada em seguida, a nova tecnologia foi batizada como sequenciamento de nova geração (*next-generation sequencing* - NGS). Posteriormente, com o advento de novas tecnologias o NGS passou também a ser comumente chamado como sequenciamento de segunda geração (Bleidorn, 2016).

A principal característica revolucionária da NGS foi a capacidade de produção de quantidades enormes de dados a um custo acessível. Esse novo patamar na geração de dados permitiu a execução de projetos de genômica e transcriptômica antes impossíveis e uma consequente explosão das bases de dados. A estratégia se baseia na imobilização de fragmentos do mesmo tamanho em uma fase sólida, que são amplificados gerando clusters. Esses fragmentos funcionam como fita molde, onde a DNA polimerase se liga e inicia a duplicação enzimática. A reação utiliza os nucleotídeos complexados com fluoróforos (desoxinucleotídeos trifosfatados - dNTPs) gerando fluorescência em comprimentos de onda que variam de acordo com o nucleotídeo adicionado. Essas imagens são captadas e a leitura da sobreposição das imagens revela a composição da sequência do DNA (Metzker, 2010) (Figura 5).

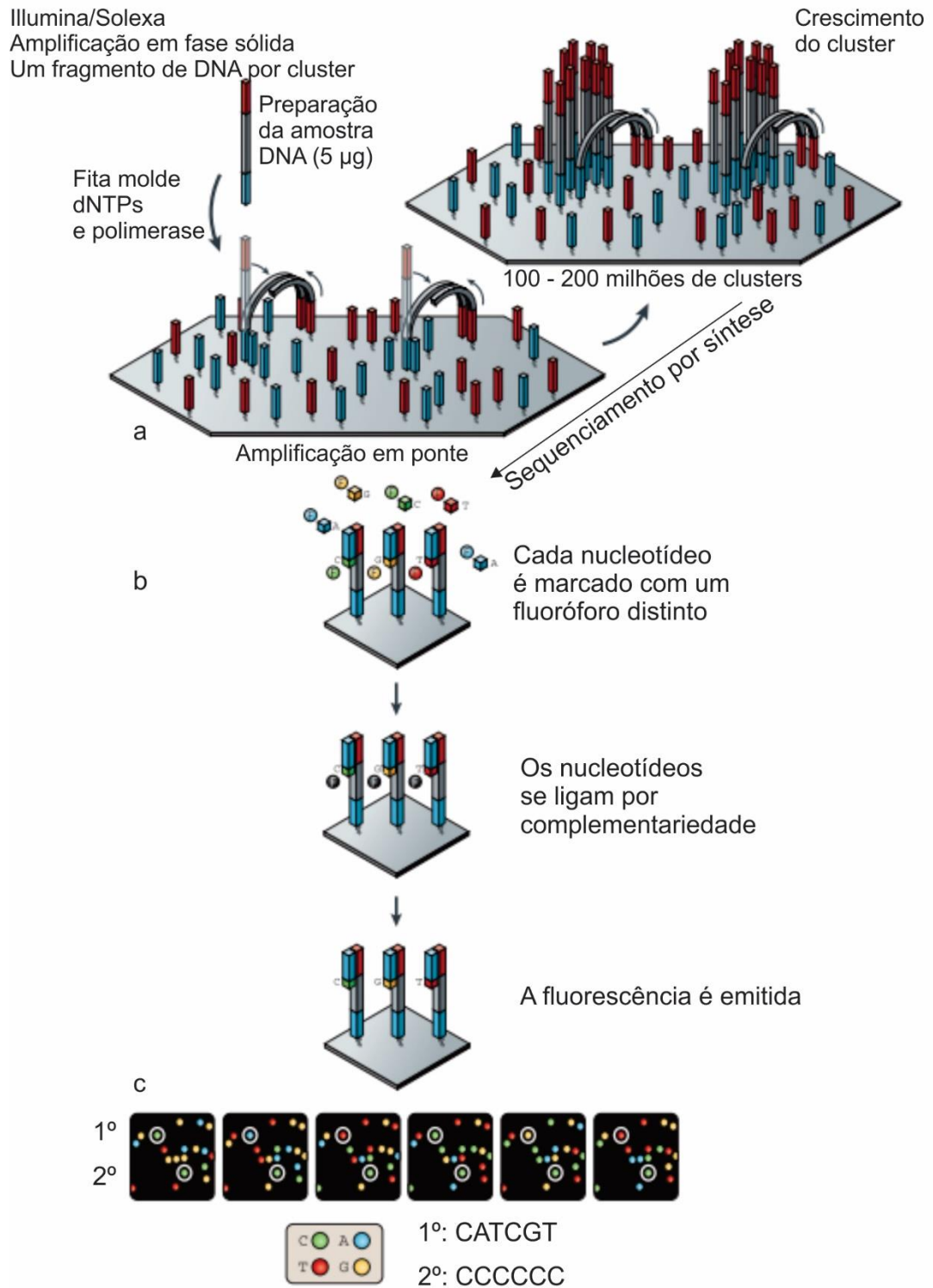


Figura 5. Representação esquemática do sequenciamento por síntese empregado pelos equipamentos Illumina. a) Preparo e amplificação das amostras. b) Sequenciamento por síntese. c) Fluorescência gerada em cores de acordo com o nucleotídeo adicionado à fita. Adaptado de Metzker (2010).

1.9 Montagem de genomas

Uma vez realizado o sequenciamento, os *reads* precisam ser montados visando reconstruir a sequência genômica. As primeiras metodologias para

reconstruir as sequências genômicas eram baseadas em sobreposição (Simpson e Pop, 2015). Os primeiros projetos de sequenciamento genômico foram realizados com a tecnologia de Sanger (Sanger, 1975), fornecendo *reads* mais longos e em menor quantidade, o que envolvia um procedimento de montagem mais intuitivo, em particular por se conhecer a priori a posição dos fragmentos. Entretanto, esta abordagem demanda intenso trabalho manual, de difícil automação e aumento de escala (Sanger *et al.*, 1977). Uma abordagem alternativa, denominada *shotgun sequencing*, foi proposta por Staden em 1979, e caracteriza-se por fragmentar o DNA, geralmente através de ações mecânicas, seguido de sequenciamento em paralelo (Staden, 1979). Os resultados fornecidos por esta técnica rapidamente tornaram clara a necessidade de computadores para detectar a sobreposição entre os fragmentos (Gingeras *et al.*, 1979).

A inovação do processo de sequenciamento *shotgun* rapidamente levou ao desenvolvimento de estudos teóricos sobre a montagem de sequências por *softwares*, os *assemblers* (Peltola *et al.*, 1984). Lander e Waterman criaram um modelo capaz de inferir o número de segmentos genômicos contíguos que podem ser reconstruídos a partir de um determinado grupo de *reads*. A função é dependente de três variáveis: o tamanho dos *reads* (i), o número total de bases presentes nos *reads* dividido pelo tamanho do genoma (i.e. cobertura) (ii) e, a sobreposição mínima entre dois *reads* (Lander e Waterman, 1988). Outro trabalho identificou dificuldades teóricas na montagem de genomas (Gallant *et al.*, 1980) e concluiu que o genoma reconstruído se configurava do menor grupo que contivesse todos os *reads* como subgrupos, embora na época este cálculo fosse computacionalmente inviável (Hunkapiller e Kaiser, 1991). Uma abordagem proposta consistia da obtenção de soluções heurísticas (aproximadas) denominadas *superstring* (Räihä e Ukkonen, 1981), o que culminou na criação do primeiro montador de genomas (Tarhio e Ukkonen, 1988). Contudo, a abordagem por *superstring* ignorava uma peculiaridade comum em genomas mais complexos: a presença de repetições, segmentos de DNA que se repetem de forma idêntica (ou quase idêntica) (Treangen e Salzberg, 2011). Portanto, a reconstrução correta de uma sequência de DNA pode não ser necessariamente a mais curta. Com base nisso, foram

desenvolvidas diferentes abordagens teóricas para a montagem de genomas nos anos que se seguiram.

1.9.1 Algoritmos *greedy*

Algoritmos *greedy* caracterizam-se por uma abordagem simples, na qual os *reads* são montados de acordo com sua qualidade de sobreposição. As sequências montadas (i.e. *contigs*), crescem pela adição de novos *reads* ou pela fusão com outros *contigs*, ao passo em aqueles que apresentem conflitos de sobreposição com os *contigs* já montados são descartados ou separados (Wajid e Serpedin, 2012). A característica peculiar desses algoritmos é fazer sempre a melhor escolha do ponto de vista local da sequência em construção (Pop, 2009).

Apesar da aparente simplicidade do método, a aproximação obtida pode apresentar qualidade relevante e alguns dos *assemblers* mais utilizados incorporaram o conceito, como por exemplo o Phrap (Green *et al.*, 1999), principal *assembler* utilizado no projeto genoma humano (Lander *et al.*, 2001) e o TIGR (Sutton *et al.*, 1995), utilizado na montagem do primeiro genoma de um organismo vivo, a bactéria *Haemophilus influenza* (Fleischmann *et al.*, 1995). Não obstante, mesmo com suas inquestionáveis contribuições, os algoritmos *greedy* possuem limitações que estimularam o advento de novos métodos de montagem de genomas.

1.9.2 Overlap layout consensus (OLC)

Inicialmente os *assemblers* consistiam basicamente de alinhadores de sequências. Contudo, em 1995, Kececioğlu e Myers criaram o que pode-se considerar a primeira representação em grafos na montagem de sequências – o grafo de sobreposição (Kececioğlu e Myers, 1995). Em um grafo de sobreposição, vértices representam *reads* e arestas representam sobreposições. Pode-se encontrar o arranjo da sequência através de um caminho ideal, percorrendo cada vértice, formando um ciclo hamiltoniano (Gallant *et al.*, 1980). Além disso, foi proposto ainda um algoritmo capaz de reorganizar o grafo (Figura 6), reduzindo a redundância de sobreposição e incrementando a capacidade computacional (Gallant, 1983).

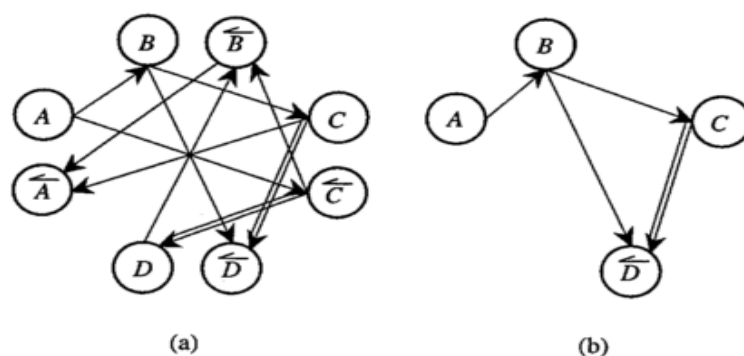


Figura 6. Grafos de sobreposição não orientados (a). Um subgrafo orientado (b). Adaptado de Kececioğlu e Myers (1995).

A implementação das novas teorias ganhou corpo com a publicação do genoma de *Drosophila melanogaster* utilizando o montador Celera (Myers *et al.*, 2000), que implementava a estratégia de montagem por OLC e a qualificação da confiabilidade dos *contigs*. Uma vez que um *contig* era representado por um grafo de sobreposição incontestável, esse classificava-se como *unitig*, e era então avaliado pelo algoritmo para aferir se a sequência representava uma região repetitiva ou única (*U-unitig*). Myers define matematicamente um *U-unitig* como um subgrafo de maior intervalo pertencente ao grafo de sobreposição de todos os fragmentos (Myers *et al.*, 2000). Destacam-se também entre os programas que se utilizam do método OLC os *assemblers* Newbler (Egholm *et al.*, 2005), ARACHNE (Batzoglou *et al.*, 2002; Jaffe *et al.*, 2003), PCAP (Huang e Yang, 2005), dentre outros.

Apesar do sucesso alcançado pela abordagem OLC, esta estratégia não conseguiu lidar com a vasta quantidade de *reads* produzidos pelos instrumentos de sequenciamento de segunda geração (e.g. Illumina HiSeq), em particular porque a busca de sobreposição entre *reads* se tornou computacionalmente impraticável, dada a ordem de centenas de milhões de *reads* curtos que temos atualmente (Miller *et al.*, 2010).

1.9.3 Grafos de Brujin

O arcabouço teórico dos grafos de Brujin data do final dos anos 1980, quando Pevzner estudou a reconstrução de sequências genômicas utilizando apenas *k*-mers (segmentos de tamanho fixo da sequência) (Pevzner, 1989). Este

trabalho teórico e seu aperfeiçoamento subsequente (Idury e Waterman, 1995; Pevzner *et al.*, 2001) levaram à criação dos assemblers de genomas baseados em grafos de Brujin. Nessa técnica, cada *read* é quebrado em uma sequência de k-mers sobrepostos e os k-mers são adicionados ao grafo como vértices. K-mers originários de posições adjacentes são vértices conectados por arestas. Figura 7 (Simpson e Pop, 2015). O problema do assembly então pode ser formulado como: descobrir o caminho através do grafo que visita cada aresta uma única vez – um circuito Euleriano (Compeau *et al.*, 2011; Zimin *et al.*, 2013).

Na prática, erros de sequenciamento ou nas amostras embaralham o grafo, de forma que um caminho Euleriano completo por todo o grafo torna-se praticamente inviável (Pevzner e Tang, 2001). Mesmo nos casos em que este caminho é encontrado, é improvável que reflita perfeitamente a sequência genômica completa, principalmente devido às repetições (Kingsford *et al.*, 2010).

A abordagem por grafos de Brujin apresenta significativa vantagem computacional quando comparada com as estratégias baseadas em sobreposição, principalmente por não se basear na busca e identificação de regiões sobrepostas (Li *et al.*, 2012b). Trabalhos demonstraram que sequências genômicas bacterianas podem ser eficientemente montadas a partir de reads curtos utilizando-se métodos baseados nos grafos de Brujin (Chaisson *et al.*, 2009). Entretanto a qualidade dos assemblies era limitada pelo tamanho dos *reads* curtos (Ekblom e Wolf, 2014). Dessa forma a acentuada redução dos custos no sequenciamento foi um ponto de inflexão para a área (Shendure e Ji, 2008).

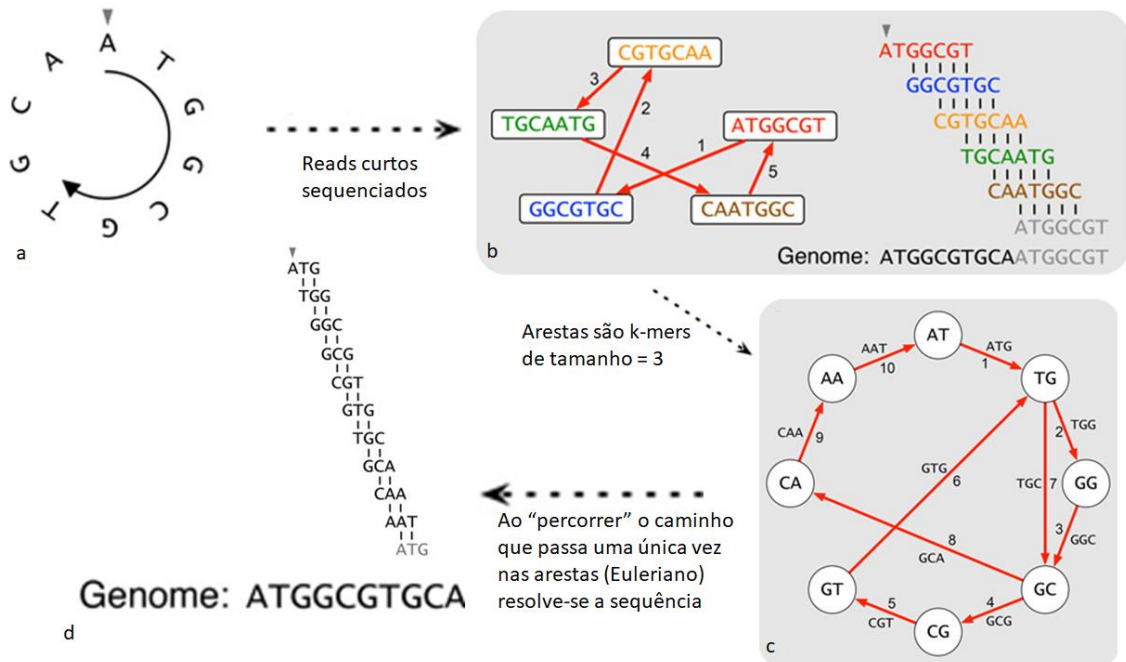


Figura 7. Representação esquemática dos Grafos de Brujin. a) DNA a ser sequenciado (10 pb). b) Reads curtos ilustrativos obtidos pelo sequenciamento (7 pb) e a esquematização da problemática da reconstrução da sequência original. c) Grafo obtido ao se utilizar um k-mer = 3. d) Ordenação dos k-mers reconstruindo a sequência original. Adaptado de Compeau *et al.*, (2001).

3. OBJETIVO GERAL

Aplicar as tecnologias de sequenciamento de segunda geração e análises de bioinformática para entender a arquitetura e composição genômica de bactérias isoladas de vermicompostos com potencial de promover o crescimento vegetal.

3.1 OBJETIVOS ESPECÍFICOS

- 1) Sequenciar e analisar os genomas de bactérias promotoras do crescimento vegetal isoladas de vermicompostos;
- 2) Determinar a espécie dos isolados selecionados com base nos dados genômicos e predizer as relações filogenéticas dos isolados com outros disponíveis publicamente;
- 3) Montagem dos genomas;
- 4) Predizer computacionalmente os genes de cada genoma, enfatizando genes relevantes à promoção do crescimento vegetal;
- 5) Correlacionar os achados *in silico* com dados *in vitro*.
- 6) Estudar comparativamente estes genomas com outros publicamente disponíveis;
- 7) Fazer um levantamento das vias metabólicas potencialmente envolvidas com o fenótipo de interesse.

4. METODOLOGIA

4.1 Isolamento das bactérias

Em um trabalho anterior (Aguiar, 2012), diferentes vermicompostos foram utilizados como fonte de prospecção de bactérias. Os isolados foram incubados em diferentes meios. O primeiro grupo de isolamento (grupo I) foi obtido a partir de diluições e plaqueamento em meio sólido de composição complexa (meio DYGS). Contribuíram para este universo de isolados amostras de vermicomposto de esterco bovino e diferentes porções do intestino de minhoca.

No segundo grupo de isolamento (grupo II), obtiveram-se isolados bacterianos com capacidade de fixar nitrogênio atmosférico, a partir de diluições seriadas e crescimento em meio JNFb semi-sólido isento de fontes de nitrogênio. Neste ensaio, procedeu-se a contagem e o subsequente isolamento de bactérias diazotróficas oriundas dos diferentes materiais vermicompostados: esterco bovino, esterco com bagaço de cana, esterco com torta de girassol, esterco bovino com torta de girassol e bagaço de cana e por último esterco bovino com torta de filtro de cana. A partir das mesmas diluições de amostras utilizadas para obtenção de isolados do grupo II obtiveram-se, adicionalmente, isolados microbianos dos diferentes vermicompostos, isolados em um meio sólido contendo ácidos húmicos como única fonte de C disponível para crescimento e as cepas assim obtidas foram incluídas no grupo III.

4.2 Seleção de bactérias enviadas para o sequenciamento

A seleção de bactérias que tiveram seus genomas sequenciados foi feita a partir da coleção do Prof. Fábio Olivares (LBCT-CBB-UENF), que é um dos coordenadores do INCT Fixação Biológica de Nitrogênio e possui vasta experiência com microrganismos promotores do crescimento vegetal. Tendo em vista os dados obtidos em trabalhos anteriores, escolhemos para o sequenciamento os isolados: 4GII, 5BGII, 6GII, 14GII, 17GIII, 21GII, 22GI, 23GI, 24GI, 25GI, 25GII, 25GIII, 26GII, 26GIII, 65, 98, 101 e 103. As bactérias, isoladas a partir dos diferentes sistemas supracitados (GI, GII e GIII), foram testadas em outro trabalho (Aguiar, 2012) quanto a capacidade de solubilizar P e Zn, degradar celulose, fixar N₂ e produzir auxina, funções de interesse biotecnológico no contexto da promoção do crescimento vegetal.

4.3 Extração do DNA

O DNA genômico dos isolados foi extraído utilizando o kit KIAamp (QIAGEN), seguindo-se as recomendações do fabricante. O DNA extraído foi visualizado através de gel de agarose 1,2%, contendo GelRed (Huang *et al.*, 2010), e quantificado utilizando Nanodrop 3000 (Thermo Scientific, USA). E todas as amostras apresentaram boa qualidade e quantidade suficiente para o sequenciamento. Posteriormente, as amostras foram liofilizadas no liofilizador Liotop K105 e preparadas para o envio.

4.4 Sequenciamento dos genomas

Os genomas dos isolados escolhidos nas etapas iniciais do projeto foram sequenciados usando a plataforma Illumina HiSeq 2500 do Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD - Unicamp). A *library* sequenciada foi criada com o kit TruSeq Nano DNA LT Library Prep (Illumina) consistia de reads pareados de 100pb de tamanho com insert size de 300, totalizando um tamanho de fragmento de 500 pb. A qualidade dos *reads* foi analisada utilizando-se o programa FastQC (Andrews, 2010). *Reads* de baixa qualidade e adaptadores foram removidos utilizando-se o programa Trimmomatic 0.36 com os parâmetros HC:10, AVGQUAL:30 (Bolger *et al.*, 2014).

4.5 Qualidade e pré-processamento dos *reads*

O sequenciamento na plataforma Illumina envolve duas etapas: análise de imagem e *base calling*. A análise das imagens objetiva corrigir o posicionamento da câmera entre os ciclos e aberrações cromáticas nas lentes. Atualmente essa correção é feita alinhando as imagens dos ciclos subsequentes a uma imagem de referência do ciclo inicial (Kao *et al.*, 2009). Essa etapa também busca identificar variações na intensidade de fluorescência e ruído. Para cada uma dessas regiões com fluorescência o processo de *base calling* converte os sinais fluorescentes no arquivo contendo as sequências e o score de qualidade. Os valores de qualidade phred (Ewing *et al.*, 1998) são os padrões utilizados para determinar a confiabilidade de cada base. Consistem, basicamente, de um valor logarítmico de probabilidade de erro, onde para a variável Q; a qualidade é P; e para probabilidade de erro temos: $Q = -10 \cdot \log_{10}$

(P) (Tabela 1).

Tabela 1. O score de qualidade phred e seu significado.

Phred (Q)	Probabilidade de erro	Precisão do <i>base call</i>
10	1 em 10	90%
20	1 em 100	99%
30	1 em 1.000	99,9%
40	1 em 10.000	99,99%
50	1 em 100.000	99,999%
60	1 em 1.000.000	99,9999%

4.6 Priorização de genomas

Devido ao grande volume de dados, foram priorizadas neste trabalho as cepas que mostraram resultados iniciais mais promissores tais como: qualidade do assembly, relevância da cepa, relevância do conteúdo genômico e resultados experimentais *in vivo* e *in vitro*.

4.7 Montagem dos genomas

Os genomas foram montados utilizando-se os softwares SPAdes 3.11.1 (Bankevich *et al.*, 2012) e Velvet 1.2.10 (Zerbino e Birney, 2008) ambos baseados em Grafos *de Brujin*. O software Velvet é composto por dois módulos principais, denominados Velveth e Velvg. O Velveth é o montador em si, ao passo que o Velvg é o *scaffolder*. Na etapa inicial define-se os parâmetros ao Velveth, como cobertura e tamanho do *kmer*. Neste momento foi utilizado o VelvetOptimiser (Gladman e Seemann, 2012), um script capaz de testar sucessivos parâmetros no Velveth, de forma automática, auxiliando na decisão sobre as melhores opções para obter o melhor assembly. O processo de scaffolding foi realizado com o SSPACE 2.0 (Boetzer *et al.*, 2011) visando incrementar a contiguidade. Finalmente, para averiguar as métricas das montagens foi utilizado o software QUAST 4.6.0 (Gurevich *et al.*, 2013).

4.8 Anotação dos genomas

Os genomas foram anotados utilizando-se o software PROKKA 1.12 (Seemann, 2014) e o pipeline PGAP – Prokaryotic Genome Annotation Pipeline do NCBI (Tatusova *et al.*, 2016). O PROKKA utiliza vários outros programas que fazem desde a predição *ab initio*, ou seja, inspeciona o genoma em busca de *open reading frames* (ORFs) até a comparação com base de dados para obter a função do gene predito. Para a busca de plasmídeos foram utilizados os softwares plasmidSPADEs 3.8.0 (Antipov *et al.*, 2016) e PlasmidFinder 1.3 (Carattoli *et al.*, 2014).

O PGAP foi lançado em 2008 e reformulado recentemente (Tatusova *et al.*, 2016). Combina a predição *ab initio* com a detecção de genes baseada em homologia. A plataforma é capaz de prever regiões codificadoras, RNAs estruturais, tRNAs, e pequenos RNAs não-codificadores. Exercendo papel crucial no pipeline (Figura 8), o algoritmo GeneMarkS+ combina estratégias de alinhamento e predições estatísticas e incorpora essas informações no processo de predição gênica (Besemer *et al.*, 2001). Neste projeto utilizamos as três abordagens descritas para garantir a melhor anotação possível dos genomas sequenciados.

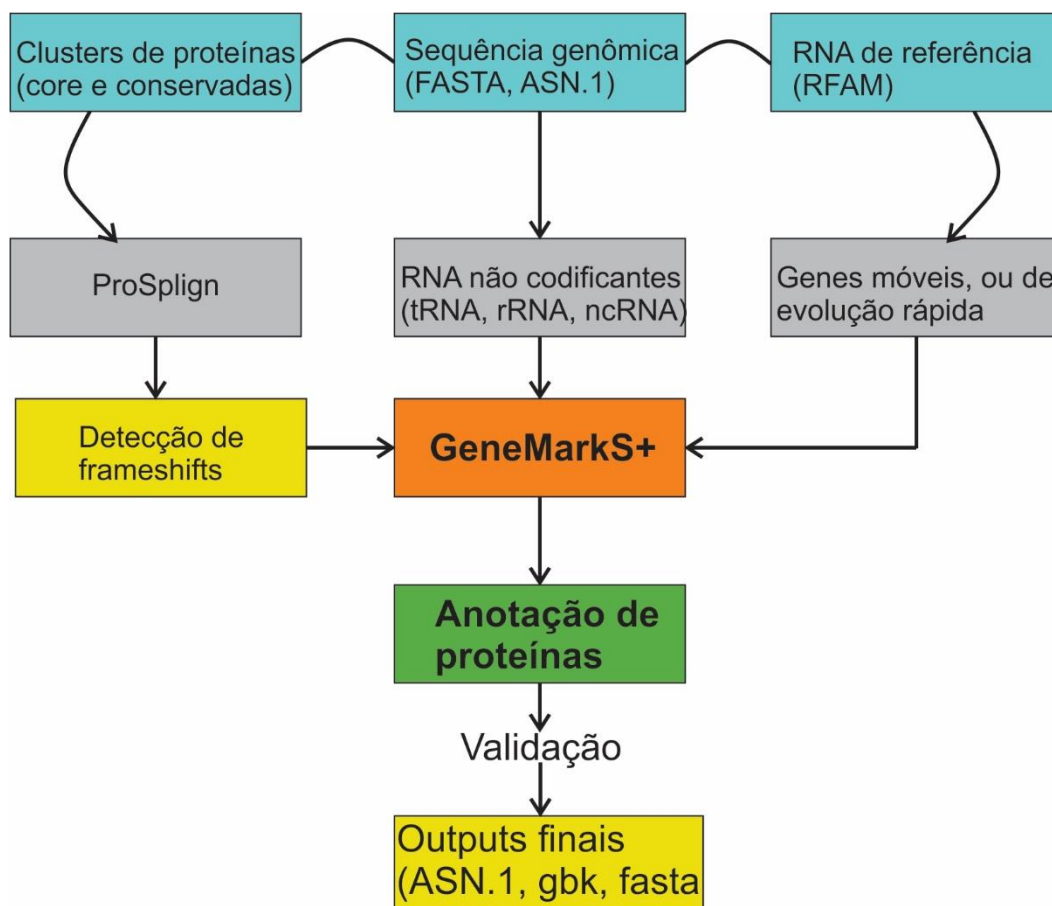


Figura 8. Componentes principais do PGAP. Adaptado de Tatusova *et al.*, (2016).

4.8.1 Anotação de operons

Em procaríotos, genes funcionalmente relacionados podem estar organizados de forma contígua em operons, que são tipicamente transcritos e em mRNAs policistrônicos (Merino *et al.*, 2008). A caracterização dos operons presentes em uma bactéria é capaz de fornecer relevantes informações sobre o microrganismo (Price *et al.*, 2006). Vários métodos computacionais foram desenvolvidos e utilizados para anotar operons em organismos modelo como *E. coli* (Salgado *et al.*, 2000) e *B. subtilis* (de Hoon *et al.*, 2004), que resultaram no desenvolvimento de alguns bancos de dados. Dentre estas, a base de dados DOOR se destaca, sendo capaz de fornecer o grau de similaridade entre operons relacionados em diferentes organismos (Mao *et al.*, 2009; Mao *et al.*, 2014). Apesar de outras plataformas estarem disponíveis (Perteza *et al.*, 2009; Dehal *et al.*, 2010), optamos por utilizar o DOOR2 neste trabalho, por se tratar de um projeto em continuidade e mais abrangente em suas possibilidades.

4.8.2 Anotação de vias metabólicas

A base de dados KEGG (Kanehisa *et al.*, 2011) visa organizar genes em vias metabólicas e de sinalização. Esta base de dados contém ainda informações e ortologia, permitindo o estudo de espécies não-modelo. Uma ferramenta que integra a informação existente na KEGG *Pathway* com um genoma recém-montado é o Blast-KOALA, que atua buscando grupos ortólogos e utiliza essa informação para associar genes a vias metabólicas de referência (Kanehisa *et al.*, 2016). Neste trabalho, utilizamos as vias do KEGG e a ferramenta Blast-KOALA para inferir as vias metabólicas dos genomas de interesse.

4.9 Visualização e integração

Os genomas montados e suas anotações foram visualizados com os softwares Artemis (Rutherford *et al.*, 2000) e Bandage (Wick *et al.*, 2015).

Os genes preditos e seus produtos foram extensivamente comparados com outros disponíveis publicamente utilizando a ferramenta BLAST (Altschul *et al.*, 1997), em associação com bases de dados como RefSeq (O'Leary *et al.*, 2015) e Uniprot (Consortium, 2016) Também foram realizadas análises de arquiteturas de domínios protéicos utilizando-se o banco de dados Pfam (Finn *et al.*, 2013) e o algoritmo HMMer 3.1 (Eddy, 1998).

4.10 Análise filogenéticas

Mais de 130.000 genomas bacterianos completos estão depositados no Genbank (NCBI, 2016). Entretanto, as relações filogenéticas entre muitos desses genomas sequenciados permanecem inexploradas (Wu *et al.*, 2009). Quando um novo procarioto tem seu genoma sequenciado, é usual utilizar a filogenia do RNAr 16s, a fim de situar o organismo no contexto filogenético (Lang *et al.*, 2013). O RNAr 16s constitui-se de uma ferramenta valiosa para esse propósito, pois possui regiões altamente conservadas (Janda e Abbott, 2007), além da grande disponibilidade, mesmo para bactérias não culturáveis e amostras ambientais. Entretanto, há de se ter cuidado, pois nem sempre a filogenia obtida com RNAr reflete aquela derivada de outros genes marcadores (Eisen, 1995).

4.10.1 Obtenção das sequências de RNAr 16s diretamente dos reads

O software EMIRGE 0.61.1 (Miller *et al.*, 2011) utiliza um algoritmo de mapeamento probabilístico (Dempster *et al.*, 1977) capaz de reconstruir a sequência do gene da subunidade menor do RNA (16s ou 18s). O programa utiliza-se dos *reads* e sua qualidade, além de uma base de dados curada manualmente (Miller *et al.*, 2011). No caso deste trabalho, foi utilizada a base de dados SILVA SSU/LSU 123 (Quast *et al.*, 2013).

4.10.2 Reconstrução filogenética

Árvores filogenéticas derivadas de dados moleculares precisam ser inferidas a partir de sequências (Yang e Rannala, 2012). O primeiro passo para a reconstrução filogenética é o alinhamento múltiplo de sequências, essencial para garantir que a filogenia obtida tenha valor biológico (Phillips *et al.*, 2000). Os alinhamentos múltiplos deste trabalho foram obtidos com o software MUSCLE 3.8.31 usando os parâmetros *default* (Edgar, 2004). A extração das regiões filogeneticamente relevantes, e a remoção de regiões ambíguas é importante para otimizar a análise (Castresana, 2000; Talavera e Castresana, 2007). Este procedimento foi realizado manualmente com auxílio do software JALVIEW (Waterhouse *et al.*, 2009).

As reconstruções filogenéticas foram realizadas com o método *Maximum Likelihood* (ML) (Felsenstein, 1981), a predição do modelo de evolução mais acurado foi feita utilizando o Prottest 3 (Darriba *et al.*, 2011) e a construção da árvore foi realizada através do software RAxML 8.2.10 com os parâmetros -m PROTGAMMALG, -#1000 -f a -x 12345 -p 12345 (Stamatakis, 2014). As árvores reconstruídas foram visualizadas com os programas Figtree 1.4.3 (Rambaut, 2009) e iTOL 3 (Letunic e Bork, 2016).

4.10.3 Análise taxonômica a partir dos reads

Para confirmar os achados filogenéticos provenientes das sequências obtidas realizou-se também a inferência filogenética diretamente a partir dos reads de sequenciamento, com auxílio do software Kaiju (Menzel *et al.*, 2016). O algoritmo traduz os *reads* nos seis possíveis *frames* de leitura e busca combinações dos aminoácidos nas bases de dados de proteínas para genomas bacterianos de referência. Os *reads* são então designados como provenientes

de determinada espécie com base na similaridade de sequências de aminoácidos. Os resultados desta análise foram visualizados com a ferramenta Krona (Ondov *et al.*, 2011).

4.10.4 Sintenia

A sintenia descreve o grau de preservação e co-localização de genes em diferentes espécies ou cepas filogeneticamente relacionadas. Com o intuito de analisar o grau de sintenia entre nossos isolados e genomas de referência, utilizamos o programa SIBELIA 3.0.6 (Minkin *et al.*, 2013). Para analisar o grau de conservação de pequenas regiões com genomas de referência utilizamos o software Synima 1.0 (Farrer, 2017).

4.10.5 Similaridade genômica

Com o aumento do número de genomas depositados, métricas que se valem da similaridade do genoma inteiro se tornaram cada vez mais difundidas, especialmente para taxonomia bacteriana. Uma dessas métricas é a identidade média de nucleotídeos (*average nucleotide identity*, ANI), que utiliza um blast prévio nas regiões codificadoras do genoma e constrói uma nova matriz de similaridade. Essa é comparada com uma matriz prévia de um mesmo taxon e a similaridade final é calculada (Richter e Rosselló-Móra, 2009). Outra técnica que também utiliza a sequência inteira do genoma, porém difere na metodologia do cálculo é a dDDH (*digital DNA-to-DNA hybridization*) (Auch *et al.*, 2010). Nesta abordagem utiliza-se o software blast para obtenção de *High Scoring Pairs* (HSPs) (pares de fragmentos alinhados), essas HSPs são submetidas a um algoritmo que constrói uma matriz e cria árvores filogenéticas pelo método UPGMA (Henz *et al.*, 2004). Os resultados são então convertidos para percentual de similaridade com intervalo de erro (Auch *et al.*, 2010). Neste trabalho utilizamos ambas metodologias (ANI e dDDH).

4.11 Pan-genoma

O termo pan-genoma foi cunhado em 2005 para descrever o conjunto de genes compartilhados em diferentes genomas de interesse (Medini *et al.*, 2005). Considerando que genomas bacterianos podem ganhar ou perder genes rapidamente, inclusive em resposta a alterações e estímulos ambientais (Vernikos *et al.*, 2015), a obtenção de um panorama de conservação gênica entre organismos da mesma espécie, bem como a busca por genes únicos, constitui-

se de valiosa ferramenta para o melhor entendimento de uma dada espécie.

As análises de pan-genoma deste trabalho foram realizadas com o software BPGA 1.3 (Chaudhari *et al.*, 2016), que agrupa as proteínas codificadas em todos os genomas utilizando o USEARCH (Edgar, 2010). Em seguida, é feita uma comparação todos contra todos utilizando BLASTP (Altschul *et al.*, 1997) com *cutoff* de 95 % de identidade. Os resultados são comparados e os genes presentes em todos os genomas (*core genome*), em mais de um genoma, porém não em todos (*accessory genome*) e em apenas um genoma (*unique*) são definidos.

4.12 Inspeção manual de regiões genômicas

Para identificar rearranjos gênicos específicos, realizamos a inspeção manual com auxílio do software BRIG 0.95 (Alikhan *et al.*, 2011), que determina regiões similares entre os genomas e as dispõem em uma comparação circular. As regiões pouco conservadas ou com evidência de transferência horizontal foram analisadas com as ferramentas AntiSMASH 4 (Blin *et al.*, 2017) e IslandViewer 4 (Bertelli *et al.*, 2017), que identificam vias do metabolismo secundário e possíveis eventos de transferência horizontal, respectivamente.

4.13 Testes *in vivo*

Sementes de milho (*Zea mays* var. UENF/506-11) tiveram sua superfície desinfetada com etanol (C₂H₆O) 70% por 30 segundos, em seguida lavados com hipoclorito de sódio 5 % (NaClO) por 20 minutos. Em seguida, as sementes foram lavadas com água destilada estéril 5 vezes sob pressão por 3 minutos e transferidas para placas de petri contendo 1.5 % ágar-água para pré-germinação por 4 dias.

As plântulas germinadas com 2,0 a 2,5 cm de comprimento radicular foram cuidadosamente transferidas, em uma cabine de fluxo laminar, para tubos de 2 cm de diâmetro e 20 cm de altura contendo 10 g de vermiculita estéril (uma semente por tubo). Enquanto isso o inóculo bacteriano (isolados 6GII, 14GII, 21GII, 22GI, 23GI, 24GI e 25GI) foi cultivado em meio DYGS líquido por 36 h, 30 °C, 120 rpm. A inoculação foi realizada através da aplicação da suspensão bacteriana (10⁸ células·mL⁻¹) sobre as plântulas. Sistemas inoculados com 1 mL do meio DYGS estéril foram usados como controle negativo.

O ensaio foi realizado em laboratório, com temperatura média de 30 °C e

12 h de fotoperíodo (claro/escuro). Após 10 dias, as plantas foram coletadas e os seguintes parâmetros foram registrados: altura (cm), comprimento radicular (cm), massa fresca da raiz, massa fresca da parte aérea (mg), massa seca da raiz (mg) e massa seca da parte aérea (mg). Esse ensaio foi realizado em 4 replicatas. Análises estatísticas foram realizadas utilizando o software SAEG (Universidade Federal de Viçosa, Brasil) (Ribeiro Júnior, 2001). As médias obtidas foram comparadas pelo teste de Tukey.

4.14 Testes *in vitro*

Com o intuito de averiguar propriedades de promoção do crescimento vegetal, testes *in vitro* foram realizados.

4.14.1 Solubilização de P e Zn

O inóculo bacteriano foi cultivado por 36 h em meio líquido Dygs a 150 rpm e 30 °C até atingir a densidade de 10^8 células·mL⁻¹ (O.D._{540nm} = 1.0) (Döbereiner *et al.*, 1995). Para o ensaio qualitativo de solubilização de fosfato, uma alíquota de 10 µl da suspensão de bactérias foi inoculado em placas de Petri contendo 10 g·L⁻¹ glicose (C₆H₁₂O₆), 5 g·L⁻¹ cloreto de amônia (NH₄Cl), 1 g·L⁻¹ cloreto de sódio (NaCl), 1 g·L⁻¹ sulfato de magnésio heptahidratado (MgSO₄·7H₂O), 15 g·L⁻¹ de ágar, fosfato de cálcio (Ca₃(PO₄)₂) 1 g·L⁻¹ em 1 L água destilada em pH 7,0, e incubada a 30 °C por 7 dias.

A solubilização de Zn foi avaliada a partir de uma alíquota de 10 µL de bactérias na mesma condição descrita acima. Essas foram inoculadas em placas de Petri contendo meio sólido constituído por: 10 g·L⁻¹ de glicose (C₆H₁₂O₆), 1 g·L⁻¹ sulfato de amônio ((NH₄)₂SO₄), 0,2 g·L⁻¹ cloreto de potássio (KCl), 0,1 g·L⁻¹ de fosfato dipotássico (K₂HPO₄), 0,2 g·L⁻¹ sulfato de magnésio heptahidratado (MgSO₄·7H₂O), 1 g·L⁻¹ de óxido de zinco (ZnO), 15 g·L⁻¹ ágar, 1 L água destilada; o meio foi incubado por 7 dias a 30 °C (Saravanan *et al.*, 2007).

Em ambos ensaios a presença de halo em torno das colônias foi considerada indicativo de crescimento bacteriano. As dimensões dos halos e das colônias foram utilizadas para calcular o Índice de Solubilização (IS), que é o diâmetro do halo (d1) dividido pelo diâmetro da colônia (d2). Os valores de IS podem ser usados para classificar a capacidade de solubilização de um isolado como: baixo (IS < 2), intermediário (2 < IS < 4) e alto (IS > 4) (Marra *et al.*, 2015).

O ensaio quantitativo de solubilização de P foi realizado em meio líquido

com a mesma composição do meio do ensaio qualitativo, a única diferença foi a ausência de ágar no meio líquido. Foram adicionadas duas fontes distintas de P: fosfato de cálcio ($\text{Ca}_3(\text{PO}_4)_2$) $1 \text{ g}\cdot\text{L}^{-1}$ e rocha fosfatada de fluorapatita ($\text{Ca}_{10}(\text{PO}_4)_6\text{F}_2$) at $1 \text{ g}\cdot\text{L}^{-1}$. Após 7 dias, uma alíquota de 5,0 mL do meio reacional foi centrifugada por 15 min a 3200 rpm, o sobrenadante foi colhido e o pH aferido para verificar acidificação. A quantidade total de P foi determinada no sobrenadante por espectrofotômetro usando o método de ácido ascórbico-azul de molibdênio (Murphy e Riley, 1962). Todos os testes foram realizados em triplicata e expressos em $\text{mg PO}_4^{2-}\cdot\text{L}^{-1}$.

4.14.2 Produção de compostos indólicos

Bactérias previamente crescidas foram transferidas para tubos de vidro contendo 5mL do meio Dygs com e sem adição de Trp ($100 \text{ mg}\cdot\text{L}^{-1}$), seguido por incubação por 72 h abrigado da luz, $30 \text{ }^\circ\text{C}$ e 150 rpm. Para avaliar a síntese de indol (Sarwar e Kremer, 1995), 150 μL de suspensão bacteriana foram transferidos para microplacas e adicionou-se 100 μL de reagente de Salkowski (1 mL de tricloro de ferro hexahidratado ($\text{FeCl}_3\cdot 6\text{H}_2\text{O}$), $92.5 \text{ g}\cdot\text{L}^{-1}$, em 50 mL ácido perclórico (HClO_4) $350 \text{ g}\cdot\text{L}^{-1}$ em água). A placa foi incubada por 30 min protegida da luz e as amostras foram analisadas sob luz UV a 492 nm, em um espectrofotômetro UV mini 1240 (Shimadzu, Japan).

4.14.3 Inibição do crescimento de fungos fitopatogênicos

O ensaio de inibição foi realizado em uma placa de Petri de 9 cm de diâmetro contendo meio sólido ágar batata dextrose. Hifas de *Fusarium solani* e *F. oxysporum* em ativo crescimento foram retiradas e inoculadas na placa. Os fungos testados foram gentilmente cedidos pela Profa. Valdirene Gomes (LFBM/CBB/UENF). Uma suspensão bacteriana foi inoculada no interior do meio em 4 posições equidistantes em relação ao inóculo fúngico no centro da placa. O grupo controle se constituiu dos fungos na ausência da bactéria e foram realizados em paralelo, os tratamentos foram feitos por 10 dias e três replicatas independentes foram realizadas. Amostras da zona de transição entre as estruturas fúngicas e a área da colônia bacteriana foram colocadas em lâminas de vidro e lamínula e observadas sob microscopia invertida de contraste de fase em um microscópio Zeiss Axio 10 Observer A1 e fotodocumentadas com câmera digital AxioCam MRC 5.

5. RESULTADOS E DISCUSSÃO

5.1 Testes preditivos *in vitro*

Em um trabalho anterior (Aguiar, 2012), bactérias isoladas de diferentes vermicompostos foram testadas quanto à diversas características promotoras do crescimento vegetal como: solubilização de fósforo e zinco, degradação de celulose, FBN e produção de auxina. Foi selecionado um grupo de 18 isolados para posterior sequenciamento, a Tabela 2 evidencia os testes de promoção *in vitro* realizados por Aguiar (2012) para essas 18 bactérias, todas testaram positivamente (+) para pelo menos 1 das características de promoção do crescimento vegetal aferidas (solubilização de P, solubilização de Zn, degradação de celulose, FBN e produção de auxina). Algumas das bactérias apresentaram resultados mais promissores nesses testes como a 22GI e a 24GI que foram positivas para todos os testes aferidos, ao passo que outros isolados mostraram resultados distintos testando majoritariamente negativos (-) como a 26GIII (Tabela 2).

Tabela 2. Perfil bioquímico da promoção do crescimento em testes *in vitro*.

Cepas	Solub. P	Solub. Zn	Deg. celulose	FBN	Prod. auxina
4GII	+	-	+	+	+
5BGII	+	+	-	+	+
6GII	+	+	-	+	+
14GII	+	+	-	+	+
17GIII	+	-	-	+	+
21GII	+	-	-	+	+
22GI	+	+	+	+	+
23GI	+	+	+	-	+
24GI	+	+	+	+	+
25GI	+	-	+	+	+
25GII	+	+	-	+	+
25GIII	+	+	-	+	+
26GII	+	+	-	+	+
26GIII	-	-	-	+	+
65	+	-	-	+	+
98	+	+	+	-	-
101	+	+	+	+	-
103	-	+	-	+	-

Experimentos de *screenings in vitro* são comumente aplicados como

etapa inicial de triagem para estudos posteriores, almejando escolher para as etapas sequenciais apenas os isolados mais promissores. Essa mesma abordagem foi realizada por Sayed e colaboradores (2014), onde um screening de 531 isolados concluiu que 66 desses possuíam o fenótipo desejado (El-Sayed *et al.*, 2014).

5.2 Sequenciamento

Os *reads* foram obtidos após o sequenciamento, quantificados e analisados no que tange a qualidade das bases (Phred+33), o resultado encontra-se disposto na Tabela 3.

Tabela 3. Quantidade e qualidade dos reads obtidos pelo sequenciamento.

Isolados	Nº de reads	%bases \geq Q30
4GII	60.186.890	88,82
5BGII	15.943.718	87,44
6GII	12.446.212	90,95
14GII	16.892.038	90,24
17GIII	17.299.154	87,88
21GII	23.203.014	89,13
22GI	21.445.242	89,21
23GI	12.765.434	91,35
24GI	14.578.952	90,93
25GI	14.664.222	91,46
25GII	23.070.346	96,23
25GIII	31.250.186	85,84
26GII	47.695.756	74,00
26GIII	38.174.643	88,01
65	23.547.288	87,55
98	29.058.964	86,84
101	26.326.372	86,49
103	11.470.538	87,90

5.3 Análises da qualidade dos *reads* sequenciados.

A qualidade dos reads foi analisada com o programa fastQC. A qualidade de uma base sequenciada é codificada por um sistema Phred, representado no eixo y. No eixo x encontra-se a posição das bases dentro dos reads de 100 pb. A Figura 9 mostra a qualidade dos reads *forward* da cepa UENF-22GI como

forma de ilustrar a qualidade obtida no sequenciamento. A ausência de indicadores de qualidade (barra amarela) nos 8 primeiros reads deve à adição de tags (bar codes) do sequenciamento, que foram aparados por meio do software Trimmomatic (Bolger *et al.*, 2014). Reads com qualidade média inferior a 30 foram excluídos com o uso do mesmo software.

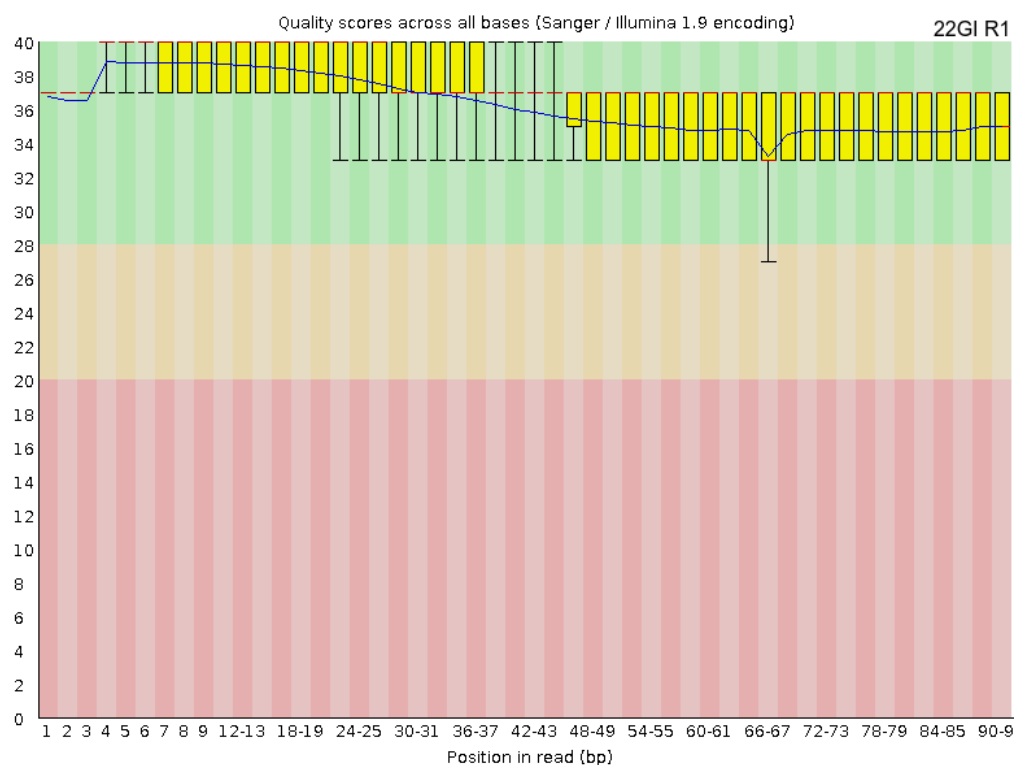


Figura 9. Boxplot com a distribuição de qualidade de uma das bibliotecas sequenciadas (22GI forward reads) ao longo das coordenadas dentro dos reads. Esta figura é representativa de todas as bibliotecas sequenciadas no projeto.

5.4 Inferências filogenéticas obtidas a partir dos reads - 16s rRNA

A análise das sequências do gene da porção 16s do rRNA dos isolados nos permitiu realizar inferências preliminares acerca de suas identidades através de BLASTn (Altschul *et al.*, 1997) contra a base de dados de referência de genomas microbianos RefSeq do NCBI (Tatusova *et al.*, 2013). Com base nesses resultados (Tabela 4) nossos isolados se concentram nos gêneros *Stenotrophomonas* (4GII, 5BGII, 17GIII, 25GIII, 26GII e 26GIII) e *Enterobacter* (6GII, 14GII, 21GII, 23GI, 24GI e 25GI). Outras cepas incluem *Paraburkholderia* (101 e 65), *Azospirillum* (98), *Serratia* (103 e 22GI) e *Staphylococcus* (25GII)

respectivamente. Uma vez que os isolados 65, 98, 101 e 103 foram obtidos de um mesmo grupo de amostras (fruteiras) foram designados para ser analisados na dissertação de mestrado do aluno Gustavo Rodrigues, defendida recentemente no âmbito do Programa de Pós-graduação em Biociências e Biotecnologia da UENF. Já as cepas de *Stenotrophomonas* (4GII, 5BGII, 17GIII, 25GIII, 26GII e 26GIII) são tema do projeto de dissertação de mestrado do estudante Francisnei Pedrosa, no mesmo programa de pós-graduação. Por fim, as amostras destacadas em vermelho foram escolhidas como foco do presente trabalho.

Tabela 4. Tamanho das sequências do gene do rRNA 16s obtidas a partir dos reads, hit mais provável com respectivo NR e grau de identidade.

Cepa	Tamanho (pb)	Blastn top hit	Acession	Identidade (%)
4GII	1623	<i>Stenotrophomonas maltophilia</i> ATCC19861	NR_040804.1	99
5BGII	1272	<i>Stenotrophomonas maltophilia</i> ATCC19861	NR_040804.1	99
6GII	1517	<i>Enterobacter ludwigi</i> EN-119	NR_042349.1	96
14GII	1491	<i>Enterobacter aerogenes</i> KCTC 2190	NR_102493.1	99
17GIII	1451	<i>Stenotrophomonas maltophilia</i> ATCC19861	NR_040804.1	98
21GII	1380	<i>Enterobacter bugandensis</i>	NR_148649.1	98
22GI	1275	<i>Serratia marcescens</i> WW4	NR_102509.1	99
23GI	1440	<i>Enterobacter aerogenes</i> KCTC2190	NR_102493.1	99
24GI	1413	<i>Enterobacter ludwigi</i> EN-119	NR_042349.1	99
25GI	1478	<i>Enterobacter ludwigi</i> EN-119	NR_042349.1	99
25GII	1372	<i>Staphylococcus aureus</i> N315	NR_075000.1	99
25GIII	1554	<i>Stenotrophomonas maltophilia</i> ATCC13637	NR_112030.1	99
26GII	1272	<i>Stenotrophomonas pavanii</i> LMG25348	NR_117008.1	99
26GIII	1376	<i>Stenotrophomonas maltophilia</i> IAM12423	NR_041577.1	99
65	1250	<i>Paraburkholderia silvatlantica</i> SRMrh-20	NR_043306.1	98
98	1198	<i>Azospirillum brasilense</i> ATCC29145	NR_042845.1	97
101	1254	<i>Paraburkholderia silvatlantica</i> SRMrh-20	NR_043306.1	99
103	1375	<i>Serratia marcescens</i>	NR_036866.1	99

Os gêneros encontrados em nossa análise da sequência do 16s já foram previamente reportados como potenciais bactérias promotoras do crescimento vegetal (Goswami *et al.*, 2016).

5.5 Inferências filogenéticas obtidas diretamente a partir dos reads

A análise dos resultados demonstrados na Tabela 4 indicou a predominância dos gêneros *Stenotrophomonas* e *Enterobacter* entre os isolados sequenciados. Almejando confirmar esses dados, verificamos a filogenia utilizando os reads totais. A Figura 10 ilustra o panorama de todos os reads sequenciados: 91 % são *proteobacterias*, 8 % *terrabacteria* e 1 % (56101 reads) não foram designados. Dentro do filo *terrabacteria* a única espécie significativamente representada foi *S. aureus*. Outro filo encontrado é o das *proteobacterias* compostas unicamente pela classe *Gammaproteobacterias*. No âmbito das *Gammaproteobacterias*, particularmente das ordens *Xanthomonadales* (rosa) e *Enterobacteriales* (marrom). Dentro da ordem *Xanthomonadales* o único gênero que aparenta estar presente é o das *Stenotrophomonas*. Já em *Enterobacteriales* há dois gêneros: *Enterobacter* e *Serratia*.

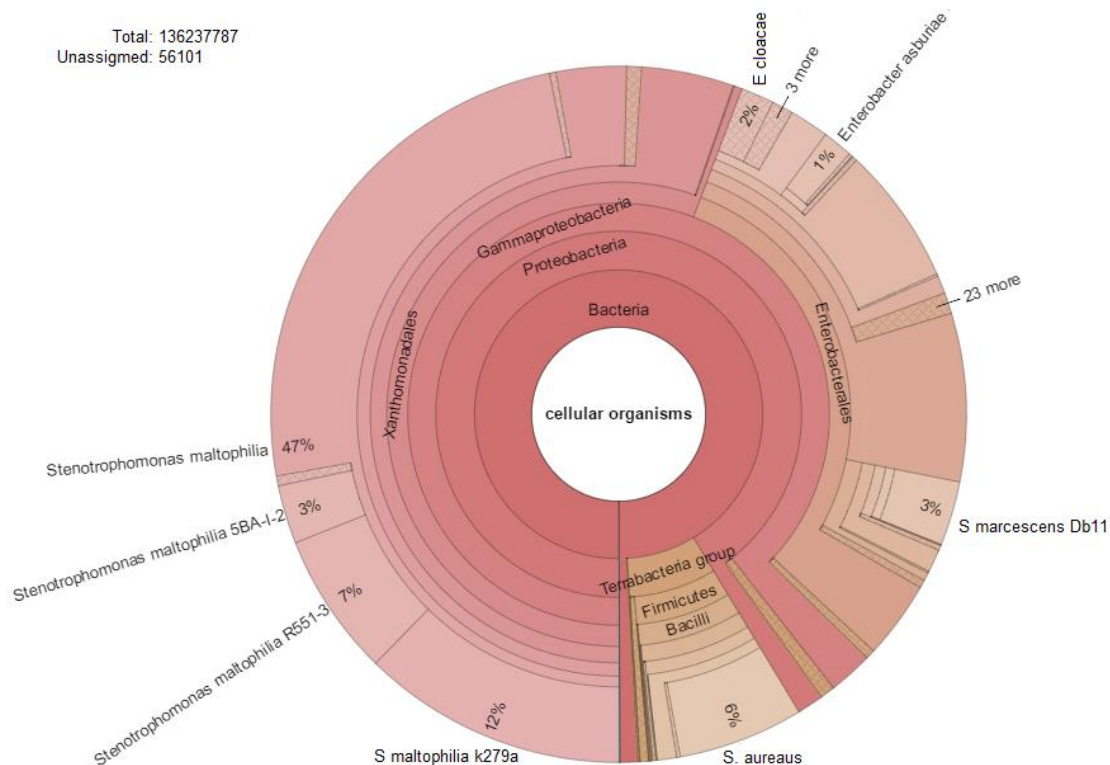


Figura 10. Representação circular da filiação taxonômica dos reads totais computados com KAIJU (Menzel *et al.*, 2016) e visualizados através do KRONA (Ondov *et al.*, 2011).

5.6 Montagem dos genomas

Os genomas foram montados utilizando um pipeline para otimizar os parâmetros de assembly como escolha do k-mer e coverage-cutoff, que refletem, respectivamente, a porção do reads utilizada na montagem e o limiar mínimo de cobertura para que um determinado contig montado seja levado em consideração. O software VelvetOptimiser (Gladman e Seemann, 2012) foi utilizado para aferir esses parâmetros de modo eficiente e os valores obtidos foram então utilizados no software SPAdes (Bankevich *et al.*, 2012) que via de regra apresentou os melhores resultados (Tabela 5). As métricas foram avaliadas com o software QUAST (Gurevich *et al.*, 2013).

O pré-tratamento dos *reads* variou de acordo com cada situação e, principalmente, com a qualidade e tamanho da biblioteca. O processo de filtragem dos *reads* por qualidade visa garantir que somente bases sequenciadas com maior qualidade sejam usadas nas análises subsequentes, diminuindo assim as taxas de erro.

Tabela 5. Estatísticas das montagens obtidas para os isolados selecionados

Métricas	Nº scaffs	Tamanho total	Tamanho médio scaffs	Maior scaff	GC (%)	N50	L50
Cepas							
6GII	86	4.936.020	57.395	1.084.112	56,25	455.047	4
14GII	65	4.874.681	75.995	1.540.454	56,15	478.433	3
21GII	38	4.952.588	130.331	1.665.897	55,62	935.866	2
22GI	17	5.001.584	294.210	3.077.593	59,78	3.077.593	1
23GI	26	4.733.744	187.067	2.648.146	56,07	2.648.146	1
24GI	25	4.730.472	189.218	2.419.931	56,06	2.419.931	1
25GI	30	4.730.933	157.697	2.419.897	56,06	2.419.897	1
25GII	17	2.791.796	164.233	567.966	32,71	275.001	3

As métricas analisadas estão relacionadas na Tabela 5 e se relacionam com a quantidade de fragmentos obtidos e o tamanho destes. Dado um grupo de sequências de tamanhos variados ordenados do maior para a menor, o N50 é o tamanho N para o qual 50% das bases estão presentes nos *scaffolds* (scaffs) (Nagarajan e Pop, 2013). O L50 descreve a quantidade de *scaffolds* contendo 50 % do tamanho total dos genomas, contando-se do maior para o menor. A mesma lógica pode ser aplicada ao N90 (valor que representa 90 % do tamanho total do genoma). Todas as montagens obtidas foram satisfatórias, com destaque para o *assembly* da cepa UENF-22GI, que apresentou um genoma de cerca de 5 Mb, montado em 17 fragmentos e possuindo um N50 de 3 Mb.

5.6.1 Presença de genes ortólogos conservados

Uma forma de aferir a qualidade de um *assembly* é determinar a presença de genes conservados para determinado táxon. O software BUSCO (Simao *et al.*, 2015) implementa uma ferramenta de busca por ortólogos únicos através de diferentes bases de dados de referência preparadas pelos próprios autores (<http://busco.ezlab.org/>). Utilizando as predições obtidas na Tabela 4 escolhemos a base de dados da ordem *Enterobacteriales* para analisar os genomas dos isolados 6GII, 14GII, 21GI, 22GI, 23GI, 24GI e 25GI e da base de dados da ordem *Bacillales* para a cepa 25GII. As bases de dados de *Enterobacteriales* e *Bacillales* são compostas por 781 e 526 genes em cópia única, respectivamente (Figura 11).

BUSCO Assessment Results

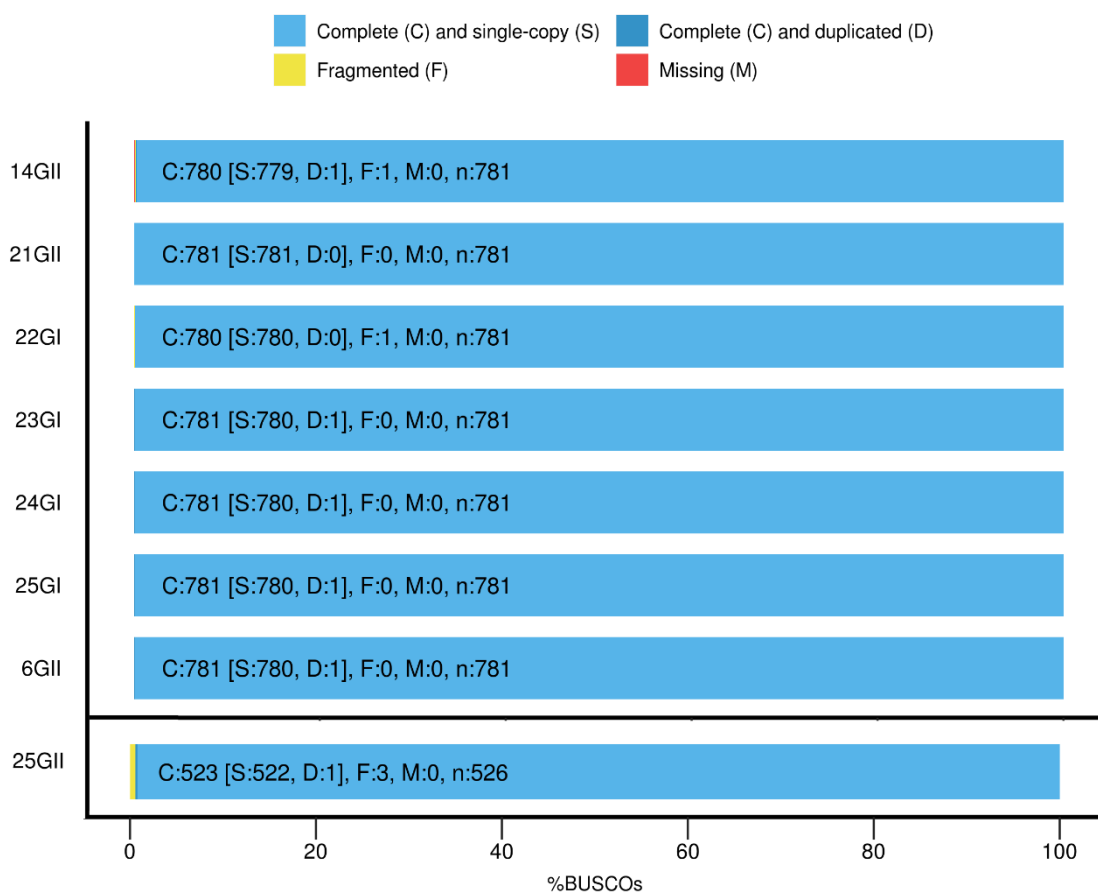


Figura 11. Presença de genes ortólogos únicos da ordem *Enterobacteriales* nos isolados 6GII, 14GII, 21GI, 22GI, 23GI, 24GI, 25GI e *Bacillales* em 25GII.

Fomos capazes de recuperar a totalidade destes genes em nossos genomas, o que evidencia a boa qualidade das montagens. Há um número crescente de ferramentas projetadas para ajudar a validar *assemblies*, algumas estratégias são: programas capazes de lidar com regiões altamente repetitivas (Bresler *et al.*, 2012), montadores que utilizam proteínas ortólogas para otimizar montagens (Li e Copley, 2012) e ferramentas que corrigem erros de montagem (Kelley e Salzberg, 2010). Iniciativas também foram implementadas no sentido de comparar os montadores e seus *outputs*, projetos como: GAGE (genome assembly gold-standard evaluations) (Salzberg *et al.*, 2012) e Assemblathon (Earl *et al.*, 2011). Diversas métricas foram propostas como forma de validar um assembly, algumas aferem as regiões e proporções nas quais os reads estão presentes na montagem (Clark *et al.*, 2013; Rahman e Pachter, 2013), enquanto outra infere regiões errôneas na montagem (Hunt *et al.*, 2013). De uma forma geral a utilização de um conteúdo gênico esperado, como forma de validação da

montagem do genoma é uma abordagem elegante e bem consolidada (Zdobnov *et al.*, 2016).

5.7 Resultados da cepa *Serratia marcescens* UENF-22GI

Tendo em vista os resultados obtidos nos testes *in vitro* (Tabela 2) e na montagem do genomas (Tabela 5), aprofundamos as análises, inicialmente, na cepa UENF-22GI. O isolado denominado *Serratia marcescens* UENF-22GI (SMU) foi montado em 17 scaffolds (length \geq 500 bp), com tamanho total de 5.001.584 pb, 59,8% de conteúdo de GC, N50 de 3.077.593 pb e tamanho médio de 294.210 (Tabela 5). A Figura 12 mostra a visualização circular do *assembly* obtido.

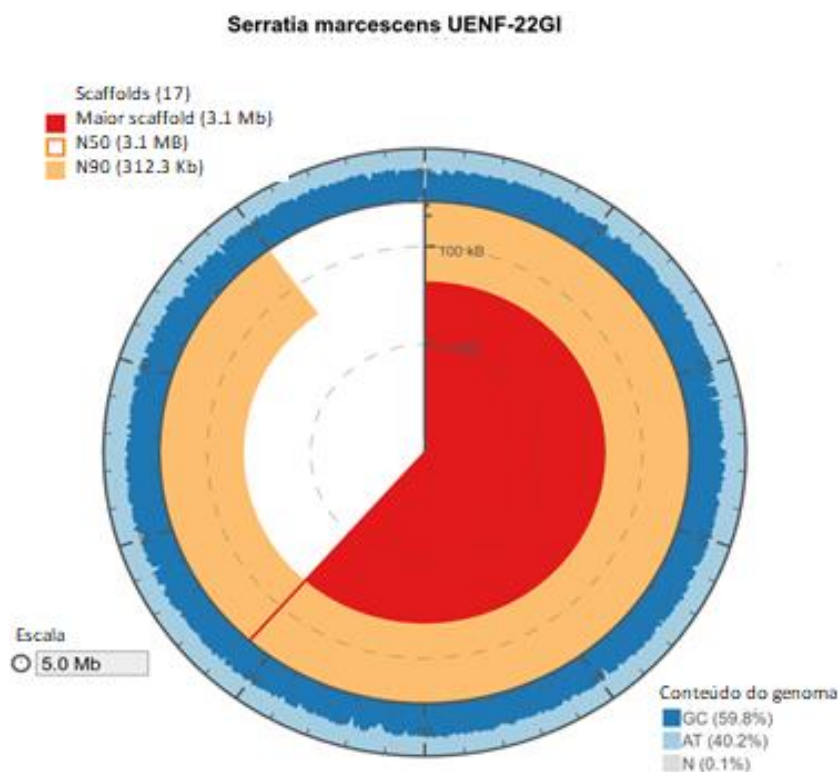


Figura 12. Visualização circular do genoma de *S. marcescens* UENF-22GI. O conteúdo de GC ao longo de todo o genoma (círculo azul mais externo). O N50, que coincide com o maior scaffold (em vermelho) e o N90 em laranja.

Genomas de referência se constituem de uma relevante fonte de informações a um projeto de sequenciamento *de novo*, pois estão completos e, geralmente, bem anotados (Buza *et al.*, 2015). Assim sendo, computamos as

estatísticas da nossa montagem obtida e comparamos com outros genomas de cepas filogeneticamente próximas de SMU Tabela 6.

A diversidade de sítios de isolamento de *S. marcescens* explicita a plasticidade do genoma dessa bactéria (Iguchi *et al.*, 2014), SMU apresenta tamanho de genoma e conteúdo de GC similares aos genomas de referência mostrados. Entretanto, a anotação recuperou apenas 11 rRNAs enquanto todas as cepas de referência apresentadas possuem 22 genes, o que pode estar associado a conservação dessas moléculas, acarretando acúmulo de reads similares dificultando a montagem. O genoma de SMU não apresenta plasmídeos, que foram buscados através de plasmidSPADES e plasmidFinder (Carattoli *et al.*, 2014; Antipov *et al.*, 2016).

Tabela 6. Comparativo do genoma de SMU e outras cepas de *S. marcescens* filogeneticamente próximas.

	SMU	WW4	Db11	B3R3	RSC-14	UMH8
Tamanho genoma (Mb)	5,01	5,24	5,11	5,47	5,12	5,15
Status	Incompleto	completo	completo	completo	completo	completo
GC (%)	59,7	59,6	59,5	59,03	59,6	59,6
Nº. de plasmídeos	0	1	0	1	0	0
Nº CDS	4.528	4.830	4.723	5.150	4.593	4.736
Nº tRNAs	84	81	87	96	88	89
Nº rRNAs	11	22	22	22	22	22
Isolamento	Vermicomposto	Máquina de papel	<i>Drosophila</i>	<i>Zea mays</i>	<i>Solanum nigrum</i>	<i>Homo sapiens</i>

5.8 Caracterização *in vitro* do potencial promotor do crescimento vegetal de SMU

Após a associação de resultados anteriores *in vitro* animadores (Tabela 2) e uma satisfatória montagem do genoma da cepa SMU (Tabela 5) decidimos expandir as análises *in vitro* desse isolado para caracterizar os mecanismos de promoção do crescimento vegetal.

5.8.1 Solubilização de P e Zn

Conforme discutido anteriormente, a disponibilidade de P e Zn no solo são fatores que limitam a produtividade agrícola. Para aferir a capacidade de solubilizar esses nutrientes, meios contendo fontes de P e Zn insolúveis foram utilizados para o cultivo da bactéria, utilizando-se da formação do halo para avaliar o nível de crescimento microbiano. Os valores de IS foram $2,11 \pm 0,47$, para P (Figura 13a) e $2,47 \pm 0,22$, para Zn (Figura 13b), esses valores de IS são considerados intermediários para a solubilização. Em um trabalho recente os índices de solubilização reportados para isolados potenciais PGPR apresentaram IS menores do que os aqui reportados para solubilização de P (Meena *et al.*, 2015), alguns valores foram reportados como sendo maiores do que os aqui relatado, porém a fórmula utilizada nesse ensaio foi diferente da que utilizamos aqui, acarretando aumento não fidedigno dos valores.

Em um ensaio quantitativo, verificamos também que há acentuada diferença na quantidade de P liberado no meio, tanto no substrato de fosfato de cálcio quanto em rocha de Araxá (Figura 13c), concomitantemente à diminuição do pH (Figura 13d). Na literatura a solubilização microbiana de P e Zn é relacionada à produção de ácidos que, ao diminuir o pH facilitam a mobilização desses nutrientes (Çakmakçı *et al.*, 2017), os resultados indicam que SMU é capaz de realizar esse mecanismo.

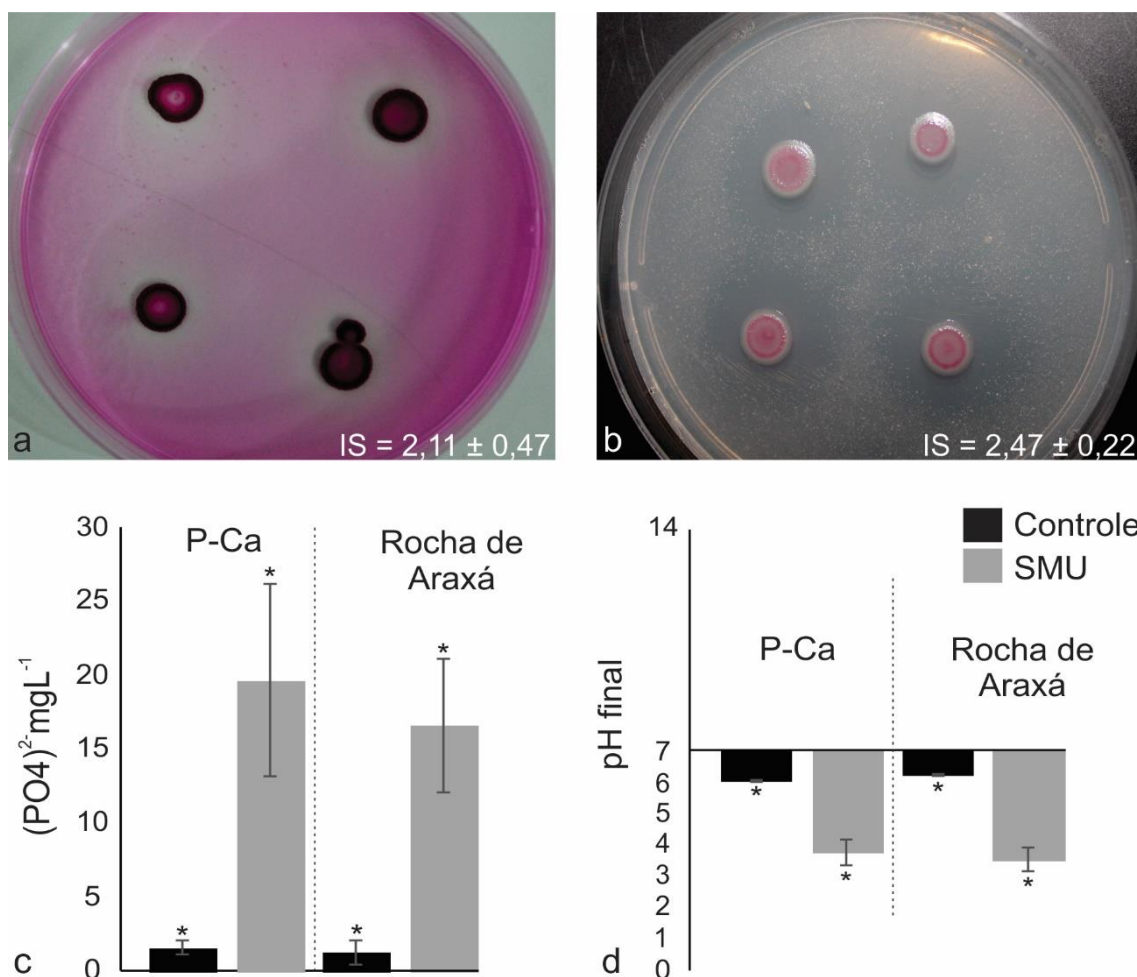


Figura 13. a) Solubilização de fosfato de cálcio ($\text{Ca}_3(\text{PO}_4)_2$). b) Solubilização de óxido de zinco (ZnO). c) Quantificação da solubilização de fosfato de cálcio e da rocha de Araxá d) Valor de pH ao final o tratamento de fosfato de cálcio e rocha de Araxá. As barras pretas são os sistemas controle, as cinzas os inoculados na presença de SMU.

5.8.2 Síntese de compostos indólicos

Dentre os fitormônios produzidos por PGPRs, a auxina é sem dúvida a mais estudada, sendo diretamente relacionada ao crescimento vegetal (Maheshwari *et al.*, 2015). A anotação do genoma de SMU revelou a presença do gene considerado central da via AIP, o *ipdC*, que codifica a enzima indol-3-piruvato Descarboxilase (Figura 14a). Em testes *in vitro* analisamos a conversão do indol, um teste indireto que indica a capacidade da bactéria produzir AIA. Verificamos que SMU é capaz de converter indol tanto na presença quanto na ausência de suplementação de Trp (Figura 14b), indicando que pode estar presente mais de uma via de produção de auxina em SMU.

Já foi mostrado que a inativação do gene *ipdC* na diminuição drástica da produção de IAA (Malhotra e Srivastava, 2008). Além disso também

encontramos nas anotações dois genes que codificam bombas de efluxo de auxina, indicando que SMU realiza transporte de IAA para o exterior das células.

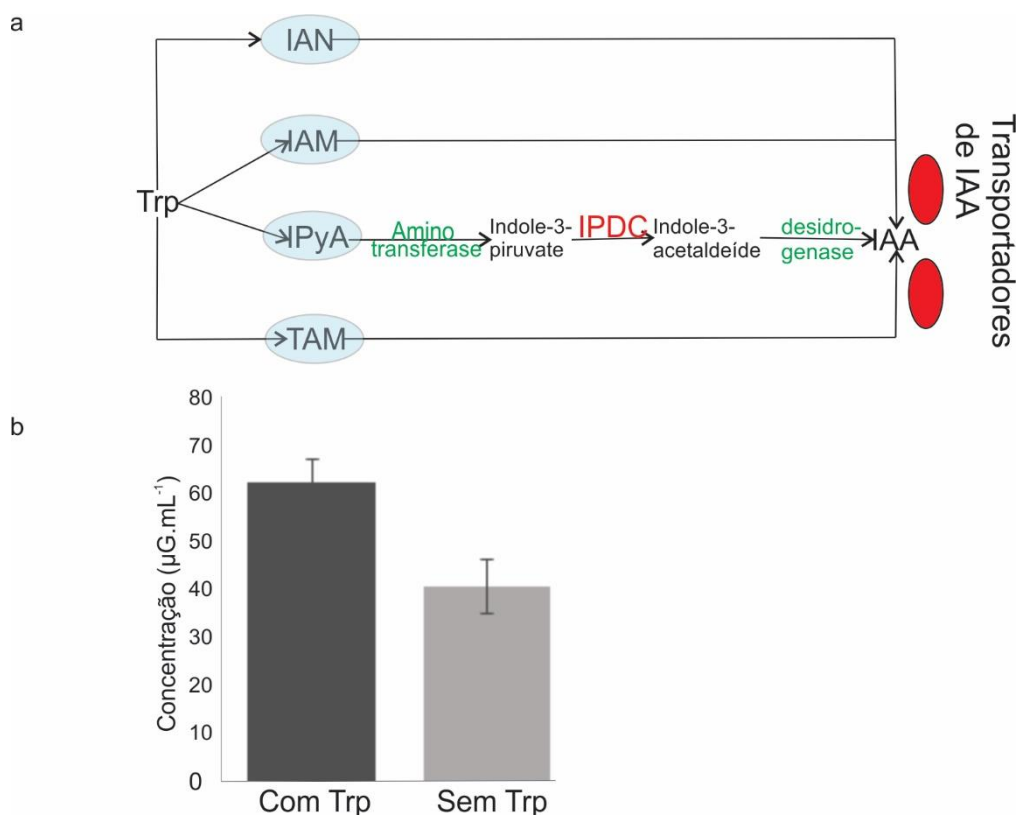


Figura 14. a) Vias de produção de auxina conhecidas em bactérias, em vermelho o gene presente em SMU. b) Produção de compostos indólicos por SMU com suplementação de Trp ou não.

5.8.3 Inibição de fungos fitopatogênicos

Fungos fitopatogênicos representam um enorme obstáculo para a agricultura (Oerke, 2006). *Fusarium* representa um gênero comumente associado à vegetais superiores e de distribuição ubíqua em ecossistemas terrestres (Ploetz, 2006). Em nossos ensaios mostramos que SMU é capaz de inibir o crescimento de *F. oxysporum* e *F. solani* (Figura 15), ambos fitopatogênicos que infectam culturas como feijão e soja respectivamente, além de muitas outras. Além disso evidenciamos que os mecanismos principais pelos quais se dá o biocontrole do patógeno fúngico são: a produção de prodigiosina (o pigmento avermelhado disseminado na placa) (Figura 15b, e) e produção de biofilme, que permite a adesão direta às hifas (Figura 15c, f). Um outro trabalho reporta um isolado de *S. marcescens* também capaz de inibir os fungos *Fusarium oxysporum* e *Fusarium graminearum* (Singh e Jha, 2016).

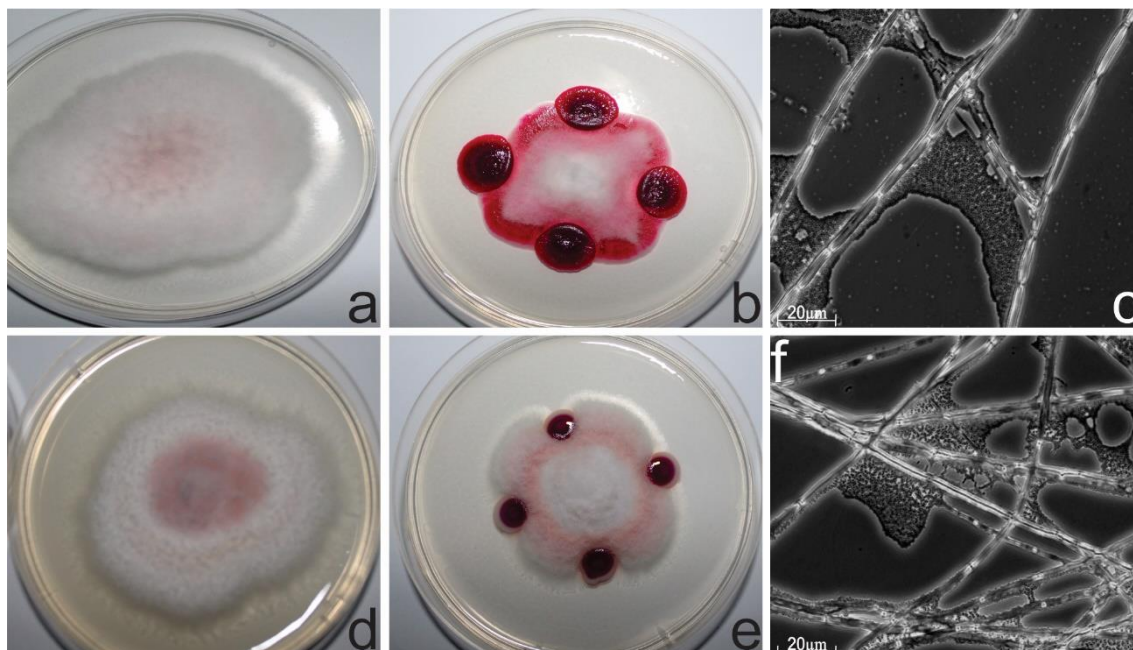


Figura 15. Ensaio de inibição de duas espécies de fungos fitopatogênicos por SMU. Sistemas controle foram compostos somente pelos fungos *F. oxysporum* e *F. solani* (a, d, respectivamente). No sistema de co-inoculação SMU foi inserida em 4 pontos equidistantes de *F. oxysporum* e *F. solani* (b, e, respectivamente). A aderência da bactéria às hifas fúngicas foi demonstrada através de microscopia óptica em *F. oxysporum* e *F. solani* (c, d, respectivamente)

5.9 Atividade promotora do crescimento de plântulas de milho

Realizamos um experimento gnotobiótico para avaliar a capacidade de SMU em promover o crescimento vegetal utilizando milho como modelo. A caracterização do crescimento em si é o efeito visível da soma de todas características benéficas que um PGPR fornece *in vivo*. Aplicamos uma suspensão de células de SMU em sementes pré-germinadas de milho. Após 10 dias foram medidos: massa fresca da raiz (MFR), massa fresca da parte aérea (MFPA), massa seca da raiz (MSR), massa seca da parte aérea (MSPA), comprimento radicular (CR) e altura (A).

A presença de SMU levou a incrementos substanciais na massa seca e fresca, tanto do sistema radicular quanto na parte aérea. Também foram incrementados os comprimentos das raízes e a altura da plântula (Figura 16). O incremento de biomassa foi de 100 % na altura (A) e no comprimento radicular (CR). Para MFR 80 %, para MFPA 64 % e 150 % para MSR e MSPA. Alguns estudos mostraram efeitos benéficos de outras cepas de *S. marcescens* em plantas, como a atenuação de estresse osmótico em trigo (Singh e Jha, 2016) e a promoção do crescimento em gengibre (Dinesh *et al.*, 2015).

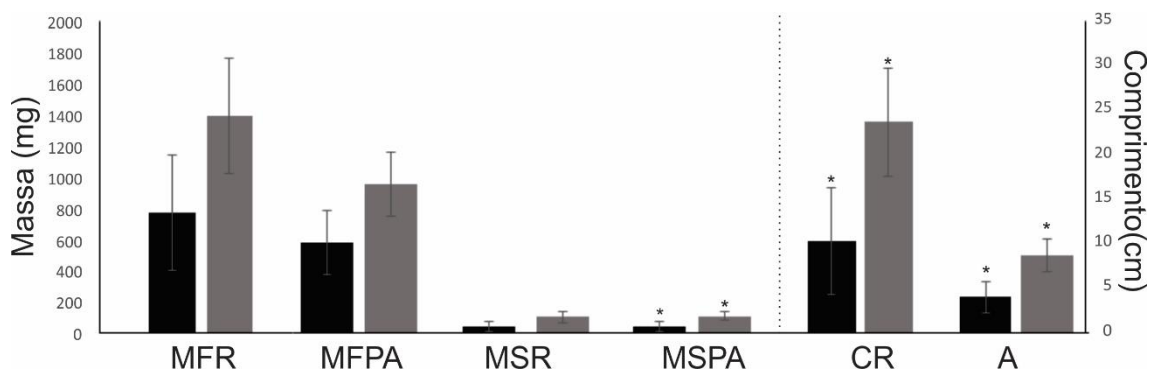


Figura 16. Efeito da inoculação de SMU no crescimento de plântulas de milho. Plântulas germinadas (com comprimento radicular de 2 a 2,5 cm) foram transferidas para todos de vidro contendo vermiculita estéril (uma plântula por tubo). A inoculação foi feita através da aplicação de 1 mL de suspensão bacteriana (10^8 células·mL⁻¹) sobre as plântulas. Sistemas inoculados com 1 mL de meio DYGS estéril foram usados como controle negativos. As seguintes medidas foram aferidas após 10 dias: Massa fresca da raiz (MFR), massa fresca da parte aérea (MFPA), massa seca da raiz (MSR), massa seca parte aérea (MSPA), comprimento radicular (CR) e altura (A). Todos os experimentos foram conduzidos em triplicata e diferenças significativas foram aferidas pelo teste de Tukey ($p < 0.05$).

5.10 Reconstrução do pan-genoma de SMU

O isolado SMU foi classificado como *Serratia marcescens* de acordo com a similaridade observada para a molécula do rRNA 16s (Tabela 4). Realizamos então a análise de pan-genoma de SMU com outros isolados de *S. marcescens* a fim de melhor entender a diversidade genômica da espécie. O passo inicial para a construção de um pan-genoma é a determinação dos genomas utilizados na análise. Construímos um pan-genoma a nível de espécie utilizando 35 genomas completos de *S. marcescens* (táxon 615 NCBI). Foram utilizados tanto genomas de isolados ambientais, quanto outros de relevância clínica (Tabela 7).

O pan-genoma em si se configura então, da relação do total de famílias gênicas presentes em todos os 35 genomas analisados. Um total de 16.456 famílias gênicas foram encontradas. Destas, 2.107 pertencem ao core genoma, ou seja, estão presentes em todos os isolados. O genoma acessório é composto por 7.656 genes, esses genes são aqueles presentes em mais de um genoma, porém não em todos. Todo o restante são genes únicos de cada cepa, desses 57 estão presentes somente no genoma de SMU (Figura 17).

Um outro estudo descrevendo um pan-genoma de *S. marcescens* foi realizado recentemente e identificou 13.614 genes presentes no pan-genoma, sendo 3.372 core genes e 10.215 acessórios e apenas 27 genes únicos

(Moradigaravand *et al.*, 2016), nesse estudo foram utilizadas 205 cepas clínicas do Reino Unido e Irlanda. Ao comparar os valores obtidos no nosso pan-genoma e nesse trabalho evidenciamos que o número de genomas por si não se relaciona com a quantidade de famílias gênicas encontradas, uma vez que obtivemos um pan-genoma com mais famílias gênicas, mesmo construindo-o com 35 isolados, porém de diversos sítios de isolamento.

Tabela 7. Genomas de *S. marcescens* públicos utilizados na construção do pan-genoma.

Nome	ID	Nome	ID
Db11	NZ_HG326223.1	UMH12	NZ_CP018930.1
WW4	NC_020211.1	SMB2099	NZ_HG738868.1
FDAARGOS_65	NZ_CP026050.1	1274	NZ_CP019927.1
SM39	NZ_AP013063.1	S217	NZ_CP021984.1
CAV1492	NZ_CP011642.1	AS1	NZ_CP010584.1
RSC-14	NZ_CP012639.1	MSU97	GCA_001902635.1
SmUNAM836	NZ_CP012685.1	ATCC14041	GCA_000695485.1
B3R3	NZ_CP013046.2	ATCC13880	GCA_000735445.1
UMH1	NZ_CP018915.1	VGH107	GCA_000342205.1
UMH2	NZ_CP018924.1	EGD-HP20	GCA_002265665.1
UMH3	NZ_CP018925.1	CDC_813-60	GCA_000743395.1
UMH5	NZ_CP018917.1	LCT-262	GCA_000442375.1
UMH6	NZ_CP018926.1	LCT-213	GCA_000264275.1
UMH7	NZ_CP018919.1	LCT-166	GCA_000442455.1
UMH8	NZ_CP018927.1	SM39	NZ_AP013063.1
UMH9	NZ_CP018923.1	FS14	NZ_CP005927.1
UMH10	NZ_CP018928.1	JKS296	GCA_900215445.1
UMH11	NZ_CP018929.1		

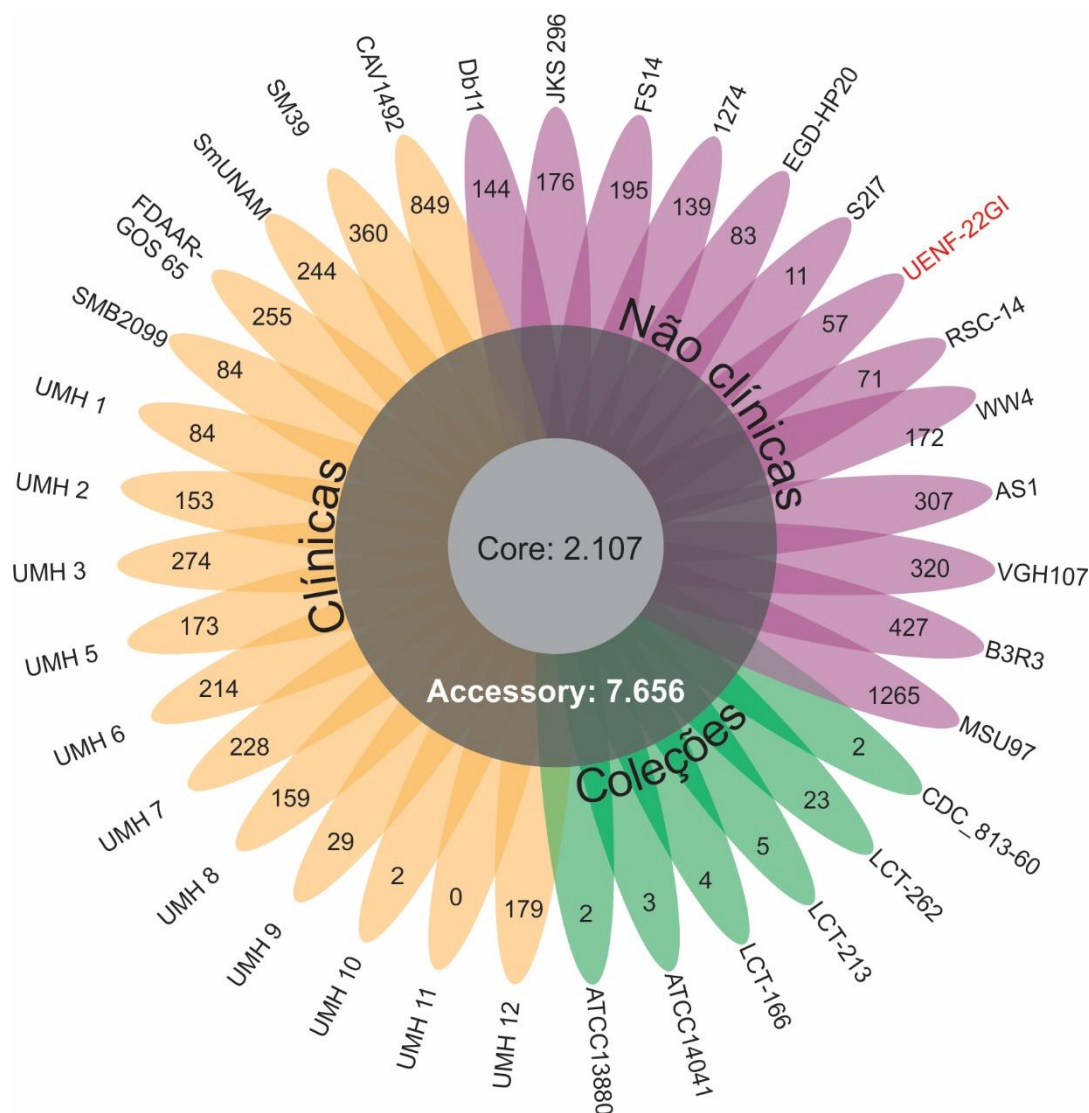


Figura 17. Pan-genoma de *Serratia marcescens*. Em verde isolados pertencentes a coleções de bactérias, em laranja cepas clínicas e em roxo isolados não clínicos, inclusive SMU, grafada em vermelho, que possui 57 genes únicos.

5.11 Análise filogenética de SMU

Para a reconstrução filogenética de *S. marcescens* foram utilizadas as sequências de 10 proteínas codificadas por genes cópia-única presentes no core genoma, que estão também presentes como cópia-única no banco de dados de referência de bactérias do BUSCO (Simao *et al.*, 2015). Utilizamos todas as cepas presentes no pan-genoma e *S. plymuthica* AS9 como grupo externo.

A análise filogenética obtida permite inferir conclusões importantes sobre as relações intra-espécie. A árvore filogenética se divide em dois grandes ramos (Figura 18 seta vermelha): um dominado por cepas clínicas com alguns isolados não-clínicos (AS1, 1274, RSC-14 e Db11) e; um segundo grupo exclusivamente

com isolados não-clínicos, com exceção de UMH8. SMU se encontra nesse grande grupo de isolados não-clínicos, estando mais próxima de *Serratia* sp. JKS296. O mapa de calor na Figura 18 representa a similaridade do genoma inteiro computada por dDDH. Pode-se perceber que, em linhas gerais, a similaridade do genoma acompanha os achados obtidos na filogenia. Os valores baixos de bootstraps nos ramos terminais são indicativos da alta similaridade entre isolados no respectivo clado.

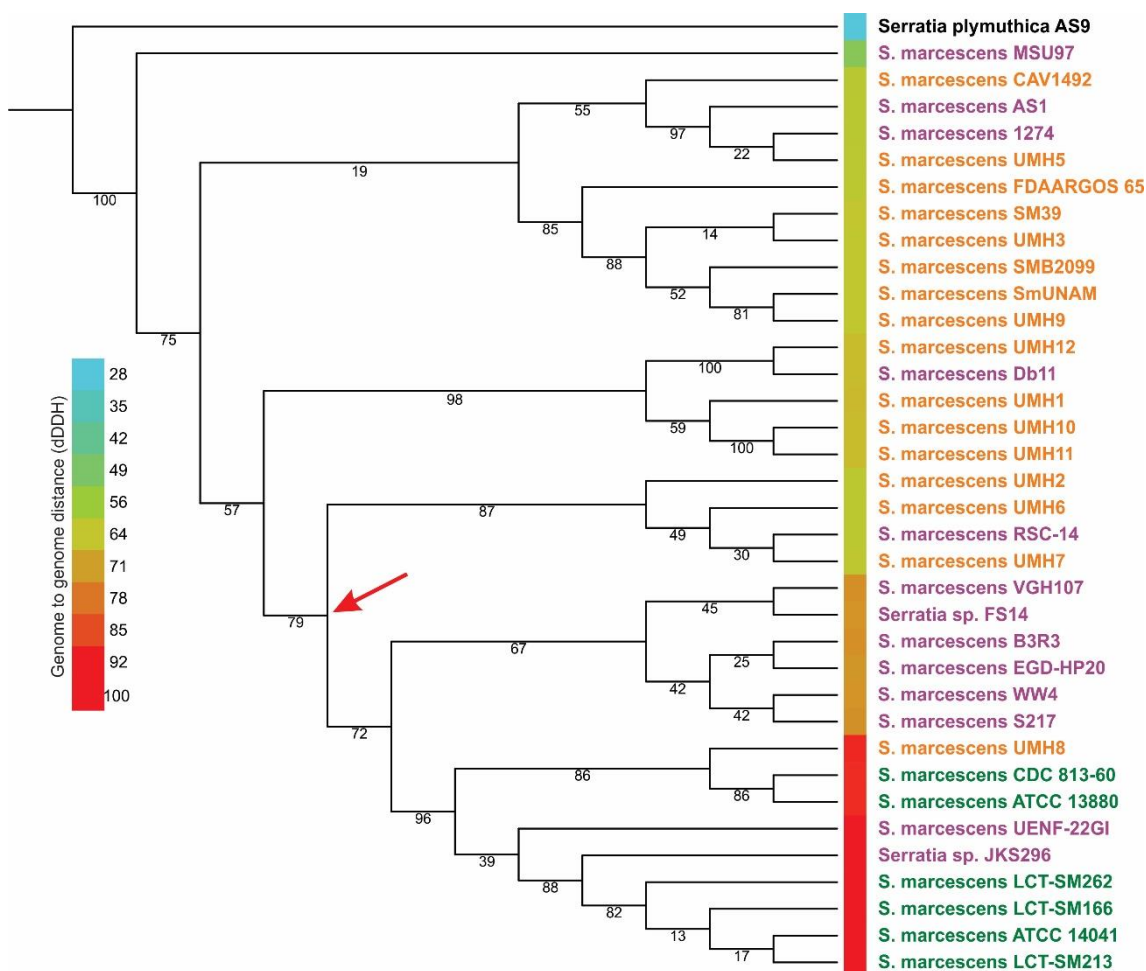


Figura 18. Árvore filogenética multilocus computada por máxima verossimilhança dos 35 isolados de *S. marcescens*. As sequências das 10 proteínas foram obtidas e alinhadas separadamente. Posteriormente os alinhamentos foram concatenados e a filogenia inferida utilizando-se RAxML (Stamatakis, 2014) com suporte estatístico de bootstrap de 1000 vezes representado em porcentagem. Em laranja isolados clínicos, em verde, pertencentes à coleções e em roxo não clínicos.

5.12 Correlação dos achados in vitro com genes presentes em SMU

Sabendo dos potenciais mecanismos de promoção do crescimento empregados por SMU, buscamos anotar genes específicos que desempenhem

papéis centrais nos fenótipos observados. Os resultados estão sumarizados na Tabela 8 e estão agrupados de acordo com o processo biológico. Nas próximas seções abordaremos cada um dos grupos destacados da referida tabela enfatizando as funções dos genes no contexto PGPR.

Tabela 8. Genes relacionados a características de promoção crescimento vegetal em SMU.

Solubilização de P e Zn	
ID do gene (AK961_)	Nome da proteína
07090, 07095, 07100, 07105, 07110	pqqB, pqqC, pqqD, pqqE, pqqF
10840	(PQQ)-dependente glicose desidrogenase
17880	Gluconolactonase
08395	2-gluconate desidrogenase
17580, 17575, 17570	2-keto-gluconato desidrogenase
21125, 21130, 21135, 21140, 21145	phoU, pstB, pstA, pstC, pstS
Tolerância contra metais tóxicos	
01055	Arsenate redutase
03990	ArsR regulador transcricional do operon arsRBC
03995	arsB proteína transmembrana de efluxo
04000	arsC1 arsenato redutase
01905, 01915	Proteína de resistência a cobre
09290	Proteína de resistência a cobre copD
07435, 07440	Transportador de cromato
12615	Cobalto-zinco-cádmio efflux system
Relacionados à IAA e espermidina	
01575	ipdC
00655, 12310	Proteínas de efluxo de auxina
18130, 18125	speAB
18275, 18270	speDE
Formação de biofilme	
20475, 20470, 20465, 20460, 20480, 20485	bcsA, bcsB, bcsC, bcsZ
20490, 20495, 20500	bcsQ, bcsR
01650, 01655, 1660, 01665	bcsE, bcsF, bcsG
13115	pgaA, pgaB, pgaC, pgaD
	adrA
Biocontrole e competição	
20530, 01270, 12935, 05475	chiA, chiB, chiD, chiA1
13300, 13305, 13310, 13315,	pigA, pigB, pigC, pigD, pigE, pigF, pigG, pigH, pigI, pigJ, pigK
13320, 13325, 13330, 13335,	pigL, pigM, pigN
13340, 13345, 13350, 13355,	
13360, 13365	
16235	Proteína de resistência a Kasugamicina ksgA
03740, 14040	Proteína de resistência a biciclomicina
02590	Sistema de efluxo de biciclomicina
13395	Proteína de resistência a Fosmidomicina
07380	Inibidor de Barnase
08005	Proteína de resistência a ácido fusárico

Um traço típico da estrutura genômica procariótica, os operons, geralmente compreendem genes associados a uma função específica (e.g. componentes de uma via metabólica ou complexo proteico). A organização

peculiar facilita a resposta adaptativa do procarionto às condições do meio (Land *et al.*, 2015), pois permite a expressão simultânea de todos os genes localizados adjacente na mesma fita, além disso todos os genes pertencentes a um operon estão sob controle de um mesmo promotor, o qual atua como sítio de reconhecimento da maquinaria transcricional (Ralston, 2008). Na Figura 19 estão representados alguns dos operons importantes que foram anotados manualmente como sendo correlacionados com o perfil de promoção do crescimento vegetal proporcionado por SMU.

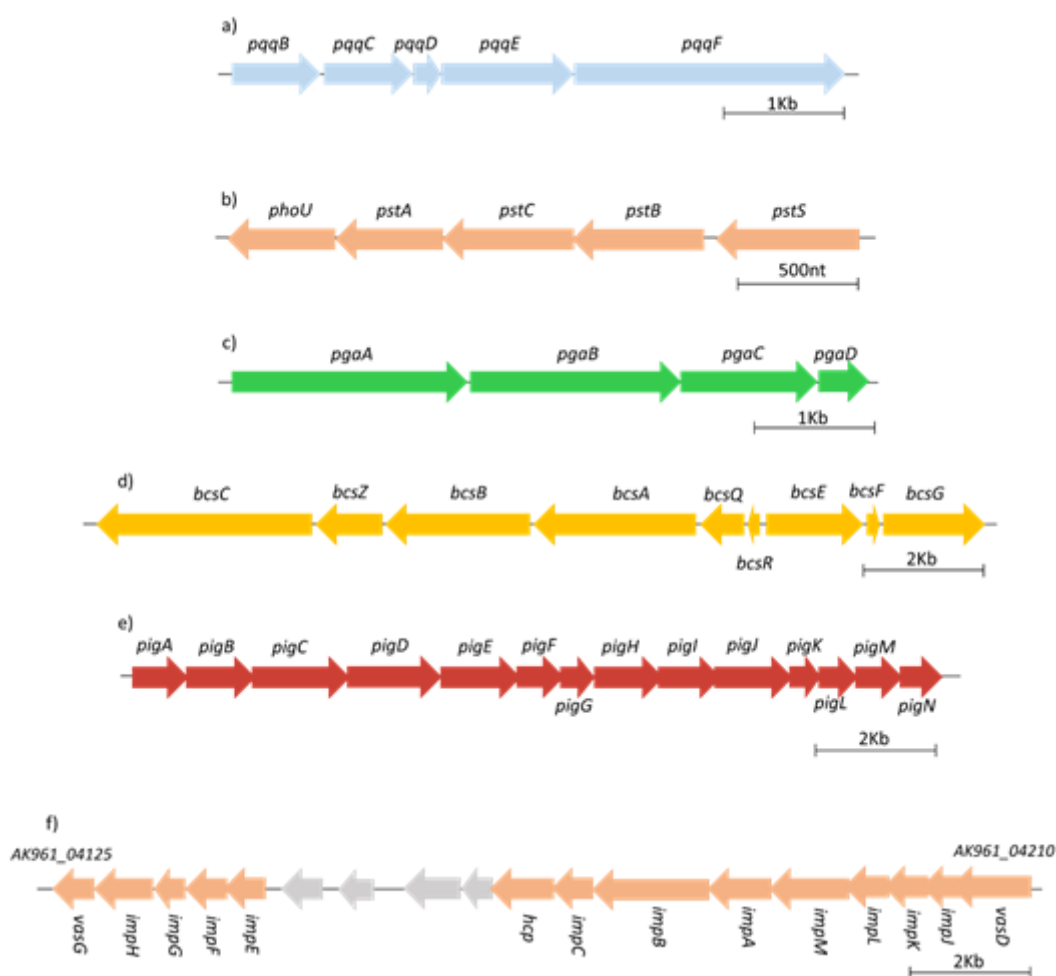


Figura 19. Operons correlacionados com funções PGPR presentes no genoma de SMU. a) Biossíntese de PQQ. b) Transportador Pst específico para fosfato. c) e d) Produção de biofilme, respectivamente poli-N-acetilglicosamina e celulose. e) Prodiginosina. f) Sistema de secreção tipo 6.

5.12.1 Genes relacionados a solubilização de P e Zn

Em ambientes tropicais, a principal forma de P são minerais fosfatados pouco solúveis. Esses compostos não estão acessíveis diretamente pelos

vegetais (An e Moe, 2016). Assim, a conversão microbiana de formas insolúveis de íons orgânicos solúveis ($H_2PO_4^-$) é um mecanismo essencial para aumentar a disponibilidade de P no solo (Alori *et al.*, 2017).

A produção e secreção de ácidos foi descrita como sendo uma das principais estratégias de solubilização, não apenas de P (An e Moe, 2016), mas também Zn (Solanki *et al.*, 2016). Dentre essas substâncias destaca-se o ácido glucônico, que é produzido por três reações oxidativas e secretado por proteínas periplasmáticas ligadas à membrana (Krishnaraj e Goldstein, 2001).

O genoma de SMU possui vários genes relacionados à produção de ácido glucônico, que se inicia na oxidação da glicose por uma glicose desidrogenase pirroquinolina-quinona (PQQ) dependente ligada à membrana periplasmática (AK961_10840). O produto intermediário glucono-1,5-lactona é hidrolisado a gluconato pela gluconalactonase (AK961_17880) e oxidado pela 2-gluconato-desidrogenase (AK961_08395) a 2-cetogluconato, que é oxidado a 2,5-diceto gluconato pela 2-ceto-gluconato desidrogenase, um complexo enzimático com as seguintes subunidades codificadas no mesmo operon: menor (AK961_17580), maior (AK961_17575) e citocromo (AK961_17570). Em um modelo semelhante ao descrito para *Gluconobacter oxydans* (AB985494).

A síntese do ácido glucônico é dependente do cofator PQQ (Duine, 1991), que é tipicamente produzido por proteínas codificadas pelo operon *pqqBCDEF*, que está totalmente conservado no genoma de SMU (AK961_07090, AK961_07095, AK961_07100, AK961_07105 e AK961_07110, respectivamente (Tabela 8 e Figura 19a) . O genoma de SMU também apresenta o operon *pstABCS* (AK961_21130, AK961_21135, AK961_21140, AK961_21145) (Tabela 8 e Figura 19b), que codifica um sistema específico para transporte de fosfato.

O Zn é outro nutriente essencial para o desenvolvimento da planta (Choudhary *et al.*, 2017), outro trabalho indica que a solubilização deste micronutriente é realizada pelos mesmos genes envolvidos na solubilização de P inorgânico (Intorne *et al.*, 2009), essencialmente pela acidificação do meio. Desta forma, nossos dados genômicos corroboram as evidências *in vitro* (Figura 13), indicando que SMU solubiliza P e Zn através da acidificação do solo.

5.12.2 Tolerância a metais tóxicos

Bactérias encontram no solo uma série de condições desfavoráveis,

sendo uma delas a contaminação por metais. Para lidar com esse tipo de condição, diferentes estratégias podem ser empregadas, como por exemplo através de reações de redução, oxidação, metilação ou conversão dos metais pesado em formas menos tóxicas (Das *et al.*, 2016). Algumas PGPRs podem atenuar o impacto negativo destes metais nas plantas (Hassan *et al.*, 2017). O genoma de SMU possui vários genes relacionados à essas funções (Tabela 8): arsenato redutase (AK961_01055), *arsRBC* (AK961_03990, AK961_03995, AK961_04000), proteína de resistência a cobre (AK961_01905, AK961_01915, AK961_09290), *cusRS* (AK961_11430, AK961_11425), transportador de cromato ChrA (AK961_07435, AK961_07440), cromato redutase (AK961_21085), resistência a cobalto-cádmio-zinco *czcD* (AK961_12615). Esses resultado estão de acordo com outro genoma de *S. marcescens* que se mostrou capaz de atenuar o estresse por cádmio em plantas (Khan *et al.*, 2017).

5.12.3 IAA e espermidina

O gene *ipdC* foi encontrado em SMU (Tabela 8). Esse gene codifica a enzima responsável pela conversão do indol-3-piruvato em indol-3-acetaldeído, uma etapa crucial da via IPyA (Figura 14a). Em um outro trabalho foi reportado que o knockout do gene *ipdC* diminui drasticamente a produção de IAA em *A. brasilense* (Malhotra e Srivastava, 2008). Além disso, também identificamos duas proteínas de efluxo de auxina (AK961_00655, AK961_12310), sugerindo que SMU é capaz de exportar IAA.

Poliaminas (e.g. espermidinas) são essenciais para viabilidade celular e são correlacionadas com o desenvolvimento de raízes laterais, resistência a patógenos e atenuação de estresses oxidativos, osmóticos e acídicos (Xie *et al.*, 2014). Sendo assim, a produção de espermidina por SMU possivelmente contribui também para a promoção do crescimento vegetal.

5.12.4 Formação de biofilme

Biofilmes bacterianos são comunidades multicelulares envoltas por uma matriz extracelular polimérica (Flemming *et al.*, 2016), que são essenciais para sobrevivência, interação célula-célula e colonização da raiz (Kasim *et al.*, 2016). No genoma de SMU encontramos diversos genes relacionados à produção de biofilme (Tabela 8), como *pgaABCD* (AK961_01650, AK961_01655, AK961_01660, AK961_01665, respectivamente) (Figura 19c). Esse operon é

responsável pela produção de Poli- β -1,6-N-Acetil-D-Glucosamina (PGA), que está associada com adesão a superfície, adesão intercelular e formação de biofilme em diversas espécies (Echeverz *et al.*, 2017).

Outro polissacarídeo fundamental para biofilmes em *Enterobacteriaceae* é a celulose. A biossíntese de celulose já foi amplamente descrita em bactéria e há grande diversidade de operons de síntese de celulose (Römling e Galperin, 2015). Em *proteobacteria* a biossíntese de celulose é feita principalmente pelos operons *bcsABZC* e *bcsEFG*, em associação aos genes *bcsQ* e *bcsR* (Tabela 8), assim como descrito para *E. coli* (Krasteva *et al.*, 2017). O operon *bcsABZC* (AK961_20475, AK961_20470, AK961_20465, AK961_20460) e *bcsEFG* (AK961_20490, AK961_20495, AK961_20500) estão proximais um do outro no genoma de SMU, em fitas opostas. Localizado upstream ao operon *bcsABZC* estão os genes regulatórios *bcsQ* (AK961_20480) e *bcsR* (AK961_20485) (Figura 19d) . Além disso, também encontramos o gene *adrA* (AK961_13115), que codifica uma molécula de sinalização que se liga a BcsA ativando a síntese de celulose (Cowles *et al.*, 2016). Dessa forma propomos que PGA e celulose são constituintes do biofilme produzido por SMU e que esse biofilme está relacionado com o biocontrole de patógenos fúngicos (Figure 15c, f) e colonização da raiz.

5.12.5 Biocontrole e competição

Quitinas são polissacarídeos complexos abundantes na natureza, que são acessados por bactérias através da atividade de enzimas hidrolíticas (quitinases), fornecendo uma importante fonte de carboidratos. Além, disso a parede celular de fungos contém quitina, tornando-os alvos em potencial de quitinases bacterianas (Paspaliari *et al.*, 2017). Tendo em vista o impacto de doenças fúngicas na agricultura, as quitinases têm recebido cada vez mais atenção como mecanismo de biocontrole pela comunidade científica (Santamarina *et al.*, 2017). Em acordo com estas observações, *S. marcescens* já foi descrita como sendo uma bactéria produtora de quitinases (Vaikuntapu *et al.*, 2016).

Encontramos no genoma de SMU 4 quitinases (Tabela 8), e as classificamos através de análises com Blastp contra a base de dados Swissprot. AK961_20530 possui 99 % de identidade com chitinase A (P07254),

AK961_01270 possui 100 % de identidade com quitinase B (P11797), AK961_12935 possui 29 % identidade e 93 % cobertura com quitinase D (P27050) e AK961_05475 possui 31 % identidade e 88 % cobertura com quitinase A1 (P20533). Também encontramos outros genes provavelmente envolvidos no metabolismo de quitina: AK961_01260 e AK961_12890, que codificam uma proteína de transporte e quitobiase, respectivamente.

Outro operon encontrado no genoma de SMU, *pig* (Tabela 8), codifica o pigmento vermelho observado geralmente em isolados ambientais de *S. marcescens*, a prodigiosina (Mahlen, 2011). Em SMU, este operon tem 14 genes (Figura 19e), em uma estrutura muito semelhante à encontrada em *S. marcescens* ATCC274 (outro isolado ambiental) (Harris *et al.*, 2004). Já foi reportado que a prodigiosina possui atividade capaz de restringir o crescimento de vários fungos (Duzhak *et al.*, 2012), bactérias (Danevčič *et al.*, 2016), protozoários (Genes *et al.*, 2011) e até mesmo vírus (Zhou *et al.*, 2016). O mecanismo de ação foi recentemente sugerido como sendo relacionado à afinidade da prodigiosina pela membrana plasmática, o que causaria danos à membrana externa (Darshan e Manonmani, 2016). Contudo, os detalhes bioquímicos deste processo permanecem desconhecidos

SMU também possui um gene que confere resistência a ácido fusárico (AK961_08005), uma micotoxina capaz de agir contra outros microrganismos, como *P. fluorescens* (Crutcher *et al.*, 2017). A presença desse gene pode conferir uma vantagem indireta na atividade fungicida de SMU frente a *Fusarium*.

Espécies de *Streptomyces* são ubíquas no solo e são notáveis por produzir ampla gama de antibióticos (Barka *et al.*, 2016). Sendo assim a presença de genes conferindo resistência a esses compostos é um traço interessante para PGPR. O genoma de SMU possui algumas proteínas de resistência contra antimicrobianos de *Streptomyces* como: bicyclomicina (AK961_03740, AK961_14040, AK961_02590), fosmidomicina (AK961_13395) e kasugamicina (AK961_16235). Além disso SMU possui Sistema de secreção do tipo 6 (T6SS) (AK961_04125-AK961_04210), descrito como capaz de mediar interações antagonísticas contra outras bactérias (Russell *et al.*, 2014).

5.13 Análise de transferência horizontal

Com o intuito de comparar o genoma com os genomas mais próximos

encontrados em nossas análises filogenéticas (Figura 18), realizamos o alinhamento global dos genomas, o que nos permitiu inferir a existência de uma região do genoma de SMU que era pouco conservada em outras cepas próximas, porém presente em JSK296 e ATCC14041. Esta região foi denominada Gap1, essa possui 38 genes e 52,198 bp (Figura 20a). Nossos dados revelam que essa região é uma ilha genômica resultante de transferência horizontal mediada por um fago. A maior parte dos genes dentro da região do Gap1 estão relacionados à produção de um peptídeo não-ribossomal antimicrobiano. Anotamos domínios relacionados à síntese desses peptídeos nos genes (AK961_03500, AK961_03515, AK961_03520 e AK961_03545). O gene AK961_03560 parece codificar um mecanismo de detoxificação, provavelmente o mecanismo anti-toxina que o protege dos efeitos do peptídeo não ribossomal. O gene AK961_03610 codifica uma integrase de origem fágica, que delimita a região transferida horizontalmente (Figura 20b).

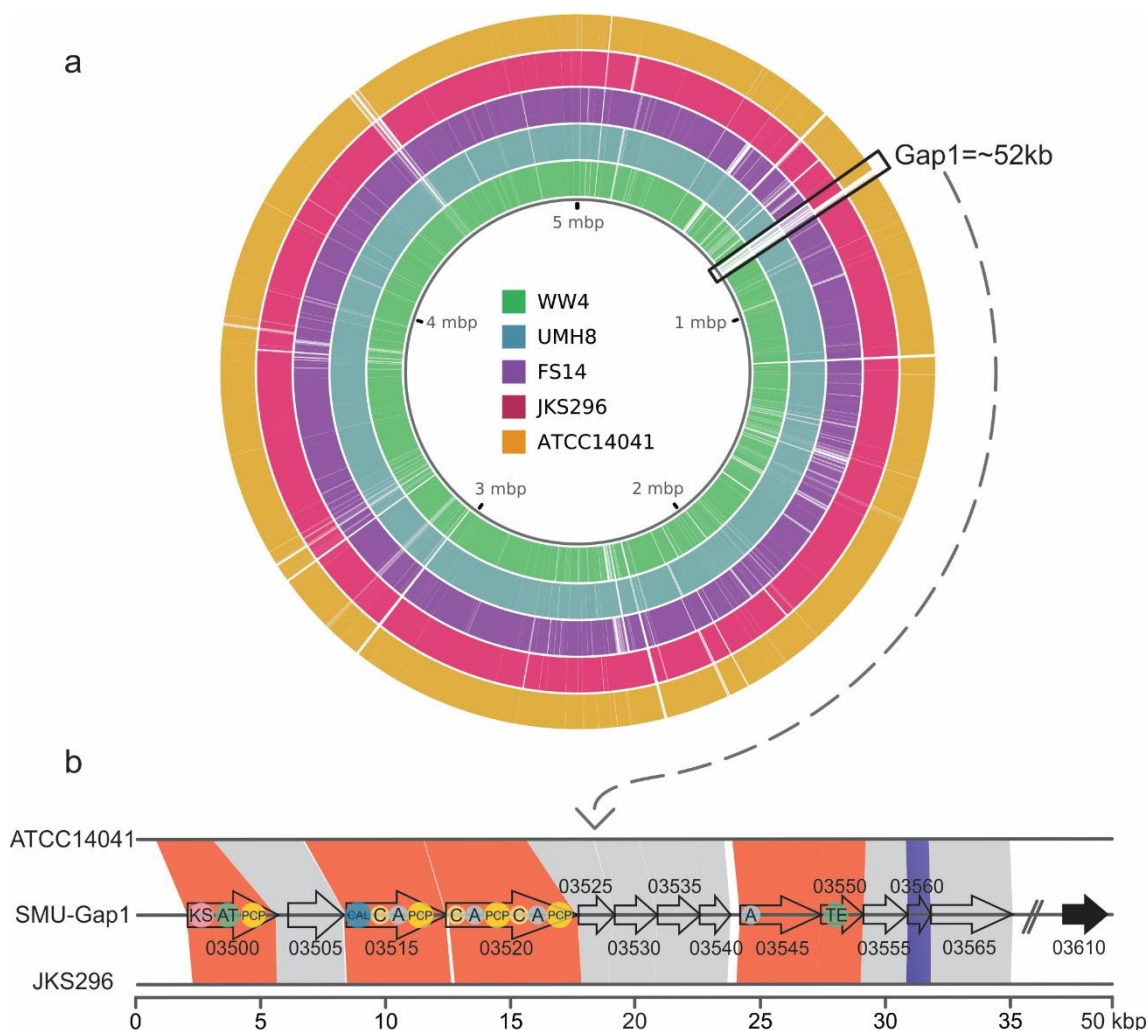


Figura 20. Análise de transferência horizontal. a) Alinhamento do genoma de SMU com genomas de referência. b) Sintonia da região do Gap1 mostrando os genes associados à produção do peptídeo não ribossomal e seus respectivos domínios.

5.14 Inferência taxonômica por ANI em *Enterobacter* sp.

A filogenia obtida pela sequência do RNA ribossomal 16S não foi conclusiva para nossos isolados 6GII, 14GII, 21GI, 23GI, 24GI e 25GI. Sendo assim averiguamos inicialmente a ANI entre nossos isolado e todos os genomas depositados para *Enterobacter* sp. (1.357). A partir da matriz de indêntidade obtida (Figura 21) filtramos os isolados com ANI \geq 95 %, o patamar que tem sido considerado para se considerar isolados da mesma espécie (Richter e Rosselló-Móra, 2009). Entretanto baseado em nossas análises prévias esses valores carecem de confirmação e serão posteriormente confirmados utilizando dDDH.

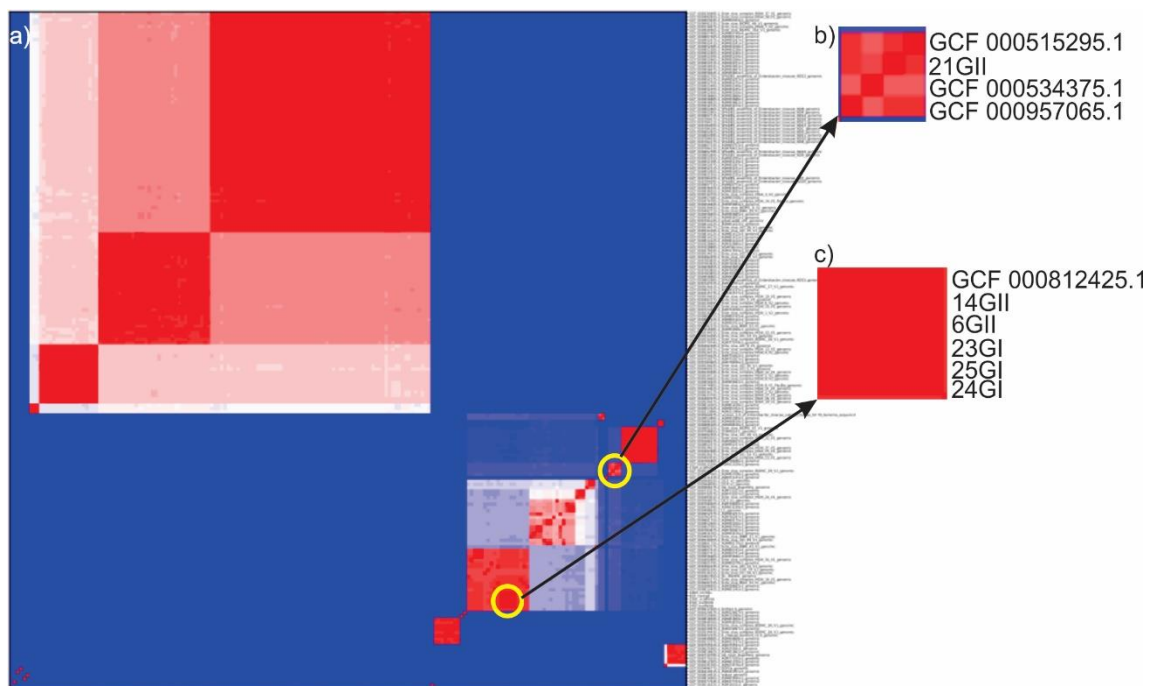


Figura 21. Matriz de ANI representativa. Ilustra a metodologia aplicada para obtenção de isolados filogeneticamente mais próximos das amostras pré-classificadas como *Enterobacter* spp.. Os 1.357 genomas depositados foram divididos em 6 devido a questões computacionais. a) Output de 235 genomas, os círculos em amarelo destacam clusters nos quais nossos isolados estão inseridos. b) Cluster de 21GII. c) Cluster de 6GII, 14GII, 23GI, 24GI e 25GI.

5.15 Testes *in vivo*

Almejando aferir a capacidade dos isolados 6GII, 14GII, 21GII, 23GI, 24GI e 25GI em promover o crescimento vegetal, utilizamos o mesmo sistema realizado para 22GI utilizando plântulas de milho. Os resultados estão dispostos na Figura 22 e evidenciam o potencial desses isolados em modular positivamente o crescimento da planta, o que pode ser constatado através de significativo aumento das medidas obtidas.

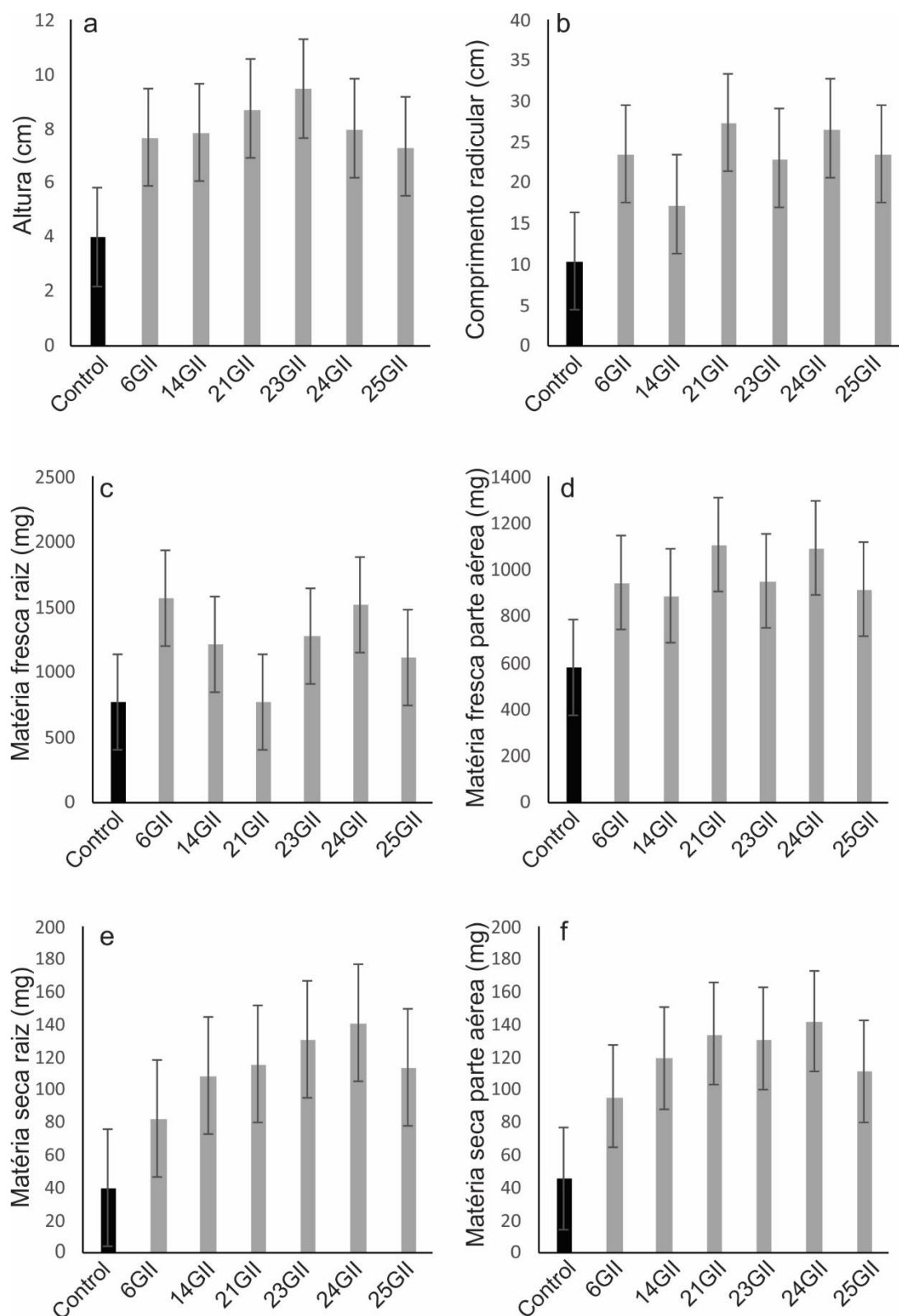


Figura 22. Promoção do crescimento de plântulas de milho pelos isolados 6GII, 14GII, 21GII, 23GI, 24GI e 25GI. Plântulas germinadas (com comprimento radicular de 2 a 2,5 cm) foram transferidas para todos de vidro contendo vermiculita estéril (uma plântula por tubo). A inoculação foi feita através da aplicação de 1 mL de suspensão bacteriana (10^8 células·mL⁻¹) sobre as plântulas. a) Altura. b) Comprimento radicular. c) Massa fresca da raiz. d) Massa fresca parte aérea. e) Matéria seca da raiz. f) Massa seca parte aérea. Sistemas inoculados com 1 mL de meio Dygs estéril foram usados como controle negativos, as medições foram feitas após 10 dias. Os experimentos foram realizados em triplicata.

6. CONSIDERAÇÕES FINAIS

Nossos resultados geraram ampla quantidade de dados, a partir dos quais obtivemos o genoma de um isolado de *S. marcescens*, denominado UENF-22GI. Traçamos um paralelo consistente entre o genoma de SMU e outros genomas públicos e aprofundamos a filogenia intra-espécie de *S. marcescens* conseguindo obter um grau de diferenciação entre isolados clínicos e não-clínicos.

Analizamos extensivamente o genoma dessa bactéria e catalogamos a presença de diversos genes e suas funções no contexto PGPR. Esses genes PGPR de SMU são capazes de promover o crescimento vegetal tanto diretamente quanto indiretamente e nós validamos essa capacidade através de testes *in vitro* e *in vivo*.

7. CONCLUSÕES

1. SMU traz em seu genoma genes capazes de promover de forma direta e indireta o crescimento vegetal.
2. Existem no genoma de SMU diversas proteínas de resistência a antimicrobianos e biocontrole, abrindo a possibilidade do estudo mais aprofundado dessa cepa contra outros patógenos
3. Há no genoma de SMU uma região decorrente de transferência horizontal relacionada à produção de um peptídeo não ribossomal.
4. *Serratia marcescens* UENF-22GI é uma bactéria encontrada em material vermicompostado que é capaz de promover o crescimento vegetal.

8. REFERÊNCIAS

- AGUIAR, K. P. (2012). Prospecção De Bactérias Promotoras Do Crescimento Vegetal Associadas a Vermicompostos., Dissertação de mestrado, Produção Vegetal - UENF. Campos dos Goitacazes., v., n., p. 86.
- AGUIAR, N. O. et al. (2012). Bioactivity of humic acids isolated from vermicomposts at different maturation stages, Plant and soil, v. 362, n. 1-2, p. 161-174.
- AHMED, E.; S. J. HOLMSTRÖM (2014). Siderophores in environmental research: roles and applications, Microbial biotechnology, v. 7, n. 3, p. 196-208.
- ALBERS, C. N. et al. (2008). Effect of different humic substances on the fate of diuron and its main metabolite 3, 4-dichloroaniline in soil, Environmental science & technology, v. 42, n. 23, p. 8687-8691.
- ALIKHAN, N.-F. et al. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons, BMC genomics, v. 12, n. 1, p. 402.
- ALKARAWI, H. H.; G. ZOTZ (2014). Phytic acid in green leaves of herbaceous plants—temporal variation in situ and response to different nitrogen/phosphorus fertilizing regimes, AoB Plants, v. 6, n.
- ALORI, E. T.; B. R. GLICK; O. O. BABALOLA (2017). Microbial Phosphorus Solubilization and Its Potential for Use in Sustainable Agriculture, Frontiers in microbiology, v. 8, n.
- ALTSCHUL, S. F. et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic acids research, v. 25, n. 17, p. 3389-3402.
- AMALRAJ, E. L. D. et al. (2015). Potential Microbial Consortium for Plant Growth Promotion of Sunflower (*Helianthus annuus* L.), Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, v. 85, n. 2, p. 635-642.
- AN, R.; L. A. MOE (2016). Regulation of Pyrroloquinoline Quinone-Dependent Glucose Dehydrogenase Activity in the Model Rhizosphere-Dwelling Bacterium *Pseudomonas putida* KT2440, Appl Environ Microbiol, v. 82, n. 16, p. 4955-4964.
- ANDREWS, S. (2010). FastQC: a quality control tool for high throughput sequence data.
- ANTIPOV, D. et al. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data, bioRxiv, v., n., p. 048942.
- ANWER, M. A. (2017). Flourescent Pseudomonads: A Potential Bio-control Agents Against Plant Diseases. Biopesticides and Bioagents, Apple Academic Press: 93-114.
- ARCAND, M. M.; K. D. SCHNEIDER (2006). Plant-and microbial-based

- mechanisms to improve the agronomic effectiveness of phosphate rock: a review, Anais da Academia Brasileira de Ciências, v. 78, n. 4, p. 791-807.
- ARORA, N. K.; M. VERMA; J. MISHRA (2017). Rhizobial bioformulations: Past, Present and Future. Rhizotrophs: Plant Growth Promotion to Bioremediation, Springer: 69-99.
- ASHRAF, M. A. et al. (2017). Advances in microbe-assisted reclamation of heavy metal contaminated soils over the last decade: a review, Journal of environmental management, v. 198, n., p. 132-143.
- AUCH, A. F.; H.-P. KLENK; M. GÖKER (2010). Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs, Standards in Genomic Sciences, v. 2, n. 1, p. 142.
- BACA, B.; C. ELMERICH (2007). Microbial production of plants hormones by microorganisms, Associative nitrogenfixation bacteria and cyanobacteria. Series: nitrogen fixation: origins, applications, and research progress, v. 4, n., p. 11.
- BAKKER, P. A.; C. M. PIETERSE; L. VAN LOON (2007). Induced systemic resistance by fluorescent *Pseudomonas* spp, Phytopathology, v. 97, n. 2, p. 239-243.
- BALDANI, J. et al. (1986). Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium, Int J Syst Evol Microbiol, v. 36, n. 1, p. 86-93.
- BALDOTTO, L. E. B.; F. L. OLIVARES (2008). Phylloepiphytic interaction between bacteria and different plant species in a tropical agricultural system, Canadian Journal of Microbiology, v. 54, n. 11, p. 918-931.
- BANKEVICH, A. et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, Journal of computational biology, v. 19, n. 5, p. 455-477.
- BARKA, E. A. et al. (2016). Taxonomy, physiology, and natural products of Actinobacteria, Microbiology and Molecular Biology Reviews, v. 80, n. 1, p. 1-43.
- BARKER, A. V.; D. J. PILBEAM (2015). Handbook of plant nutrition, CRC press.
- BARTLING, D. et al. (1992). Cloning and expression of an Arabidopsis nitrilase which can convert indole-3-acetonitrile to the plant hormone, indole-3-acetic acid, The FEBS Journal, v. 205, n. 1, p. 417-424.
- BASHAN, Y. (1998). Azospirillum plant growth-promoting strains are nonpathogenic on tomato, pepper, cotton, and wheat, Canadian Journal of Microbiology, v. 44, n. 2, p. 168-174.
- BASHAN, Y. et al. (2014). Advances in plant growth-promoting bacterial

inoculant technology: formulations and practical perspectives (1998–2013), Plant and soil, v. 378, n. 1-2, p. 1-33.

BASHAN, Y.; H. LEVANONY (1988). Adsorption of the rhizosphere bacterium *Azospirillum brasilense* Cd to soil, sand and peat particles, Microbiology, v. 134, n. 7, p. 1811-1820.

BATZOGLOU, S. et al. (2002). ARACHNE: a whole-genome shotgun assembler, Genome Res, v. 12, n. 1, p. 177-189.

BEIER, S.; S. BERTILSSON (2013). Bacterial chitin degradation—mechanisms and ecophysiological strategies, Frontiers in microbiology, v. 4, n., p. 149.

BELGAROU, N. et al. (2016). The secretion of the bacterial phytase PHY-US417 by *Arabidopsis* roots reveals its potential for increasing phosphate acquisition and biomass production during co-growth, Plant biotechnology journal, v. 14, n. 9, p. 1914-1924.

BERG, G. et al. (2014). Unraveling the plant microbiome: looking back and future perspectives, v., n.

BERG, G. et al. (2013). Next-generation bio-products sowing the seeds of success for sustainable agriculture, Multidisciplinary Digital Publishing Institute.

BERTALAN, M. et al. (2009). Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5, BMC genomics, v. 10, n. 1, p. 450.

BERTELLI, C. et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets, Nucleic acids research, v. 45, n. W1, p. W30-W35.

BESEMER, J.; A. LOMSADZE; M. BORODOVSKY (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions, Nucleic acids research, v. 29, n. 12, p. 2607-2618.

BIBRA, M.; R. N. KRISHNARAJ; R. K. SANI (2017). An Overview on Extremophilic Chitinases. Extremophilic Enzymatic Processing of Lignocellulosic Feedstocks to Bioenergy, Springer: 225-247.

BLEIDORN, C. (2016). Third generation sequencing: technology and its potential impact on evolutionary biodiversity research, Systematics and biodiversity, v. 14, n. 1, p. 1-8.

BLIN, K. et al. (2017). antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification, Nucleic acids research, v. 45, n. W1, p. W36-W41.

BODIRSKY, B. L. et al. (2014). Reactive nitrogen requirements to feed the world in 2050 and potential to mitigate nitrogen pollution, Nat Commun, v. 5, n., p. 3858.

- BODIRSKY, B. L. et al. (2015). Global food demand scenarios for the 21 st century, Plos one, v. 10, n. 11, p. e0139201.
- BOETZER, M. et al. (2011). Scaffolding pre-assembled contigs using SSPACE, Bioinformatics, v. 27, n. 4, p. 578-579.
- BOLGER, A. M.; M. LOHSE; B. USADEL (2014). Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics, v., n., p. btu170.
- BRESLER, M. A. et al. (2012). Telescope: de novo assembly of highly repetitive regions, Bioinformatics, v. 28, n. 18, p. i311-i317.
- BRINCH-PEDERSEN, H.; L. D. SØRENSEN; P. B. HOLM (2002). Engineering crop plants: getting a handle on phosphate, Trends in plant science, v. 7, n. 3, p. 118-125.
- BUZA, K.; B. WILCZYNSKI; N. DOJER (2015). RECORD: reference-assisted genome assembly for closely related genomes, International journal of genomics, v. 2015, n.
- ÇAKMAKÇI, R. et al. (2017). The Role of Soil Beneficial Bacteria in Wheat Production: A Review. Wheat Improvement, Management and Utilization, InTech.
- CANELLAS, L. P. et al. (2002). Humic acids isolated from earthworm compost enhance root elongation, lateral root emergence, and plasma membrane H⁺-ATPase activity in maize roots, Plant physiology, v. 130, n. 4, p. 1951-1957.
- CARATTOLI, A. et al. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing, Antimicrobial agents and chemotherapy, v. 58, n. 7, p. 3895-3903.
- CASTRESANA, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, Mol Biol Evol, v. 17, n. 4, p. 540-552.
- CASTRO-GUERRERO, J. et al. (2012). The hisC1 gene, encoding aromatic amino acid aminotransferase-1 in *Azospirillum brasilense* Sp7, expressed in wheat, Plant and soil, v. 356, n. 1-2, p. 139-150.
- CASTRO, R. A. et al. (2014). Isolation and enzyme bioprospection of endophytic bacteria associated with plants of Brazilian mangrove ecosystem, SpringerPlus, v. 3, n. 1, p. 382.
- CATTANI, I. et al. (2009). The role of natural purified humic acids in modifying mercury accessibility in water and soil, Journal of environmental quality, v. 38, n. 2, p. 493-501.
- CELI, L.; E. BARBERIS (2007). Abiotic Reactions of Inositol Phosphates in Soil Inositol Phosphates: Linking Agriculture and the Environment: 207-220.
- CHAISSON, M. J.; D. BRINZA; P. A. PEVZNER (2009). De novo fragment

assembly with short mate-paired reads: Does the read length matter?, Genome Res, v. 19, n. 2, p. 336-346.

CHANG, P. et al. (2014). Plant growth-promoting bacteria facilitate the growth of barley and oats in salt-impacted soil: implications for phytoremediation of saline soils, International journal of phytoremediation, v. 16, n. 11, p. 1133-1147.

CHAUDHARI, N. M.; V. K. GUPTA; C. DUTTA (2016). BPGA- an ultra-fast pan-genome analysis pipeline, Sci Rep, v. 6, n., p. 24373.

CHEN, Y. et al. (2015). Earthworms modify microbial community structure and accelerate maize stover decomposition during vermicomposting, Environmental Science and Pollution Research, v. 22, n. 21, p. 17161-17170.

CHIBUIKE, G. U.; S. C. OBIORA (2014). Heavy metal polluted soils: effect on plants and bioremediation methods, Applied and Environmental Soil Science, v. 2014, n.

CHOUDHARY, D. K.; A. VARMA; N. TUTEJA (2017). Plant-microbe Interaction: an Approach to Sustainable Agriculture, Springer.

CLARK, S. C. et al. (2013). ALE: a generic assembly likelihood evaluation framework for assessing the accuracy of genome and metagenome assemblies, Bioinformatics, v. 29, n. 4, p. 435-443.

COMBES-MEYNET, E. et al. (2011). The Pseudomonas secondary metabolite 2, 4-diacetylphloroglucinol is a signal inducing rhizoplane expression of Azospirillum genes involved in plant-growth promotion, Molecular plant-microbe interactions, v. 24, n. 2, p. 271-284.

COMPANT, S.; C. CLÉMENT; A. SESSITSCH (2010). Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization, Soil Biology and Biochemistry, v. 42, n. 5, p. 669-678.

COMPEAU, P. E.; P. A. PEVZNER; G. TESLER (2011). How to apply de Bruijn graphs to genome assembly, Nat Biotechnol, v. 29, n. 11, p. 987-991.

CONN, V. M.; C. M. FRANCO (2004). Effect of microbial inoculants on the indigenous actinobacterial endophyte population in the roots of wheat as determined by terminal restriction fragment length polymorphism, Appl Environ Microbiol, v. 70, n. 11, p. 6407-6413.

CONSORTIUM, U. (2016). UniProt: the universal protein knowledgebase, Nucleic acids research, v. 45, n. D1, p. D158-D169.

CORDELL, D.; J.-O. DRANGERT; S. WHITE (2009). The story of phosphorus: global food security and food for thought, Global environmental change, v. 19, n. 2, p. 292-305.

CORDELL, D.; S. WHITE (2015). Tracking phosphorus security: indicators of

phosphorus vulnerability in the global food system, Food Security, v. 7, n. 2, p. 337-350.

COWLES, K. N. et al. (2016). Diguanilate cyclases AdrA and STM1987 regulate *Salmonella enterica* exopolysaccharide production during plant colonization in an environment-dependent manner, Applied and environmental microbiology, v. 82, n. 4, p. 1237-1248.

CROWLEY, D. E. (2006). Microbial siderophores in the plant rhizosphere. Iron nutrition in plants and rhizospheric microorganisms, Springer: 169-198.

CRUTCHER, F. K. et al. (2017). Microbial Resistance Mechanisms to the Antibiotic and Phytotoxin Fusaric Acid, Journal of chemical ecology, v. 43, n. 10, p. 996-1006.

DA CONCEIÇÃO, P. M. et al. (2008). Recobrimento de sementes de milho com ácidos húmicos e bactérias diazotróficas endofíticas, Pesquisa agropecuária brasileira, v. 43, n. 4, p. 545-548.

DANEVČIČ, T. et al. (2016). Prodigiosin induces autolysins in actively grown *Bacillus subtilis* cells, Frontiers in microbiology, v. 7, n.

DARRIBA, D. et al. (2011). ProtTest 3: fast selection of best-fit models of protein evolution, Bioinformatics, v. 27, n. 8, p. 1164-1165.

DARSHAN, N.; H. MANONMANI (2016). Prodigiosin inhibits motility and activates bacterial cell death revealing molecular biomarkers of programmed cell death, AMB Express, v. 6, n. 1, p. 50.

DAS, S.; H. R. DASH; J. CHAKRABORTY (2016). Genetic basis and importance of metal resistant genes in bacteria for bioremediation of contaminated environments with toxic metal pollutants, Applied microbiology and biotechnology, v. 100, n. 7, p. 2967-2984.

DATTA, A. et al. (2015). An effective and beneficial plant growth promoting soil bacterium “*Rhizobium*”: a review, Annals of Plant Sciences, v. 4, n. 01, p. 933-942.

DAVLETOVA, S. et al. (2005). The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in *Arabidopsis*, Plant physiology, v. 139, n. 2, p. 847-856.

DE HOON, M. J. et al. (2004). Predicting the operon structure of *Bacillus subtilis* using operon length, intergene distance, and gene expression information. Pacific Symposium on Biocomputing.

DEAKER, R.; R. J. ROUGHLEY; I. R. KENNEDY (2004). Legume seed inoculation technology—a review, Soil Biology and Biochemistry, v. 36, n. 8, p. 1275-1288.

DEEBA, F. et al. (2016). Chitinase production in organisms: a review, Punjab Univ. J. Zool, v. 31, n. 1, p. 101-106.

DEHAL, P. S. et al. (2010). MicrobesOnline: an integrated portal for comparative and functional genomics, Nucleic acids research, v. 38, n. suppl 1, p. D396-D400.

DELONGE, M. et al. (2013). The impact of changing moisture conditions on short-term P availability in weathered soils, Plant and soil, v. 365, n. 1-2, p. 201-209.

DEMPSTER, A. P.; N. M. LAIRD; D. B. RUBIN (1977). Maximum likelihood from incomplete data via the EM algorithm, Journal of the royal statistical society. Series B (methodological), v., n., p. 1-38.

DINESH, R. et al. (2015). Isolation, characterization, and evaluation of multi-trait plant growth promoting rhizobacteria for their growth promoting and disease suppressing effects on ginger, Microbiological research, v. 173, n., p. 34-43.

DOBBELAERE, S.; J. VANDERLEYDEN; Y. OKON (2003). Plant Growth-Promoting Effects of Diazotrophs in the Rhizosphere, Critical Reviews in Plant Sciences, v. 22, n. 2, p. 107-149.

DÖBEREINER, J.; V. L. C. D. BALDANI; J. I. BALDANI (1995). Como isolar e identificar bactérias diazotróficas de plantas não-leguminosas, Embrapa Agrobiologia.

DÖBEREINER, J.; J. DAY (1976). Associative symbioses in tropical grasses: characterization of microorganisms and dinitrogen-fixing sites. Proceedings of the 1st international symposium on nitrogen fixation, Washington State University Press Pullman.

DOBEREINER, J.; S. URQUIAGA; R. M. BODDEY (1995). Alternatives for nitrogen nutrition of crops in tropical agriculture. Nitrogen Economy in Tropical Soils, Springer: 338-346.

DONOSO, R. et al. (2017). Biochemical and Genetic Bases of Indole-3-Acetic Acid (Auxin Phytohormone) Degradation by the Plant-Growth-Promoting Rhizobacterium *Paraburkholderia phytofirmans* PsJN, Appl Environ Microbiol, v. 83, n. 1, p. e01991-01916.

DÖRR, J.; T. HUREK; B. REINHOLD-HUREK (1998). Type IV pili are involved in plant–microbe and fungus–microbe interactions, Mol Microbiol, v. 30, n. 1, p. 7-17.

DUAN, M. et al. (2016). Effects of sulphamethazine and zinc on the functional diversity of microbial communities during composting, Environmental technology, v. 37, n. 11, p. 1357-1368.

DUCA, D. et al. (2014). Indole-3-acetic acid in plant–microbe interactions, Antonie Van Leeuwenhoek, v. 106, n. 1, p. 85-125.

DUINE, J. A. (1991). Quinoproteins: enzymes containing the quinonoid cofactor pyrroloquinoline quinone, topaquinone or tryptophan-tryptophan quinone, The FEBS Journal, v. 200, n. 2, p. 271-284.

DUTTA, J.; P. J. HANDIQUE; D. THAKUR (2015). Assessment of culturable tea rhizobacteria isolated from tea estates of Assam, India for growth promotion in commercial tea cultivars, Front Microbiol, v. 6, n.

DUZHAK, A. et al. (2012). Role of prodigiosin and chitinases in antagonistic activity of the bacterium *Serratia marcescens* against the fungus *Didymella applanata*, Biochemistry (Moscow), v. 77, n. 8, p. 910-916.

EARL, D. et al. (2011). Assemblathon 1: a competitive assessment of de novo short read assembly methods, Genome Res, v. 21, n. 12, p. 2224-2241.

ECHEVERZ, M. et al. (2017). Lack of the PGA exopolysaccharide in *Salmonella* as an adaptive trait for survival in the host, PLoS genetics, v. 13, n. 5, p. e1006816.

EDDY, S. R. (1998). Profile hidden Markov models, Bioinformatics, v. 14, n. 9, p. 755-763.

EDGAR, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic acids research, v. 32, n. 5, p. 1792-1797.

EDGAR, R. C. (2010). Search and clustering orders of magnitude faster than BLAST, Bioinformatics, v. 26, n. 19, p. 2460-2461.

EGHOLM, M. et al. (2005). Genome sequencing in open microfabricated high density picoliter reactors, Nature, v. 437, n., p. 376-380.

EISEN, J. A. (1995). The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species, Journal of molecular evolution, v. 41, n. 6, p. 1105-1123.

EKBLOM, R.; J. B. WOLF (2014). A field guide to whole-genome sequencing, assembly and annotation, Evol Appl, v. 7, n. 9, p. 1026-1042.

EL-SAYED, W. S. et al. (2014). In vitro antagonistic activity, plant growth promoting traits and phylogenetic affiliation of rhizobacteria associated with wild plants grown in arid soil, Front Microbiol, v. 5, n., p. 651.

ELMERICH, C. (2015). One Hundred Years Discovery of Nitrogen-Fixing Rhizobacteria. Biological Nitrogen Fixation, John Wiley & Sons, Inc: 897-912.

EMAMVERDIAN, A. et al. (2015). Heavy metal stress and some mechanisms of plant defense response, The Scientific World Journal, v. 2015, n.

EWING, B. et al. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment, Genome Res, v. 8, n. 3, p. 175-185.

FARRER, R. A. (2017). Synima: a Synteny imaging tool for annotated genome assemblies, BMC bioinformatics, v. 18, n. 1, p. 507.

FELSENSTEIN, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach, Journal of molecular evolution, v. 17, n. 6, p. 368-376.

FIBACH-PALDI, S.; S. BURDMAN; Y. OKON (2012). Key physiological properties contributing to rhizosphere adaptation and plant growth promotion abilities of *Azospirillum brasilense*, FEMS microbiology letters, v. 326, n. 2, p. 99-108.

FINN, R. D. et al. (2013). Pfam: the protein families database, Nucleic acids research, v., n., p. gkt1223.

FLEISCHMANN, R. D. et al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, science, v. 269, n. 5223, p. 496.

FLEMMING, H.-C. et al. (2016). Biofilms: an emergent form of bacterial life, Nature Reviews Microbiology, v. 14, n. 9, p. 563.

GALLANT, J.; D. MAIER; J. ASTORER (1980). On finding minimal length superstrings, Journal of Computer and System Sciences, v. 20, n. 1, p. 50-58.

GALLANT, J. K. (1983). The complexity of the overlap method for sequencing biopolymers, Journal of Theoretical Biology, v. 101, n. 1, p. 1-17.

GENES, C. et al. (2011). Mitochondrial dysfunction in *Trypanosoma cruzi*: the role of *Serratia marcescens* prodigiosin in the alternative treatment of Chagas disease, Parasites & vectors, v. 4, n. 1, p. 66.

GERKE, J. (2010). Humic (organic matter)-Al (Fe)-phosphate complexes: an underestimated phosphate form in soils and source of plant-available phosphate, Soil Science, v. 175, n. 9, p. 417-425.

GERKE, J. (2015). Phytate (inositol hexakisphosphate) in soil and phosphate acquisition from inositol phosphates by higher plants. A review, Plants, v. 4, n. 2, p. 253-266.

GINGERAS, T. et al. (1979). Computer programs for the assembly of DNA sequences, Nucleic acids research, v. 7, n. 2, p. 529-543.

GLADMAN, S.; T. SEEMANN (2012). VelvetOptimiser, Victorian Bioinformatics Consortium, Clayton, Australia: <http://bioinformatics.net.au/software.velvetoptimiser.shtml>, v., n.

GOSWAMI, D.; J. N. THAKKER; P. C. DHANDHUKIA (2016). Portraying mechanics of plant growth promoting rhizobacteria (PGPR): A review, Cogent Food & Agriculture, v. 2, n. 1, p. 1127500.

GREEN, S. J. et al. (1999). PHRAP documentation, v., n.

GROSSMANN, K. (2010). Auxin herbicides: current status of mechanism and mode of action, Pest management science, v. 66, n. 2, p. 113-120.

GULANI, C.; S. BHATTACHARYA; A. DAS (2012). Assessment of process

parameters influencing the enhanced production of prodigiosin from *Serratia marcescens* and evaluation of its antimicrobial, antioxidant and dyeing potentials, Malays J Microbiol, v. 8, n. 2, p. 116-122.

GUPTA, G. et al. (2015). Plant growth promoting rhizobacteria (PGPR): current and future prospects for development of sustainable agriculture, J Microb Biochem Technol, v. 7, n. 2, p. 096-102.

GUREVICH, A. et al. (2013). QUASt: quality assessment tool for genome assemblies, Bioinformatics, v. 29, n. 8, p. 1072-1075.

HAAS, D.; G. DÉFAGO (2005). Biological control of soil-borne pathogens by fluorescent pseudomonads, Nature Reviews Microbiology, v. 3, n. 4, p. 307.

HAFEEZ, B.; Y. KHANIF; M. SALEEM (2013). Role of zinc in plant nutrition-a review, American journal of experimental Agriculture, v. 3, n. 2, p. 374.

HAKEEM, K. R.; M. S. AKHTAR (2016). *Plant, Soil and Microbes: Volume 2: Mechanisms and Molecular Interactions*, Springer.

HAMID, R. et al. (2013). Chitinases: an update, Journal of pharmacy & bioallied sciences, v. 5, n. 1, p. 21.

HANC, A.; M. DRESLOVA (2016). Effect of composting and vermicomposting on properties of particle size fractions, Bioresour Technol, v. 217, n., p. 186-189.

HARRIS, A. K. et al. (2004). The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species- and strain-dependent genome context variation, Microbiology, v. 150, n. 11, p. 3547-3560.

HARTMANN, A. et al. (2009). Plant-driven selection of microbes, Plant and soil, v. 321, n. 1-2, p. 235-257.

HASSAN, T. U.; A. BANO; I. NAZ (2017). Alleviation of heavy metals toxicity by the application of plant growth promoting rhizobacteria and effects on wheat grown in saline sodic field, International journal of phytoremediation, v. 19, n. 6, p. 522-529.

HAYASHI, S.; P. M. GRESSHOFF; B. J. FERGUSON (2014). Mechanistic action of gibberellins in legume nodulation, Journal of integrative plant biology, v. 56, n. 10, p. 971-978.

HENZ, S. R. et al. (2004). Whole-genome prokaryotic phylogeny, Bioinformatics, v. 21, n. 10, p. 2329-2335.

HOU, W. et al. (2013). Extracellular polymeric substances from copper-tolerance *Sinorhizobium meliloti* immobilize Cu²⁺, Journal of hazardous materials, v. 261, n., p. 614-620.

HUANG, K. et al. (2013). Changes of bacterial and fungal community compositions during vermicomposting of vegetable wastes by *Eisenia foetida*,

Bioresour Technol, v. 150, n., p. 235-241.

HUANG, Q.; L. BAUM; W.-L. FU (2010). Simple and practical staining of DNA with GelRed in agarose gel electrophoresis, Clinical Laboratory Journal For Clinical Laboratories And Laboratories Related, v. 56, n. 3, p. 149.

HUANG, X.; S. P. YANG (2005). Generating a genome assembly with PCAP, Current Protocols in Bioinformatics, v., n., p. 11.13. 11-11.13. 23.

HUNGRIA, M.; I. C. MENDES (2015). Nitrogen fixation with soybean: the perfect symbiosis?, Biological nitrogen fixation, v., n., p. 1009-1024.

HUNKAPILLER, T.; R. KAISER (1991). Large-scale and automated DNA sequence determination, science, v. 254, n. 5028, p. 59.

HUNT, M. et al. (2013). REAPR: a universal tool for genome assembly evaluation, Genome Biol, v. 14, n. 5, p. R47.

IDURY, R. M.; M. S. WATERMAN (1995). A new algorithm for DNA sequence assembly, Journal of computational biology, v. 2, n. 2, p. 291-306.

IGUCHI, A. et al. (2014). Genome evolution and plasticity of *Serratia marcescens*, an important multidrug-resistant nosocomial pathogen, Genome biology and evolution, v. 6, n. 8, p. 2096-2110.

IMAM, J. (2016). Microbial interactions in plants: Perspectives and applications of proteomics, Current Protein & Peptide Science, v. 18, n.

INTORNE, A. C. et al. (2009). Identification and characterization of *Gluconacetobacter diazotrophicus* mutants defective in the solubilization of phosphorus and zinc, Archives of microbiology, v. 191, n. 5, p. 477-483.

IZAGUIRRE-MAYORAL, M. L.; C. LABANDERAY; J. SANJUÁN (2007). **Biofertilizantes en Iberoamérica: una visión técnica, científica y empresarial**, CYTED, BIOFAG.

JAFFE, D. B. et al. (2003). Whole-genome sequence assembly for mammalian genomes: Arachne 2, Genome Res, v. 13, n. 1, p. 91-96.

JAMESON, P. (2000). Cytokinins and auxins in plant-pathogen interactions—An overview, Plant Growth Regulation, v. 32, n. 2-3, p. 369-380.

JANDA, J. M.; S. L. ABBOTT (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls, J Clin Microbiol, v. 45, n. 9, p. 2761-2764.

JIJON-MORENO, S. et al. (2015). The ipdC, hisC1 and hisC2 genes involved in indole-3-acetic production used as alternative phylogenetic markers in *Azospirillum brasilense*, Antonie Van Leeuwenhoek, v. 107, n. 6, p. 1501-1517.

KAHLON, R. S. (2016). *Pseudomonas: Molecular and Applied Biology*, Springer.

- KANEHISA, M. et al. (2011). KEGG for integration and interpretation of large-scale molecular data sets, Nucleic acids research, v., n., p. gkr988.
- KANEHISA, M.; Y. SATO; K. MORISHIMA (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences, Journal of molecular biology, v. 428, n. 4, p. 726-731.
- KAO, W.-C.; K. STEVENS; Y. S. SONG (2009). BayesCall: A model-based base-calling algorithm for high-throughput short-read sequencing, Genome Res, v. 19, n. 10, p. 1884-1895.
- KASIM, W. A. et al. (2016). Effect of biofilm forming plant growth promoting rhizobacteria on salinity tolerance in barley, Annals of Agricultural Sciences, v. 61, n. 2, p. 217-227.
- KECECIOGLU, J. D.; E. W. MYERS (1995). Combinatorial algorithms for DNA sequence assembly, Algorithmica, v. 13, n. 1-2, p. 7.
- KELLEY, D. R.; S. L. SALZBERG (2010). Detection and correction of false segmental duplications caused by genome mis-assembly, Genome biology, v. 11, n. 3, p. R28.
- KELLEY, K.; F. STEVENSON (1995). Forms and nature of organic N in soil. Nitrogen Economy in Tropical Soils, Springer: 1-11.
- KHAN, A. L. et al. (2016). Indole acetic acid and ACC deaminase from endophytic bacteria improves the growth of *Solanum lycopersicum*, Electronic Journal of Biotechnology, v. 21, n., p. 58-64.
- KHAN, A. R. et al. (2017). Complete genome analysis of *Serratia marcescens* RSC-14: A plant growth-promoting bacterium that alleviates cadmium stress in host plants, Plos one, v. 12, n. 2, p. e0171534.
- KINGSFORD, C.; M. C. SCHATZ; M. POP (2010). Assembly complexity of prokaryotic genomes using short reads, BMC bioinformatics, v. 11, n. 1, p. 21.
- KLOEPPER, J.; M. SCHROTH (1981). Plant growth-promoting rhizobacteria and plant growth under gnotobiotic conditions, Phytopathology, v. 71, n. 6, p. 642-644.
- KOBAYASHI, M. et al. (1995). Occurrence of enzymes involved in biosynthesis of indole-3-acetic acid from indole-3-acetonitrile in plant-associated bacteria, *Agrobacterium* and *Rhizobium*, Proceedings of the National Academy of Sciences, v. 92, n. 3, p. 714-718.
- KOCHAR, M.; A. UPADHYAY; S. SRIVASTAVA (2011). Indole-3-acetic acid biosynthesis in the biocontrol strain *Pseudomonas fluorescens* Psd and plant growth regulation by hormone overexpression, Research in microbiology, v. 162, n. 4, p. 426-435.
- KRASTEVA, P. V. et al. (2017). Insights into the structure and assembly of a bacterial cellulose secretion system, Nature communications, v. 8, n. 1, p. 2065.

KRAUSE, A. et al. (2006). Complete genome of the mutualistic, N₂-fixing grass endophyte *Azoarcus* sp. strain BH72, Nat Biotechnol, v. 24, n. 11.

KRISHNARAJ, P.; A. GOLDSTEIN (2001). Cloning of a *Serratia marcescens* DNA fragment that induces quinoprotein glucose dehydrogenase-mediated gluconic acid production in *Escherichia coli* in the presence of stationary phase *Serratia marcescens*, FEMS microbiology letters, v. 205, n. 2, p. 215-220.

KUMAR, A. et al. (2018). Plant Growth-Promoting Rhizobacteria (PGPR): Perspective in Agriculture Under Biotic and Abiotic Stress. Crop Improvement Through Microbial Biotechnology, Elsevier: 333-342.

LAND, M. et al. (2015). Insights from 20 years of bacterial genome sequencing, Functional & integrative genomics, v. 15, n. 2, p. 141-161.

LANDER, E. S. et al. (2001). Initial sequencing and analysis of the human genome, Nature, v. 409, n. 6822, p. 860-921.

LANDER, E. S.; M. S. WATERMAN (1988). Genomic mapping by fingerprinting random clones: a mathematical analysis, Genomics, v. 2, n. 3, p. 231-239.

LANG, J. M.; A. E. DARLING; J. A. EISEN (2013). Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices, Plos one, v. 8, n. 4, p. e62510.

LEI, X. G. et al. (2007). Phytase: source, structure and application. Industrial enzymes, Springer: 505-529.

LEI, X. G. et al. (2013). Phytase, a new life for an “old” enzyme, Annu. Rev. Anim. Biosci., v. 1, n. 1, p. 283-309.

LETUNIC, I.; P. BORK (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees, Nucleic acids research, v. 44, n. W1, p. W242-W245.

LEVEAU, J. H. (2007). The magic and menace of metagenomics: prospects for the study of plant growth-promoting rhizobacteria, European Journal of Plant Pathology, v. 119, n. 3, p. 279-300.

LI, P. P.; X. J. WANG; Z. J. CUI (2012a). Survival and performance of two cellulose-degrading microbial systems inoculated into wheat straw-amended soil, Journal of Microbiology and Biotechnology, v. 22, n. 1, p. 126-132.

LI, Y. I.; R. R. COPLEY (2012). Scaffolding low quality genomes using orthologous protein sequences, Bioinformatics, v. 29, n. 2, p. 160-165.

LI, Z. et al. (2012b). Comparison of the two major classes of assembly algorithms: overlap–layout–consensus and de-bruijn-graph, Briefings in functional genomics, v. 11, n. 1, p. 25-37.

LIM, S. L. et al. (2015). The use of vermicompost in organic farming: overview, effects on soil and economics, J Sci Food Agric, v. 95, n. 6, p. 1143-1156.

LINDOW, S. E.; M. T. BRANDL (2003). Microbiology of the phyllosphere, Appl Environ Microbiol, v. 69, n. 4, p. 1875-1883.

LIU, L. et al. (2012). Comparison of next-generation sequencing systems, BioMed Research International, v. 2012, n.

LIU, Y.; X. PAN; J. LI (2015). A 1961–2010 record of fertilizer use, pesticide application and cereal yields: a review, Agronomy for sustainable development, v. 35, n. 1, p. 83-93.

LODEWYCKX, C. et al. (2002). Endophytic bacteria and their potential applications, Critical Reviews in Plant Sciences, v. 21, n. 6, p. 583-606.

LONG, Y. et al. (2015). Arabidopsis BIRD zinc finger proteins jointly stabilize tissue boundaries by confining the cell fate regulator SHORT-ROOT and contributing to fate specification, The Plant Cell, v. 27, n. 4, p. 1185-1199.

LUO, W.; B. GU (2008). Dissolution and mobilization of uranium in a reduced sediment by natural humic substances under anaerobic conditions, Environmental science & technology, v. 43, n. 1, p. 152-156.

MACDONALD, G. K. et al. (2011). Agronomic phosphorus imbalances across the world's croplands, Proceedings of the National Academy of Sciences, v. 108, n. 7, p. 3086-3091.

MADISON, J. D. et al. (2017). Characterization of Batrachochytrium dendrobatidis inhibiting bacteria from amphibian populations in Costa Rica, Frontiers in microbiology, v. 8, n., p. 290.

MAHESHWARI, D. K.; S. DHEEMAN; M. AGARWAL (2015). Phytohormone-producing PGPR for sustainable agriculture. Bacterial metabolites in sustainable agroecosystem, Springer: 159-182.

MAHLEN, S. D. (2011). Serratia infections: from military experiments to current practice, Clinical microbiology reviews, v. 24, n. 4, p. 755-791.

MAHMOOD, Q. et al. (2013). Natural treatment systems as sustainable ecotechnologies for the developing countries, Biomed Res Int, v. 2013, n.

MALHOTRA, M.; S. SRIVASTAVA (2008). An ipdC gene knock-out of Azospirillum brasilense strain SM and its implications on indole-3-acetic acid biosynthesis and plant growth promotion, Antonie Van Leeuwenhoek, v. 93, n. 4, p. 425-433.

MAO, F. et al. (2009). DOOR: a database for prokaryotic operons, Nucleic acids research, v. 37, n. suppl 1, p. D459-D463.

MAO, X. et al. (2014). DOOR 2.0: presenting operons and their functions through dynamic and integrated views, Nucleic acids research, v. 42, n. D1, p. D654-D659.

MARRA, L. M. et al. (2015). Initial pH of medium affects organic acids

- production but do not affect phosphate solubilization, Brazilian Journal of Microbiology, v. 46, n. 2, p. 367-375.
- MASTRETTA, C. et al. (2006). Endophytic bacteria and their potential application to improve the phytoremediation of contaminated environments, Biotechnology and genetic engineering reviews, v. 23, n. 1, p. 175-188.
- MCKENZIE, F. C.; J. WILLIAMS (2015). Sustainable food production: constraints, challenges and choices by 2050, Food Security, v. 7, n. 2, p. 221-233.
- MEDINI, D. et al. (2005). The microbial pan-genome, Current opinion in genetics & development, v. 15, n. 6, p. 589-594.
- MEENA, R. K. et al. (2015). Isolation of low temperature surviving plant growth-promoting rhizobacteria (PGPR) from pea (*Pisum sativum* L.) and documentation of their plant growth promoting traits, Biocatalysis and Agricultural Biotechnology, v. 4, n. 4, p. 806-811.
- MEHRA, P.; B. K. PANDEY; J. GIRI (2017). Improvement in phosphate acquisition and utilization by a secretory purple acid phosphatase (OsPAP21b) in rice, Plant biotechnology journal, v., n.
- MENZEL, P.; K. L. NG; A. KROGH (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nature communications, v. 7, n.
- MERINO, E.; R. A. JENSEN; C. YANOFSKY (2008). Evolution of bacterial trp operons and their regulation, Curr Opin Microbiol, v. 11, n. 2, p. 78-86.
- METZKER, M. L. (2010). Sequencing technologies—the next generation, Nature reviews genetics, v. 11, n. 1, p. 31.
- MILLER, C. S. et al. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data, Genome Biol, v. 12, n. 5, p. R44.
- MILLER, J. R.; S. KOREN; G. SUTTON (2010). Assembly algorithms for next-generation sequencing data, Genomics, v. 95, n. 6, p. 315-327.
- MINKIN, I. et al. (2013). Sibelia: a scalable and comprehensive synteny block generation tool for closely related microbial genomes. International Workshop on Algorithms in Bioinformatics, Springer.
- MOLE, B. M. et al. (2007). Global virulence regulation networks in phytopathogenic bacteria, Trends in microbiology, v. 15, n. 8, p. 363-371.
- MORADIGARAVAND, D. et al. (2016). Recent independent emergence of multiple multidrug-resistant *Serratia marcescens* clones within the United Kingdom and Ireland, Genome research, v. 26, n. 8, p. 1101-1109.
- MORRISSEY, J. P. et al. (2002). Exploitation of genetically modified inoculants for industrial ecology applications, Antonie Van Leeuwenhoek, v. 81, n. 1, p.

599-606.

MOUNT, M. (2012). *Phytopathogenic prokaryotes*, Elsevier.

MUDGE, S. R.; F. W. SMITH; A. E. RICHARDSON (2003). Root-specific and phosphate-regulated expression of phytase under the control of a phosphate transporter promoter enables *Arabidopsis* to grow on phytate as a sole P source, *Plant Science*, v. 165, n. 4, p. 871-878.

MURPHY, J.; J. P. RILEY (1962). A modified single solution method for the determination of phosphate in natural waters, *Analytica chimica acta*, v. 27, n., p. 31-36.

MYERS, E. W. et al. (2000). A whole-genome assembly of *Drosophila*, *science*, v. 287, n. 5461, p. 2196-2204.

NAGARAJAN, N.; M. POP (2013). Sequence assembly demystified, *Nature reviews genetics*, v. 14, n. 3, p. 157-167.

NANJAREDDY, K. et al. (2014). Nitrate regulates rhizobial and mycorrhizal symbiosis in common bean (*Phaseolus vulgaris* L.), *Journal of integrative plant biology*, v. 56, n. 3, p. 281-298.

NCBI, R. C. (2016). Database resources of the National Center for Biotechnology Information, *Nucleic acids research*, v. 44, n. D1, p. D7.

NELSON, E. B. (2004). Microbial dynamics and interactions in the spermosphere, *Annu. Rev. Phytopathol.*, v. 42, n., p. 271-309.

NESME, J. et al. (2016). Back to the future of soil metagenomics, *Front Microbiol*, v. 7, n.

NOBBE, F.; L. HILTNER (1896). Inoculation of the soil for cultivating leguminous plants.

O'LEARY, N. A. et al. (2015). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation, *Nucleic acids research*, v. 44, n. D1, p. D733-D745.

OERKE, E.-C. (2006). Crop losses to pests, *The Journal of Agricultural Science*, v. 144, n. 1, p. 31-43.

OLIVARES, F. (2009). Bactérias promotoras de crescimento vegetal, *Boletim Informativo da*, v., n.

OLIVARES, F. L. et al. (2015). Substrate biofortification in combination with foliar sprays of plant growth promoting bacteria and humic substances boosts production of organic tomatoes, *Scientia Horticulturae*, v. 183, n., p. 100-108.

ONDOV, B. D.; N. H. BERGMAN; A. M. PHILLIPPY (2011). Interactive metagenomic visualization in a Web browser, *BMC bioinformatics*, v. 12, n. 1, p. 385.

ONGENA, M.; G. HENRY; P. THONART (2010). The roles of cyclic lipopeptides in the biocontrol activity of *Bacillus subtilis*. Recent developments in management of plant diseases, Springer: 59-69.

OWEN, D. et al. (2015). Use of commercial bio-inoculants to increase agricultural production through improved phosphorous acquisition, Applied Soil Ecology, v. 86, n., p. 41-54.

PĂCURAR, D. I. et al. (2011). *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation, Physiological and molecular plant pathology, v. 76, n. 2, p. 76-81.

PANDEY, A. (2009). *Bacillus subtilis* NRRL B-30408 inoculation enhances the symbiotic efficiency of *Lens esculenta* Moench at a Himalayan location, Journal of Plant Nutrition and Soil Science, v. 172, n. 1, p. 134-139.

PANKIEVICZ, V. et al. (2015). Robust biological nitrogen fixation in a model grass–bacterial association, The Plant Journal, v. 81, n. 6, p. 907-919.

PARADELO, R. et al. (2012). Plant tests for determining the suitability of grape marc composts as components of plant growth media, Waste Management & Research, v. 30, n. 10, p. 1059-1065.

PASPALIARI, D. K. et al. (2017). Chitinase expression in *Listeria monocytogenes* is influenced by Imo0327, which encodes an internalin-like protein, Applied and environmental microbiology, v. 83, n. 22, p. e01283-01217.

PATTEN, C. L.; B. R. GLICK (2002). Role of *Pseudomonas putida* indoleacetic acid in development of the host plant root system, Applied and environmental microbiology, v. 68, n. 8, p. 3795-3801.

PAVEL, V. L. et al. (2013). Symbiosis in the environment biomanagement of soils contaminated with heavy metals, Eur. J. Sci. Theol, v. 9, n. 4, p. 211.

PEDRAZA, R. O. et al. (2004). Aromatic amino acid aminotransferase activity and indole-3-acetic acid production by associative nitrogen-fixing bacteria, FEMS microbiology letters, v. 233, n. 1, p. 15-21.

PEDROSA, F. O. et al. (2011). Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses, PLoS Genet, v. 7, n. 5, p. e1002064.

PELTOLA, H.; H. SÖDERLUND; E. UKKONEN (1984). SEQAID: A DNA sequence assembling program based on a mathematical model, Nucleic acids research, v. 12, n. 1Part1, p. 307-321.

PERLEY, J. E.; B. B. STOWE (1966). On the ability of *Taphrina deformans* to produce indoleacetic acid from tryptophan by way of tryptamine, Plant physiology, v. 41, n. 2, p. 234-237.

PERTEA, M. et al. (2009). OperonDB: a comprehensive database of predicted operons in microbial genomes, Nucleic acids research, v. 37, n. suppl 1, p.

D479-D482.

PEVZNER, P. A. (1989). 1-Tuple DNA sequencing: computer analysis, Journal of Biomolecular structure and dynamics, v. 7, n. 1, p. 63-73.

PEVZNER, P. A.; H. TANG (2001). Fragment assembly with double-barreled data, Bioinformatics, v. 17, n. suppl 1, p. S225-S233.

PEVZNER, P. A.; H. TANG; M. S. WATERMAN (2001). An Eulerian path approach to DNA fragment assembly, Proceedings of the National Academy of Sciences, v. 98, n. 17, p. 9748-9753.

PHILLIPS, A.; D. JANIES; W. WHEELER (2000). Multiple sequence alignment in phylogenetic analysis, Mol Phylogenet Evol, v. 16, n. 3, p. 317-330.

PICCOLO, A.; G. CELANO; G. PIETRAMELLARA (1993). Effects of fractions of coal-derived humic substances on seed germination and growth of seedlings (*Lactuca sativa* and *Lycopersicum esculentum*), Biology and Fertility of Soils, v. 16, n. 1, p. 11-15.

PINGALI, P. L. (2012). Green revolution: impacts, limits, and the path ahead, Proceedings of the National Academy of Sciences, v. 109, n. 31, p. 12302-12308.

PLOETZ, R. C. (2006). Fusarium-induced diseases of tropical, perennial crops, Phytopathology, v. 96, n. 6, p. 648-652.

POP, M. (2009). Genome assembly reborn: recent computational challenges, Brief Bioinform, v. 10, n. 4, p. 354-366.

PRADHAN, P. et al. (2015). Closing yield gaps: How sustainable can we be?, Plos one, v. 10, n. 6, p. e0129487.

PRAKAMHANG, J. et al. (2009). The communities of endophytic diazotrophic bacteria in cultivated rice (*Oryza sativa* L.), Applied Soil Ecology, v. 42, n. 2, p. 141-149.

PREMACHANDRA, D.; L. HUDEK; L. BRAU (2016). Bacterial Modes of Action for Enhancing of Plant Growth, Journal of biotechnology & biomaterials, v. 6, n. 3, p. 1-8.

PRICE, M. N.; A. P. ARKIN; E. J. ALM (2006). The life-cycle of operons, PLoS Genet, v. 2, n. 6, p. e96.

QUAST, C. et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, Nucleic Acids Res, v. 41, n. Database issue, p. D590-596.

RABOY, V. (2009). Approaches and challenges to engineering seed phytate and total phosphorus, Plant Science, v. 177, n. 4, p. 281-296.

RAGHOTHAMA, K.; A. KARTHIKEYAN (2005). Phosphate acquisition. Root

Physiology: from Gene to Function, Springer: 37-49.

RAHMAN, A.; L. PACHTER (2013). CGAL: computing genome assembly likelihoods, Genome biology, v. 14, n. 1, p. R8.

RÄIHÄ, K.-J.; E. UKKONEN (1981). The shortest common supersequence problem over binary alphabet is NP-complete, Theoretical Computer Science, v. 16, n. 2, p. 187-198.

RALSTON, A. (2008). Operons and prokaryotic gene regulation, v., n.

RAMBAUT, A. (2009). FigTree.

RANDRIAMANANTSOA, L. et al. (2013). Can the isotopic exchange kinetic method be used in soils with a very low water extractable phosphate content and a high sorbing capacity for phosphate ions?, Geoderma, v. 200, n., p. 120-129.

RAVISHANKARA, A. R.; J. S. DANIEL; R. W. PORTMANN (2009). Nitrous oxide (N₂O): the dominant ozone-depleting substance emitted in the 21st century, Science, v. 326, n. 5949, p. 123-125.

REID, A.; S. E. GREENE (2013). How microbes can help feed the world, Issues, v., n. 105, p. 33.

RIBEIRO JÚNIOR, J. (2001). Análises estatísticas no SAEG, Viçosa: UFV, v. 301, n.

RICHARDSON, A.; P. HADOBAS; J. HAYES (2000). Acid phosphomonoesterase and phytase activities of wheat (*Triticum aestivum* L.) roots and utilization of organic phosphorus substrates by seedlings grown in sterile culture, Plant, Cell & Environment, v. 23, n. 4, p. 397-405.

RICHTER, M.; R. ROSSELLÓ-MÓRA (2009). Shifting the genomic gold standard for the prokaryotic species definition, Proceedings of the National Academy of Sciences, v. 106, n. 45, p. 19126-19131.

ROMASI, E. F.; J. LEE (2013). Development of indole-3-acetic acid-producing *Escherichia coli* by functional expression of *lpdC*, *aspC*, and *lad1*, Journal of Microbiology and Biotechnology, v. 23, n. 12, p. 1.

RÖMLING, U.; M. Y. GALPERIN (2015). Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions, Trends in microbiology, v. 23, n. 9, p. 545-557.

RUBENCHIK, I.; R. L. STARKEY (1964). *Azotobacter* and its use in agriculture, Soil Science, v. 98, n. 4, p. 280.

RUSSELL, A. B.; S. B. PETERSON; J. D. MOUGOUS (2014). Type VI secretion system effectors: poisons with a purpose, Nature Reviews Microbiology, v. 12, n. 2, p. 137.

- RUTHERFORD, K. et al. (2000). Artemis: sequence visualization and annotation, Bioinformatics, v. 16, n. 10, p. 944-945.
- SAHAI, P.; V. KUMAR (2017). Carriers and Their Role in Plant Agrosystem. Probiotics and Plant Health, Springer: 291-315.
- SALGADO, H. et al. (2000). Operons in Escherichia coli: genomic analyses and predictions, Proceedings of the National Academy of Sciences, v. 97, n. 12, p. 6652-6657.
- SALZBERG, S. L. et al. (2012). GAGE: A critical evaluation of genome assemblies and assembly algorithms, Genome Res, v. 22, n. 3, p. 557-567.
- SANGER, F. (1975). The Croonian Lecture, 1975: Nucleotide Sequences in DNA, Proceedings of the Royal Society of London B: Biological Sciences, v. 191, n. 1104, p. 317-333.
- SANGER, F.; S. NICKLEN; A. R. COULSON (1977). DNA sequencing with chain-terminating inhibitors, Proceedings of the National Academy of Sciences, v. 74, n. 12, p. 5463-5467.
- SANTAMARINA, M. et al. (2017). Bioactivity of essential oils in phytopathogenic and post-harvest fungi control, Natural product research, v. 31, n. 22, p. 2675-2679.
- SANTIAGO, C. D. et al. (2017). Bacterial Compatibility in Combined Inoculations Enhances the Growth of Potato Seedlings, Microbes and environments, v. 32, n. 1, p. 14-23.
- SANTNER, J. et al. (2015). Phosphorus uptake by Zea mays L. is quantitatively predicted by infinite sink extraction of soil P, Plant and soil, v. 386, n. 1-2, p. 371-383.
- SANTOYO, G. et al. (2016). Plant growth-promoting bacterial endophytes, Microbiological research, v. 183, n., p. 92-99.
- SARAF, M.; U. PANDYA; A. THAKKAR (2014). Role of allelochemicals in plant growth promoting rhizobacteria for biocontrol of phytopathogens, Microbiological research, v. 169, n. 1, p. 18-29.
- SARAVANAN, V. S.; M. MADHAIYAN; M. THANGARAJU (2007). Solubilization of zinc compounds by the diazotrophic, plant growth promoting bacterium *Gluconacetobacter diazotrophicus*, Chemosphere, v. 66, n. 9, p. 1794-1798.
- SARWAR, M.; R. KREMER (1995). Determination of bacterially derived auxins using a microplate method, Letters in applied microbiology, v. 20, n. 5, p. 282-285.
- SCHINDLER, D. W. et al. (2008). Eutrophication of lakes cannot be controlled by reducing nitrogen input: results of a 37-year whole-ecosystem experiment, Proceedings of the National Academy of Sciences, v. 105, n. 32, p. 11254-11258.

SEAH, S.-W.; S. NATHAN; K.-L. WAN (2016). Toxicity evaluation of prodigiosin from *Serratia marcescens* in a *Caenorhabditis elegans* model. AIP Conference Proceedings, AIP Publishing.

SEEMANN, T. (2014). Prokka: rapid prokaryotic genome annotation, Bioinformatics, v. 30, n. 14, p. 2068-2069.

SHAIKH, S.; R. SAYYED (2015). Role of plant growth-promoting rhizobacteria and their formulation in biocontrol of plant diseases. Plant microbes symbiosis: applied facets, Springer: 337-351.

SHAO, J. et al. (2015). Analysis and cloning of the synthetic pathway of the phytohormone indole-3-acetic acid in the plant-beneficial *Bacillus amyloliquefaciens* SQR9, Microbial cell factories, v. 14, n. 1, p. 130.

SHARMA, S. B. et al. (2013). Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils, SpringerPlus, v. 2, n. 1, p. 587.

SHEN, H. et al. (2016). A Complex Inoculant of N₂-Fixing, P- and K-Solubilizing Bacteria from a Purple Soil Improves the Growth of Kiwifruit (*Actinidia chinensis*) Plantlets, Front Microbiol, v. 7, n.

SHENDURE, J.; H. JI (2008). Next-generation DNA sequencing, Nat Biotechnol, v. 26, n. 10, p. 1135-1145.

SHROPSHIRE, J. D.; S. R. BORDENSTEIN (2016). Speciation by symbiosis: the microbiome and behavior, mBio, v. 7, n. 2, p. e01785-01715.

SIMAO, F. A. et al. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics, v. 31, n. 19, p. 3210-3212.

SIMPSON, J. T.; M. POP (2015). The theory and practice of genome sequence assembly, Annu Rev Genomics Hum Genet, v. 16, n., p. 153-172.

SINGH, R. P.; P. N. JHA (2016). The Multifarious PGPR *Serratia marcescens* CDP-13 Augments Induced Systemic Resistance and Enhanced Salinity Tolerance of Wheat (*Triticum aestivum* L.), Plos one, v. 11, n. 6, p. e0155026.

SOLANKI, M.; N. DIDWANIA; V. NANDAL (2016). Potential of Zinc Solubilizing Bacterial Inoculants in Fodder Crops, momentum, v., n.

SON, J.-S. et al. (2014). Screening of plant growth-promoting rhizobacteria as elicitor of systemic resistance against gray leaf spot disease in pepper, Applied Soil Ecology, v. 73, n., p. 1-8.

SONG, X. et al. (2014). Heavy metal and nutrient changes during vermicomposting animal manure spiked with mushroom residues, Waste management, v. 34, n. 11, p. 1977-1983.

SOUZA, R. D.; A. AMBROSINI; L. M. PASSAGLIA (2015). Plant growth-

promoting bacteria as inoculants in agricultural soils, Genetics and molecular biology, v. 38, n. 4, p. 401-419.

SPAEPEN, S.; J. VANDERLEYDEN (2011). Auxin and plant-microbe interactions, Cold Spring Harb Perspect Biol, v. 3, n. 4.

SPAEPEN, S.; J. VANDERLEYDEN; Y. OKON (2009). Plant growth-promoting actions of rhizobacteria, Advances in botanical research, v. 51, n., p. 283-320.

SPAEPEN, S.; J. VANDERLEYDEN; R. REMANS (2007a). Indole-3-acetic acid in microbial and microorganism-plant signaling, FEMS Microbiol Rev, v. 31, n. 4, p. 425-448.

SPAEPEN, S. et al. (2007b). Characterization of phenylpyruvate decarboxylase, involved in auxin production of *Azospirillum brasilense*, J Bacteriol, v. 189, n. 21, p. 7626-7633.

STADEN, R. (1979). A strategy of DNA sequencing employing computer programs, Nucleic acids research, v. 6, n. 7, p. 2601-2610.

STAMATAKIS, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, Bioinformatics, v. 30, n. 9, p. 1312-1313.

STOYKOV, Y. M.; A. I. PAVLOV; A. I. KRASTANOV (2015). Chitinase biotechnology: production, purification, and application, Engineering in Life Sciences, v. 15, n. 1, p. 30-38.

SURYAWANSHI, R. K. et al. (2015a). In vitro antiparasitic activity of microbial pigments and their combination with phytosynthesized metal nanoparticles, Parasitology international, v. 64, n. 5, p. 353-356.

SURYAWANSHI, R. K. et al. (2015b). Mosquito larvicidal and pupaecidal potential of prodigiosin from *Serratia marcescens* and understanding its mechanism of action, Pesticide biochemistry and physiology, v. 123, n., p. 49-55.

SUTTON, G. G. et al. (1995). TIGR Assembler: A new tool for assembling large shotgun sequencing projects, Genome Science and Technology, v. 1, n. 1, p. 9-19.

TABASSUM, B. et al. (2017). Bottlenecks in commercialisation and future prospects of PGPR, Applied Soil Ecology, v. 121, n., p. 102-117.

TAGHAVI, S. et al. (2005). Horizontal gene transfer to endogenous endophytic bacteria from poplar improves phytoremediation of toluene, Appl Environ Microbiol, v. 71, n. 12, p. 8500-8505.

TALAVERA, G.; J. CASTRESANA (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments, Systematic biology, v. 56, n. 4, p. 564-577.

TARAFDAR, J.; A. GHARU (2006). Mobilization of organic and poorly soluble

phosphates by *Chaetomium globosum*, Applied Soil Ecology, v. 32, n. 3, p. 273-283.

TARHIO, J.; E. UKKONEN (1988). A greedy approximation algorithm for constructing shortest common superstrings, Theoretical Computer Science, v. 57, n. 1, p. 131-145.

TARIQ, M. et al. (2017). Antagonistic features displayed by Plant Growth Promoting Rhizobacteria (PGPR): A Review, Journal of Plant Science and Phytopathology, v. 1, n., p. 38-43.

TATUSOVA, T. et al. (2013). RefSeq microbial genomes database: new representation and annotation strategy, Nucleic acids research, v. 42, n. D1, p. D553-D559.

TATUSOVA, T. et al. (2016). NCBI prokaryotic genome annotation pipeline, Nucleic Acids Res, v. 44, n. 14, p. 6614-6624.

TILMAN, D. et al. (2011). Global food demand and the sustainable intensification of agriculture, Proceedings of the National Academy of Sciences, v. 108, n. 50, p. 20260-20264.

TILMAN, D. et al. (2002). Agricultural sustainability and intensive production practices, Nature, v. 418, n. 6898, p. 671-677.

TIMMUSK, S. et al. (2017). Perspectives and challenges of microbial application for crop improvement, Frontiers in plant science, v. 8, n., p. 49.

TRABELSI, D.; R. MHAMDI (2013). Microbial inoculants and their impact on soil microbial communities: a review, Biomed Res Int, v. 2013, n.

TRAUTMANN, N.; E. OLYNCIW (2000). Compost microorganisms, CORNELL Composting, Science, v., n.

TREANGEN, T. J.; S. L. SALZBERG (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions, Nat Rev Genet, v. 13, n. 1, p. 36-46.

TREVISAN, S. et al. (2010). Humic substances biological activity at the plant-soil interface: from environmental aspects to molecular factors, Plant signaling & behavior, v. 5, n. 6, p. 635-643.

TRIVEDI, P.; A. PANDEY; L. M. S. PALNI (2005). Carrier-based preparations of plant growth-promoting bacterial inoculants suitable for use in cooler regions, World Journal of Microbiology and Biotechnology, v. 21, n. 6, p. 941-945.

UEDA, N. et al. (2012). *Agrobacterium tumefaciens* tumor morphology root plastid localization and preferential usage of hydroxylated prenyl donor is important for efficient gall formation, Plant physiology, v. 159, n. 3, p. 1064-1072.

ULLAH, A. et al. (2015). Phytoremediation of heavy metals assisted by plant

growth promoting (PGP) bacteria: a review, Environmental and Experimental Botany, v. 117, n., p. 28-40.

ULLOA-OGAZ, A.; L. MUÑOZ-CASTELLANOS; G. NEVÁREZ-MOORILLÓN (2015). Biocontrol of phytopathogens: antibiotic production as mechanism of control, The battle against microbial pathogens: basic science, technological advances and educational programmes. FORMATEX RESEARCH CENTER, Spain, v., n., p. 305-309.

VAIKUNTAPU, P. R. et al. (2016). A new chitinase-D from a plant growth promoting *Serratia marcescens* GPS5 for enzymatic conversion of chitin, Bioresour Technol, v. 220, n., p. 200-207.

VEJAN, P. et al. (2016). Role of Plant Growth Promoting Rhizobacteria in Agricultural Sustainability-A Review, Molecules, v. 21, n. 5.

VENTER, J. C. et al. (2001). The sequence of the human genome, science, v. 291, n. 5507, p. 1304-1351.

VENTURI, V.; C. KEEL (2016). Signaling in the rhizosphere, Trends in plant science, v. 21, n. 3, p. 187-198.

VERBON, E. H.; L. M. LIBERMAN (2016). Beneficial microbes affect endogenous mechanisms controlling root development, Trends in plant science, v. 21, n. 3, p. 218-229.

VERNIKOS, G. et al. (2015). Ten years of pan-genome analyses, Current opinion in microbiology, v. 23, n., p. 148-154.

VICENTE, E. J.; D. R. DEAN (2017). Keeping the nitrogen-fixation dream alive, Proc Natl Acad Sci U S A, v. 114, n. 12, p. 3009-3011.

VORHOLT, J. A. (2012). Microbial life in the phyllosphere, Nature Reviews Microbiology, v. 10, n. 12, p. 828-840.

VYAS, P. et al. (2010). Cold-adapted and rhizosphere-competent strain of *Rahnella* sp. with broad-spectrum plant growth-promotion potential, J Microbiol Biotechnol, v. 20, n. 12, p. 1724-1734.

WAJID, B.; E. SERPEDIN (2012). Review of general algorithmic features for genome assemblers for next generation sequencers, Genomics Proteomics Bioinformatics, v. 10, n. 2, p. 58-73.

WANG, G. et al. (2004). Effect of different C sources on the solubilization of rock phosphate by three phosphate solubilizing fungi (PSF), Chinese Journal of Ecology, v. 23, n. 2, p. 32-36A.

WANG, S.; C. N. MULLIGAN (2009). Enhanced mobilization of arsenic and heavy metals from mine tailings by humic acid, Chemosphere, v. 74, n. 2, p. 274-279.

WANG, X. et al. (2015). Biocontrol and plant growth-promoting activity of

- rhizobacteria from Chinese fields with contaminated soils, Microbial biotechnology, v. 8, n. 3, p. 404-418.
- WATERHOUSE, A. M. et al. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench, Bioinformatics, v. 25, n. 9, p. 1189-1191.
- WEYENS, N. et al. (2009a). Exploiting plant–microbe partnerships to improve biomass production and remediation, Trends in biotechnology, v. 27, n. 10, p. 591-598.
- WEYENS, N. et al. (2009b). Phytoremediation: plant–endophyte partnerships take the challenge, Current Opinion in Biotechnology, v. 20, n. 2, p. 248-254.
- WICK, R. R. et al. (2015). Bandage: interactive visualization of de novo genome assemblies, Bioinformatics, v., n.
- WU, D. et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea, Nature, v. 462, n. 7276, p. 1056-1060.
- XIE, S. S. et al. (2014). Plant growth promotion by spermidine-producing *Bacillus subtilis* OKB105, Mol Plant Microbe Interact, v. 27, n. 7, p. 655-663.
- YANG, Z.; B. RANNALA (2012). Molecular phylogenetics: principles and practice, Nature reviews genetics, v. 13, n. 5, p. 303-314.
- ZAIDI, A. et al. (2009). Plant growth promotion by phosphate solubilizing bacteria, Acta microbiologica et immunologica Hungarica, v. 56, n. 3, p. 263-284.
- ZDOBNOV, E. M. et al. (2016). OrthoDB v9. 1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs, Nucleic acids research, v. 45, n. D1, p. D744-D749.
- ZELLER, E. et al. (2015). Hydrolysis of phytate and formation of inositol phosphate isomers without or with supplemented phytases in different segments of the digestive tract of broilers, Journal of nutritional science, v. 4, n.
- ZERBINO, D. R.; E. BIRNEY (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs, Genome Res, v. 18, n. 5, p. 821-829.
- ZHOU, W. et al. (2016). Antiviral activity and specific modes of action of bacterial prodigiosin against *Bombyx mori* nucleopolyhedrovirus in vitro, Applied microbiology and biotechnology, v. 100, n. 9, p. 3979-3988.
- ZIMIN, A. V. et al. (2013). The MaSuRCA genome assembler, Bioinformatics, v. 29, n. 21, p. 2669-2677.