

ANÁLISES GENÔMICAS DE DUAS RIZOBACTÉRIAS
PROMOTORAS DO CRESCIMENTO VEGETAL ISOLADAS DE
FRUTEIRAS

GUSTAVO LIMA RODRIGUES

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO – UENF

Campos dos Goytacazes –RJ
Fevereiro de 2018

ANÁLISES GENÔMICAS DE DUAS RIZOBACTÉRIAS
PROMOTORAS DO CRESCIMENTO VEGETAL ISOLADAS DE
FRUTEIRAS

GUSTAVO LIMA RODRIGUES

Dissertação apresentada ao Centro de
Biociências e Biotecnologia, da
universidade Estadual do Norte
Fluminense, como parte das exigências
para obtenção do título de Mestre em
Biociências e Biotecnologia.

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO – UENF

Campos dos Goytacazes –RJ
Fevereiro de 2018

ANÁLISES GENÔMICAS DE DUAS RIZOBACTÉRIAS
PROMOTORAS DO CRESCIMENTO VEGETAL ISOLADAS DE
FRUTEIRAS

GUSTAVO LIMA RODRIGUES

Dissertação apresentada ao Centro de
Biociências e Biotecnologia, da
universidade Estadual do Norte
Fluminense, como parte das exigências
para obtenção do título de Mestre em
Biociências e Biotecnologia.

Aprovada no dia 22/02/2018

Comissão examinadora:

Prof. Dr. Daniel Basílio Zandonadi – UFRJ

Dr. Cleiton de Paula Soares – UENF

Prof. Dr. Álvaro Fabrício Lopes Rios – UENF

Prof. Dr. Fábio Lopes Olivares – UENF (co-orientador)

Prof. Dr. Thiago Motta Venâncio – UENF (orientador)

Dedicatória

Aos meus pais Eron e Izabel, meu irmão Caio (*in memoriam*), e todos meus queridos parentes e amigos que me acompanharam nesta trajetória.

Agradecimentos

Ao meu orientador Thiago Venâncio, co-orientador Fábio Olivares e demais pesquisadores: Filipe Matteoli, Kanhu Moharana, Rajesh Gazara, Francisnei Pedrosa, Hemanoel Passarelli, Lupis Ribeiro, Daniel Bellieny, Fabrício Brum e Fabrício Almeida.

RESUMO

Dentre as consequências negativas da síntese e do uso em larga escala de insumos químicos sintéticos na produção agrícola estão a eutrofização de corpos d'água e aquecimento global. Para se reduzir o dano ambiental associado à produção agrícola, é desejável o uso de tecnologias ecologicamente sustentáveis, dentre elas o uso de rizobactérias promotoras do crescimento vegetal (PGPR, do inglês *Plant Growth Promoting Rhizobacteria*). PGPR tipicamente exercem seus efeitos benéficos por meio de suas propriedades de biofertilização, bioestimulação e biocontrole. Nosso grupo de pesquisa isolou e sequenciou o genoma de diversas PGPR do sistema solo-planta, tendo sido duas cepas, UENF412522 e UENF114111, isoladas da rizosfera de fruteiras. A análise genômica destes dois isolados é o tema desta dissertação. Ambas as cepas têm habilidade de fixar nitrogênio e solubilizar fosfato. Estas características são importantes do ponto de vista da biofertilização, o que as torna candidatas a integrar a formulação de novos inoculantes. As sequências dos *reads* de UENF412522 e UENF114111 foram montadas com diferentes montadores, seguida da predição e anotação dos genes. Para caracterização das cepas, foram realizadas reconstruções filogenéticas, identificação por rRNA 16S e busca de correlação tetra. Por meio de diferentes métodos, conseguimos classificar as cepas UENF412522 e UENF114111 como pertencentes aos gêneros *Azospirillum* e *Paraburkholderia*, respectivamente. Ambos os genomas apresentaram genes codificando componentes da fixação biológica de nitrogênio (*nif*), ácido 1-amino-ciclopropano-1-carboxílico (ACC) desaminase (*acdS*) e ativadora de citocina LOG. A cepa UENF412522 possui o gene final da biossíntese de auxina *iaaH*, os principais genes do ciclo de Calvin-Benson-Bassham, e possui uma forma não discutida na literatura do gene *treZ*, envolvido na síntese de trealose. A cepa UENF114111 possui o operon *pqq* completo, indicando que esta cepa tem a possibilidade de solubilizar fosfato. Ambas as cepas apresentaram conjuntos de genes de resistência a antibióticos de bactérias tipicamente residentes do solo e para o metabolismo de diversas fontes de carbono. Em conjunto, nossos resultados indicam que estas cepas apresentam potencial para integrar formulações de novos inoculantes agrícolas, principalmente por suas propriedades de biofertilização e de genes que provavelmente auxiliam sua permanência na rizosfera.

ABSTRACT

The eutrophication of water bodies and global warming are among the consequences of broad-scale use and synthesis of chemical fertilizers in agriculture. The use of sustainable technologies, such as plant-growth promoting rhizobacteria (PGPR), is desirable to reduce the environmental impact associated with current agriculture practices. PGPR usually benefit crops by biofertilization, biostimulation and biocontrol properties. Our research group has recently isolated and sequenced several PGPR genomes from plant-bacteria system. Among these PGPR, the strains UENF412522 and UENF114111 were isolated from fruit trees. The genome analysis from these two isolates is the subject of this dissertation. Both strains possess the ability to fix nitrogen and solubilize phosphorus. These features are important for biofertilization, making these strains good candidates to integrate new inoculant formulations. Sequencing reads from UENF412522 and UENF114111 were assembled with different assemblers, followed by gene prediction and annotation. Strain characterization was inferred through phylogeny reconstructions, 16S rRNA analysis and tetra correlation analysis. These analysis combined suggest the strains UENF412522 and UENF114111 belong, to the genera *Azospirillum* and *Paraburkholderia*, respectively. Both genomes have genes encoding components of the nitrogen biological fixation apparatus (*nif*), 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (*acdS*) and cytokinin-activating LOG. The strain UENF412522 possess the auxin biosynthesis gene *iaaH*, the main genes from Calvin-Benson-Bassham cycle, and a undiscussed form of the *treZ* gene, involved in trehalose synthesis. The strain UENF114111 possesses a complete *pqq* operon, supporting the hypothesis that this strain is able to solubilize phosphate. Both strains possess a set of resistance genes against antibiotics from soil bacteria and also genes involved in the metabolism of various carbon sources. As a whole, our results underscore the potential of these strains to integrate new inoculant formulations, mainly due to their biofertilization properties and genes that allow their persistence in the rhizosphere.

Sumário

| | |
|--|----|
| 1. INTRODUÇÃO | 1 |
| 1.1 CONSIDERAÇÕES GERAIS | 1 |
| 1.2 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO | 3 |
| 1.3 SOLUBILIZAÇÃO DE FOSFATO..... | 7 |
| 1.4 AUXINA E CITOCININA | 9 |
| 1.5 ETILENO E ACC DESAMINASE | 12 |
| 1.6 CEPAS UENF114111 E UENF412522..... | 13 |
| 2. JUSTIFICATIVA | 15 |
| 3. OBJETIVO | 15 |
| 4. MÉTODOS | 16 |
| 5. RESULTADOS E DISCUSSÃO | 18 |
| 5.1 Montagem dos genomas de UENF412522 e UENF114111 | 18 |
| 5.2 Identificação filogenética das cepas UENF412522 e UENF114111 | 19 |
| 5.3 Predição e anotação gênica | 22 |
| 5.4 Genes relacionados à promoção do crescimento vegetal e outros processos ecofisiológicos..... | 24 |
| 5.5 Genes de metabolismo de carbono..... | 32 |
| 7. CONCLUSÃO | 34 |
| 8. REFERÊNCIAS..... | 35 |

Índice de tabelas

| | |
|---|----|
| Tabela 1 – Identificação, vegetal hospedeiro e propriedades bioquímica relacionadas a promoção do crescimento vegetal nas cepas UENF114111 e UENF412522 (Rodrigues et al., 2016) | 15 |
| Tabela 2 – Métricas da montagem para cepa UENF114111 utilizando os montadores SPADes, IDBA, Velvet e Minia. Em Ctg estão evidenciadas métricas para os <i>contigs</i> e em Scf para os <i>Scaffolds</i> . Os valores indicando as melhores métricas estão sublinhados | 18 |
| Tabela 3 – Métricas da montagem para cepa UENF412522 utilizando os montadores SPADes, IDBA, Velvet e Minia. Em Ctg estão evidenciadas métricas para os <i>contigs</i> e em Scf para os <i>Scaffolds</i> . Os valores indicando as melhores métricas estão sublinhados | 19 |
| Tabela 4 – Teste para fontes de carbono API 50 CH/E com seus respectivos possíveis genes de metabolismo. O sinal de + indica habilidade de metabolizar a fonte, e - inabilidade de metabolizar, e ? gene não encontrado | 33 |

Índice de Figuras

- Figura 1 – Esquema da rizosfera e suas subdivisões. As bactérias podem colonizar o solo diretamente (ectorizosfera), se aderir à superfície da raiz (rizoplano, indicado em amarelo) ou ocupar espaços vazios em porções do córtex e da endoderme (endorizosfera). Adaptado de (McNear Jr., 2013)..... 2
- Figura 2 – Estrutura do complexo da nitrogenase (a) e seu mecanismo de ação (b). a: As proteínas Fe formam um complexo com a proteína MoFe através de seus grupamentos 4Fe-4S. O grupamento P (8Fe-7S) transfere os elétrons da proteína Fe para o cofator FeMo, aonde ocorre a fixação do nitrogênio. b: A proteína Fe oxidada é reduzida ao custo de dois ATPs, e se complexa com a proteína MoFe, transferindo-lhe dois elétrons e voltando a forma oxidada, sendo então dissociada da proteína MoFe. Adaptado de (Dixon & Kahn 2004) 5
- Figura 3 – Regulação do complexo da nitrogenase. a: Regulação em *K. pneumoniae*, sendo iniciada por GlnD. Na presença de glutamina, GlnD desuridilila GlnB, que interage com NtrB, defosforilando NtrC e bloqueando o restante da cascata. Na ausência de glutamina, GlnD uridilila GlnB, que interage com NtrB fosforilando NtrC, que é um promotor de *nifL*, *nifA*, *glnK*. NifL, na presença de oxigênio complexa-se com NifA e bloqueia a transcrição de *nif*. Na ausência de oxigênio, NifL descomplexa-se de NifA pela ação de GlnK, permitindo que NifA promova a transcrição dos outros genes *nif*. b: Regulação em *S. meliloti*, sendo iniciada por FixL. Na presença de oxigênio, FixL não fosforila FixJ, bloqueando o restante da cascata. Na ausência de oxigênio, FixL auto-fosforila e fosforila FixJ, promovendo a transcrição de *nifA* e *fixK*. NifA é regulado diretamente por oxigênio, promovendo a transcrição dos genes *nif* na sua ausência. Este sistema é auto-regulado indiretamente por FixK, que promove a transcrição de *fixT*, que inibe a fosforilação de FixJ. Adaptado de (Dixon & Kahn, 2004)..... 7
- Figura 4 – Vias de biossíntese de auxina dependentes de triptofano. Os genes conhecidos envolvidos nas rotas estão em itálico. O triptofano pode ser metabolizado em IAA por cinco vias: IAM, IAN, IPA, TAM, e TSO. Na IAM, o triptofano é convertido em IAM, que é hidrolisado por IaaH em IAA. Na via do IAN, o triptofano é convertido em indol-3-acetaldoxima, que é convertido em IAN. O IAN pode então ser desnitrificado em IAA por nitrilase ou ser convertido em IAM por nitrila hidratase. Na IPA, o triptofano é transaminado por LAO em IPA, que é descarboxilado por IpdC em IAAld, que é desidrogenado por Y4wF em IAA. Na TAM, o triptofano é descarboxilado por TDC em triptamina, que é convertida em IAAld por uma amino oxidase. Na TSO, o triptofano é convertido diretamente em IAAld. A partir do IAAld, as vias TAM e TSO compartilham o mesmo caminho metabólico que IPA. Esquema adaptado de (Imada et al., 2017; Spaepen & Vanderleyden, 2011; Spaepen et al., 2007)..... 10
- Figura 5 – Esquema de desvio da rota de etileno mediado por ACC desaminase. Em condições de estresse, o vegetal sintetiza auxina, que sinaliza a produção de SAM, que é convertido em ACC pela enzima ACC sintase. O ACC é convertido em etileno pela ACC oxidase. Neste sistema de interação planta-bactéria, a bactéria produz a enzima ACC desaminase, que cliva o ACC em α -cetobutarato e amônia, desviando a rota de síntese do etileno, ao mesmo tempo que produz auxina a partir do triptofano vegetal. Adaptado de (Glick, 2014) 13
- Figura 6 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos *core genes* para as espécies de *Azospirillum*. *R. centenum* SW foi definido como grupo externo. A cepa UENF412522 (destacada em verde) formou um grupo irmão com um clado contendo *Azospirillum sp.* B2, *Azospirillum sp.* B506, *A. oryzae* A2P, *A. humicireducens* SgZ5 e *A. lipoferum* 4B. A única espécie deste gênero contida em um clado monofilético representativo é *A. brasilense*. 21
- Figura 7 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos *core genes* para as espécies de *Paraburkholderia*. A cepa UENF114111 (destacada em verde) esta como grupo irmão de um clado contendo *P. nodosa*, *P. heleaia* e *P. mimosarum*, sem se agrupar diretamente com nenhuma outra cepa.

Observa-se um ramo longo no interior da filogenia com baixo suporte contendo *P. rhizoxinica* HKI454. Desconsiderando as relações com suporte menor do que 50% (ramos marcados com triângulo vermelho), temos uma politomia que desmembra em *P. rhizoxinica* HKI454, *P. caryophylli* Ballard720, um clado contendo as espécies com sufixo candidatos (porção inferior) e o clado contendo as demais espécies. A cepa *B. thalandsis* E254 foi definida como grupo externo..... 23

Figura 8 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos core genes para os gêneros *Paraburkholderia* e *Ralstonia*. *P. necessarius* QLW-P1DMWA-1 foi definido como grupo externo. A cepa *P. rhizoxinica* HKI454 mostrou-se um grupo basal dentro de *Paraburkholderia*, sendo a espécie mais distinta deste clado 26

Figura 9 – Distribuição dos genes *nif* posicionados em representações dos *scaffolds* das cepas UENF41522 (A) e UENF114111 (B). Na cepa UENF412522, os genes da FBN estão em oito operons dispersos em 5 *scaffolds*. Na cepa UENF114111, quatro operons contendo os genes da FBN estão adjacentes no mesmo *scaffold*..... 27

Figura 10 – Análise dos domínios de treZ da cepa UENF412522 pelo CDD-NCBI. Há um domínio de maltoquinase (malQ) anexado ao restante da estrutura de treZ. 28

Figura 11 – blastP contra a base nr para treZ de UENF412522. Observa-se a mesma forma de treZ+malQ em nove espécies de *Azospirillum*, em *Belnapia moabensis* e *Rhodospirillum centenum*. Esta fusão aparenta ser específica de alguns membros da família Rhodospirillaceae, uma vez que as cepas de Skermanella não possuem esta estrutura 29

Figura 12 – Visualização circular do genoma da cepa UENF412522. Nos anéis mais externos (cinza) estão plotados os *scaffolds*, seguido por porcentagem de GC (laranja), posição dos genes e suas respectivas identificações. Em azul os genes de resistência à antibióticos: fosmidomicina (*fsr*), biciclomicina (*bcr*), bacitracina (*uppP/BacA*), tabtoxina (*ttr*), fenazina (*ehpR*), tetraciclina (*tetA*, resistência a múltiplas drogas *MdtABCE*, *AcrAB* e *NorM*). Em vermelho os genes de metabolismo de carbono: importação-fosforilação de manose (manPTS) e frutose (fruPTS), α -glucosidase (aglucos), β -glucosidase (bglucos), ciclo de Calvin-Benson-Brassham (CBB), ribulose quinase + riboquinase (ribuloribk), trehalose (treYZ, treS, otsB+frutoquinase, *otsA*+glicoquinase), operon da xilose (*xyl*), operon do glicerol (*glp*), operon da arabinose (*ara*), glicoquinase (*glk*), frutoquinase-6 + manitol-2-desidrogenase (*manf6p*). Em verde, os genes diretamente relacionados a promoção do crescimento vegetal: síntese de ácido glucônico (pqq), citocinina (*log*), ACC desaminase (*acds*) e nitrogenase (*nifHDK*). 30

Figura 13 – Visualização circular do genoma da cepa UENF114111. Nos anéis mais externos (cinza) estão plotados os *scaffolds*, seguido por porcentagem de GC (laranja), posição dos genes e suas respectivas identificações. Em azul os genes de resistência à antibióticos: fosmidomicina (*fsr*), metilenomicina A (*mmr*), tabtoxina (*ttr*), biciclomicina (*bcr*), polimixina (*ArnA/pmr*), ácido fusarico (*fusAC*), bacitracina (*uppP/BacA*), toxina *ratA* (*pasT*), resistência a múltiplas drogas *MdtABCEG*. Em vermelho os genes de metabolismo de carbono: operon da xilose (*xylAB*), ribitol (*rbt*), β -glucosidase (bglucos), frutoquinase (*fru*), operon da ramnose (*rha*), glicerol quinase (*glp*), operon da arabinose (*ara*), trealose (treSYZ), Glutamina--frutose-6-fosfato aminotransferase (F6P-gln), manose-6-fosfato isomerase (F6P-Man6P), operon da ribose (*rbs*), glicoquinase (*glk*), sorbitol desidrogenase (sorbitol-sorbose), manitol-2-desidrogenase (Man-Fru). Em verde, os genes relacionados diretamente a promoção do crescimento vegetal: citocinina (*log*), auxina (*iaaH*) e FBN (*nif*). 31

1. INTRODUÇÃO

1.1 CONSIDERAÇÕES GERAIS

No final do século XVIII, Thomas Malthus propôs a teoria populacional malthusiana: a disponibilidade de recursos limita o tamanho populacional. Ao se alcançar o limite dos recursos, o tamanho populacional pode ser reduzido pela fome, ameaçando a integridade da sociedade humana. A partir da revolução industrial, a taxa de crescimento da população apresentou tendência exponencial (Van Bavel, 2013), enquanto a taxa de produção da agricultura era linear. Para impedir uma possível epidemia de inanição, principalmente nos países em desenvolvimento (Khush, 2001), a comunidade científica mobilizou-se durante o século XX para aumentar a produção agrícola. A revolução verde foi resultado deste esforço, multiplicando a produção agrícola mundial principalmente pelo desenvolvimento de variedades modernas de arroz, milho e soja, acoplado a novas tecnologias de manejo cultural e uso de insumos químicos sintéticos agrícolas, como pesticidas e fertilizantes (Pingali, 2012). Os países em desenvolvimento foram os mais beneficiados, chegando a triplicar a produção, o que acarretou na redução dos preços dos alimentos e aumento da competitividade no mercado internacional (Evenson & Gollin, 2003).

Por conta do pacote tecnológico da revolução verde, a agricultura moderna tornou-se extremamente dependente do uso de fertilizantes e agrotóxicos, o que contribuiu para o aumento da poluição tanto em ambientes terrestres quanto nos aquáticos (Bodirsky et al., 2014; Carpenter et al., 1998; Mabilde et al., 2017; Vitousek et al., 1997). Neste sentido, faz-se necessário o desenvolvimento e aplicação de tecnologias para uma agricultura sustentável e que cumpra a crescente demanda de alimentos para as próximas décadas (Ray et al., 2013; Tilman et al., 2011).

O uso de bactérias promotoras do crescimento vegetal (*Plant Growth-Promoting Bacteria*, PGPB) se apresenta com uma alternativa bastante promissora, principalmente pelo seu impacto ambiental reduzido, para aumentar a produção agrícola. Estas bactérias podem ter vida livre, formar associações simbióticas com plantas (e.g. nódulos em leguminosas com bactérias fixadoras de nitrogênio) e colonizar os tecidos internos dos vegetais (Glick, 2012; Wang et al., 2012). As PGPB de vida livre geralmente concentram-se na rizosfera (Figura 1) devido à disponibilidade de uma ampla fonte de nutrientes exsudados pelas raízes da planta hospedeira (Bais et al., 2006; Jones,

1998). Estas bactérias da rizosfera são frequentemente referenciadas como rizobactérias promotoras do crescimento vegetal (*Plant Growth Promoting Rhizobacteria*, PGPR).

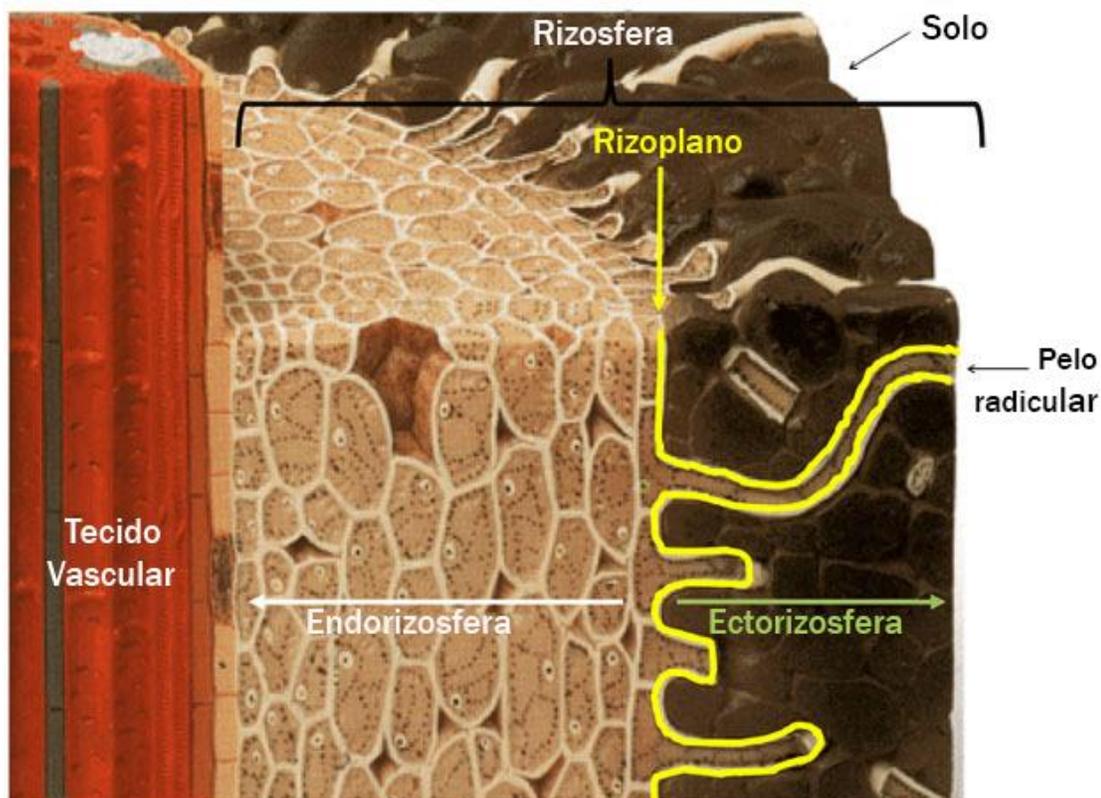


Figura 1 – Esquema da rizosfera e suas subdivisões. As bactérias podem colonizar o solo diretamente (ectorizosfera), se aderir à superfície da raiz (rizopiano, indicado em amarelo) ou ocupar espaços vazios em porções do córtex e da endoderme (endorizosfera). Adaptado de (McNear Jr., 2013).

As PGPR podem promover o crescimento através de diversos mecanismos, tais como: produção de fitormônios como auxinas (Spaepen & Vanderleyden, 2011) e citocininas (Arshad & Frankenberger, 1991); fixação de nitrogênio atmosférico (dinitrogênio, N_2) (Kuan et al., 2016); produção da enzima 1-aminociclopropano-1-carboxilato (ACC) desaminase (Honma & Shimomura, 1978), inibidora da síntese do hormônio etileno (Glick, 2014; Glick et al., 1998); compostos voláteis bacterianos (Chung et al., 2016; Schulz & Dickschat, 2007) que atuam na promoção do crescimento (Vespermann et al., 2007) e das defesas da planta (Aziz et al., 2016); liberação de sideróforos (Radzki et al., 2013) e outros compostos solubilizadores de nutrientes

(Intorne et al., 2009); antibióticos (Mansouri & Piepersberg, 1991) e agentes de controle biológico (Raddadi et al., 2007; Schwartz et al., 2004).

1.2 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO

A fixação biológica de nitrogênio (FBN) foi o primeiro mecanismo de promoção de crescimento vegetal por bactérias a ser descoberto (Elmerich, 2015). Considerando a grande eficiência destas bactérias e o fato do nitrogênio ser o principal nutriente mineral limitante do crescimento vegetal, a FBN tem grande relevância econômica: no ano de 2013 foram economizados R\$24,9 bilhões com fertilizantes nitrogenados industrializados pelo uso de bactérias fixadoras de nitrogênio (EMBRAPA), em particular na cultura da soja (Chang et al., 2015).

Além da grande importância econômica, a FBN possibilita a fertilização nitrogenada com baixo impacto ambiental. Os fertilizantes nitrogenados químicos são produzidos pela síntese de Haber-Bosch, um processo que demanda alta temperatura e pressão para reduzir o dinitrogênio atmosférico em amônia. Para atingir esta temperatura, recursos de matriz não renovável são empregados (Vicente & Dean, 2017), aumentando a emissão de carbono associada à fertilização nitrogenada. Além disso, a aplicação em excesso de fertilizantes nitrogenados causa uma série de impactos ambientais negativos. O nitrogênio não assimilado pelas culturas pode ser denitrificado na forma de óxido nitroso, este reconhecido por seu elevado efeito estufa (Ravishankara et al., 2009). A parcela do nitrogênio não volatilizada/denitrificada e não assimilada pelas culturas pode ser lixiviada para corpos d'água, contribuindo para sua eutrofização (Bodirsky et al., 2014). A FBN por outro lado, utiliza um mecanismo enzimático em condições normais de temperatura e pressão para reduzir o nitrogênio atmosférico a amônia, provendo-o de forma controlada ao ambiente e/ou hospedeiro vegetal.

As bactérias fixadoras de nitrogênio (diazotróficas) podem se estabelecer no interior dos tecidos vegetais (e.g. nódulos nas raízes e espaços endofíticos) e/ou em regiões externas (e.g. aderidas à superfície vegetal – rizoplano ou no solo, em torno das raízes – rizosfera) em uma relação simbiótica, pois a planta se beneficia da fertilização nitrogenada conduzida pelas bactérias e lhes fornecem abrigo e glicose. Quando há formação de nódulos, a relação bactéria-planta é classificada como endossimbiótica e há uma alteração morfológica/bioquímica da planta para englobar as bactérias diazotróficas em um tecido vegetal. No entanto, nas associações endofíticas,

tais modificações morfológicas perceptíveis não ocorrem, sendo os espaços intercelulares (Carvalho et al., 2014) e o xilema (James & Olivares, 1998) os locais preferenciais de colonização. Durante estes eventos de colonização, a planta ativa seus mecanismos de defesa, como acumulação de compostos fenólicos (Mandal et al., 2010) e o endurecimento da parede celular (Duijff et al., 1997), dificultando o acesso dos invasores aos tecidos internos. A colonização por bactérias endofíticas benéficas pode ser chamada de endossimbiose (Reddy & Saravanan, 2013), embora este evento geralmente se refira à colonização intracelular (Biondi et al., 2011; de Almeida et al., 2009; McFadden, 1999), havendo então alguma discordância na literatura.

A FBN depende do complexo da nitrogenase (formado por proteínas codificadas pelos genes *nif*), que é responsável pela conversão de N_2 em amônia (NH_3) (Dixon & Kahn, 2004; Howard & Rees, 2006). A nitrogenase é constituída por uma proteína maior contendo ferromolibdênio (proteína FeMo) e duas proteínas menores contendo ferro (proteína Fe, Figura 2a). As proteínas Fe possuem grupamentos 4Fe-4S em sua superfície, que se ligam reversivelmente a proteína FeMo, transferindo-lhe elétrons ao custo de 16 ATPs por mol de N-fixado. Esta etapa de acoplamento e desacoplamento é o principal limitante da velocidade catalítica da nitrogenase e quando uma proteína Fe transfere elétron, a outra proteína Fe é inibida e vice-versa (Danyal et al., 2016). A proteína FeMo possui quatro subunidades com dois centros metálicos distintos: duas subunidades com o grupamento P (contendo 8Fe-7S) e outras duas subunidades contendo um cofator FeMo ($MoFe_7S_9$ -homocitrato). A subunidade do grupamento P recebe o elétron da proteína Fe e o transfere para o cofator FeMo (Figura 2b). O cofator FeMo acomoda o gás N_2 entre os íons metálicos 7Fe e 1Mo (Einsle et al., 2002), estabelecendo ligações de Mo-N e Fe-N (Seefeldt et al., 2004), que enfraquecem a forte ligação tripla entre os átomos do dinitrogênio. Ao enfraquecer a ligação, um átomo de hidrogênio é ligado ao nitrogênio, desfazendo uma das ligações entre os nitrogênios. Quando a ligação tripla for completamente desfeita, haverá duas amônias.

O complexo da nitrogenase é vulnerável ao oxigênio, pois os grupamentos 4Fe-4S da proteína Fe ligam oxigênio irreversivelmente, impedindo o acoplamento da proteína Fe com a proteína FeMo, assim bloqueando o fluxo de elétrons para os grupos metálicos do cofator FeMo e assim impedindo a incorporação de hidrogênios no nitrogênio. Portanto, um ambiente relativamente anóxico é necessário para a atividade da nitrogenase. Um importante mecanismo para contornar a sensibilidade ao oxigênio é a formação de nódulos nas raízes das leguminosas.

Estes nódulos abrigam o rizóbio no seu interior, em um micro-habitat relativamente anóxico que favorece a ação da nitrogenase (Dixon & Kahn, 2004). No nódulo, a planta e o rizóbio sintetizam conjuntamente a leg-hemoglobina, proteína responsável por sequestrar oxigênio e ao mesmo tempo provê-lo em pequenas quantidades para o metabolismo do rizóbio (Ott et al., 2005).

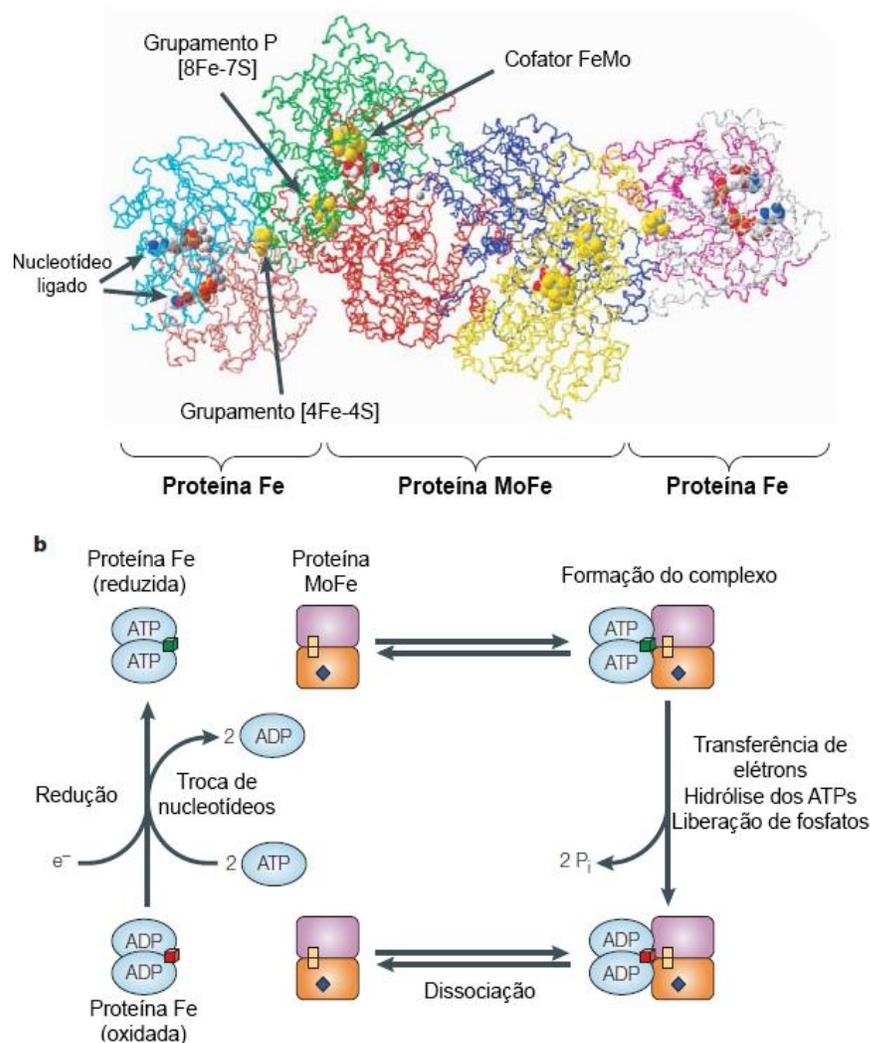


Figura 2 – Estrutura do complexo da nitrogenase (a) e seu mecanismo de ação (b). a: As proteínas Fe formam um complexo com a proteína MoFe através de seus grupamentos 4Fe-4S. O grupamento P (8Fe-7S) transfere os elétrons da proteína Fe para o cofator FeMo, onde ocorre a fixação do nitrogênio. b: A proteína Fe oxidada é reduzida ao custo de dois ATPs, e se complexa com a proteína MoFe, transferindo-lhe dois elétrons e voltando a forma oxidada, sendo então dissociada da proteína MoFe. Adaptado de (Dixon & Kahn 2004).

A regulação transcricional do complexo da nitrogenase pode variar entre organismos. Em geral, o gene *nifA*, presente nas proteobactérias diazotróficas, é o responsável pela ativação da transcrição dos demais genes *nif* que são traduzidos nos componentes do complexo da nitrogenase

(Dixon & Kahn, 2004). Em *Klebsiella pneumoniae*, um exemplo de bactéria diazotrófica de vida livre, a regulação da expressão de *nifA* (Figura 3a) se dá pelas proteínas GlnD, GlnB, NtrB, NtrC, GlnK e NifL (Dixon & Kahn, 2004; Zhang et al., 2005). A proteína GlnD, sensível ao nitrogênio pelo nível de glutamina na célula, tem como função a uridilação e desuridilação de GlnB quando há falta ou excesso de glutamina, respectivamente (Zhang et al., 2005). GlnB, no seu estado desuridilado, interage com NtrB, que desfosforila NtrC e bloqueia o restante da cascata. Já no seu estado uridilado, GlnB não interage com NtrB, o que faz com que NtrB e fosforile NtrC. NtrC fosforilado é um promotor da transcrição de *nifA* e *glnK*. NifA possui regulação pós-traducional por NifL, mediada por interação proteína-proteína (Martinez-Argudo et al., 2004). Quando oxigênio é detectado por NifL, provavelmente, através de um domínio PAS (Zhulin et al., 1997), há formação de um complexo NifA-NifL, que bloqueia a atividade de NifA. GlnK, um homólogo de GlnB, quando uridilado, interage com NifL para desfazer o complexo NifA-NifL.

Em *Sinorhizobium meliloti*, uma bactéria diazotrófica endossimbionte, a transcrição de *nifA* é regulada por disponibilidade de oxigênio ao invés de glutamina (Figura 3b). Na ausência de oxigênio, FixL se auto-fosforila e transfere o grupo fosfato para FixJ (Tuckerman et al., 2002), que no estado fosforilado é um ativador transcricional de *nifA* e *fixK*. Na presença de oxigênio, FixL perde a função de quinase, bloqueando o restante da cascata. FixK é um ativador transcricional de *fixT* e outros genes necessários para crescimento em condições micro-aeróbicas. FixT é um inibidor da atividade quinase de FixL, regulando negativamente a transcrição de *nifA* em condições micro-aeróbicas (Garnerone et al., 1999). Na regulação pós-transcricional deste modelo não há regulação por NifL e o oxigênio inibe NifA diretamente. Conforme mencionado acima, esta cascata pode variar entre espécies: em *Azorhizobium caulinodans* FixK, ao invés de FixJ, ativa a expressão de *nifA* diretamente (Kaminski et al., 1991).

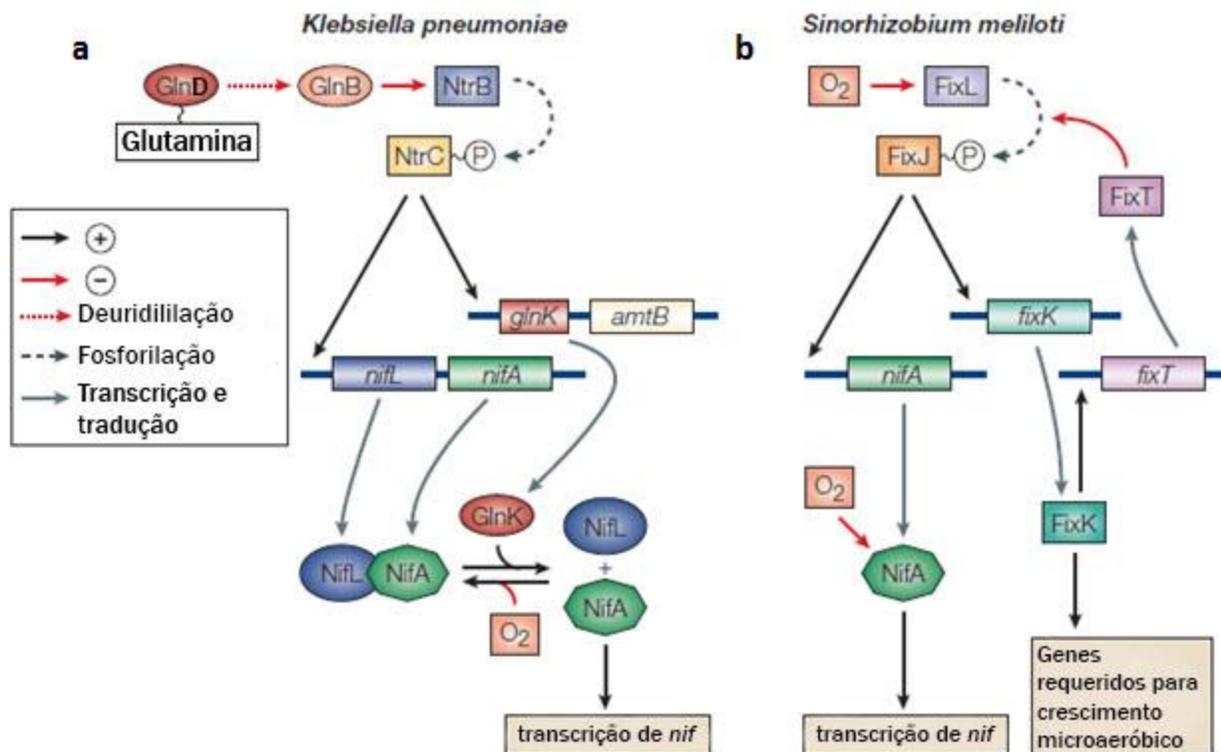


Figura 3 – Regulação do complexo da nitrogenase. a: Regulação em *K. pneumoniae*, sendo iniciada por GlnD. Na presença de glutamina, GlnD desuridilila GlnB, que interage com NtrB, defosforilando NtrC e bloqueando o restante da cascata. Na ausência de glutamina, GlnD uridilila GlnB, que interage com NtrB fosforilando NtrC, que é um promotor de *nifL*, *nifA*, *glnK*. NifL, na presença de oxigênio complexa-se com NifA e bloqueia a transcrição de *nif*. Na ausência de oxigênio, NifL descomplexa-se de NifA pela ação de GlnK, permitindo que NifA promova a transcrição dos outros genes *nif*. b: Regulação em *S. meliloti*, sendo iniciada por FixL. Na presença de oxigênio, FixL não fosforila FixJ, bloqueando o restante da cascata. Na ausência de oxigênio, FixL auto-fosforila e fosforila FixJ, promovendo a transcrição de *nifA* e *fixK*. NifA é regulado diretamente por oxigênio, promovendo a transcrição dos genes *nif* na sua ausência. Este sistema é auto-regulado indiretamente por FixK, que promove a transcrição de *fixT*, que inibe a fosforilação de FixJ. Adaptado de (Dixon & Kahn, 2004).

1.3 SOLUBILIZAÇÃO DE FOSFATO

O fósforo é conhecido por ser o segundo nutriente mineral que mais limita o crescimento vegetal nos trópicos, atrás apenas do nitrogênio (Alori et al., 2017), justificando o intenso uso de fertilizantes fosfatados na agricultura. Estima-se que 30 milhões de toneladas destes fertilizantes sejam aplicados anualmente pelo mundo (Adnan et al., 2017). No Paquistão, este tipo de fertilizante costuma ser utilizado em excesso devido sua baixa disponibilidade no solo para os vegetais, tipicamente em torno de 20% (Adnan et al., 2017), raramente excedendo 30% (Sharma

et al., 2013). A maior porção do fertilizante aplicado torna-se indisponível para absorção vegetal em decorrência de sua precipitação, ao reagir com cátions de alumínio e ferro da fase líquida do solo, e pela adsorção do fosfato aos minerais da fase sólida do solo (Sharma et al., 2013). Ao atingir o limite de carga do solo, o fosfato excedente pode ser transportado para o lençol freático e, conseqüentemente, para a água de superfície, afetando sua qualidade e agravando a eutrofização (Mabilde et al., 2017). Além disso, a presença de fertilizantes fosfatados em excesso nos campos de cultivo causam perturbações na microbiota, reduzindo a fertilidade do solo e a produção agrícola (Sharma et al., 2013).

Estima-se que as plantas consigam absorver anualmente 1% do fosfato presente no solo. Logo, com a disponibilização do fosfato atualmente imobilizado seria, teoricamente, possível nutrir os vegetais por 100 anos (Sharma et al., 2013). Os solos brasileiros possuem alta quantidade de ferro e sendo assim, uma alta capacidade de carga de fosfato, apresentando então potencial para disponibilização de fosfato. As bactérias solubilizadoras de fosfato (PSB) possuem um papel protagonista na disponibilização de fosfato imobilizado. Estas bactérias estão distribuídas em diversos gêneros, como *Burkholderia* (Simonetti et al., 2018), *Gluconacetobacter* (Intorne et al., 2009), *Erwinia*, *Ralstonia*, *Rhizobium* (Iyer et al., 2017), entre outros.

A matéria orgânica pode concentrar de 20 a 80% do fosfato total do solo na forma imobilizada (Sharma et al., 2013), apresentando grande potencial de biodisponibilização pelas PSB. A teoria do dreno explica que a biomassa de fósforo dos microrganismos é diretamente correlacionada com a decomposição de fósforo de substratos orgânicos (Alori et al., 2017; Khan et al., 2010), evidenciando a importância desta fonte de fósforo para a biodiversidade. As fontes de matéria orgânica que possuem fósforo em sua estrutura são muito diversas, de constituintes celulares como ácidos nucleicos e fosfolipídios à xenobióticos fosfatados (e.g. pesticidas, antibióticos, aditivos de detergente). Para que as plantas acessem o fósforo presente na matéria orgânica, é desejável a ação das PSB na geração fosfato iônico solúvel ou compostos orgânicos fosfatados de baixo peso molecular. Esta transformação demanda que o fosfato orgânico seja solubilizado e sofra a ação enzimática para liberar os grupamentos fosfatos. Devido a variação de pH indo do ácido ao neutro, durante a disponibilização de fosfato orgânico, a ação deste processo é dominado por fosfatases ácidas inespecíficas (Rodriguez & Fraga, 1999), que clivam os grupamentos fosfatos da matéria orgânica. Outra enzima importante na disponibilização de fosfato

é a fitase (Jorquera et al., 2011). Esta enzima age sobre o fitato, também chamado mio-inositol hexaquisfosfato. Este composto é derivado principalmente de vegetais e pode ser responsável por até 80% do fosfato orgânico do solo (Turner et al., 2002). Vegetais também podem liberar fosfato do fitato, embora a eficiência por microrganismos seja superior (Richardson & Simpson, 2011).

O principal mecanismo responsável pela solubilização de fosfato inorgânico é a síntese de compostos solubilizadores de minerais, como sideróforos, ácidos orgânicos, hidroxilas, carboxilas e prótons. Os ácidos orgânicos, produzidos em conjunto com as carboxilas e hidroxilas, quelam os cátions, reduzindo o pH e liberando fósforo. Em alguns organismos, é possível que o efluxo de prótons H^+ seja mais determinante na solubilização de fosfato do que a emissão de ácidos orgânicos (Illmer & Schinner, 1995). Este efluxo pode decorrer da assimilação de NH_4^+ (Asea et al., 1988), o que favorece formulações de inoculantes contendo uma mistura de solubilizadores de fosfato com bactérias diazotróficas. A emissão de ácidos inorgânicos também apresenta potencial para solubilizar fosfato, apesar de serem menos eficientes do que os ácidos orgânicos (Alori et al., 2017).

Dos ácidos orgânicos, o ácido glucônico é o agente de solubilização mais produzido pelas PSB. Seu mecanismo de ação se baseia em quelar os cátions ligados ao fosfato, tornando-o disponível para a assimilação vegetal. O cofator redox pirroloquinolina quinona (PQQ) atua na glicose desidrogenase, que então oxida a glicose em ácido glucônico. Nas bactérias Gram-negativas, esta etapa ocorre no espaço periplasmático (Alori et al., 2017), facilitando a exportação do ácido glucônico para o ambiente. O cluster de genes *pqqABCDE* é responsável pela síntese do cofator PQQ em diversas bactérias do solo (Chauhan et al., 2017; Oteino et al., 2015), sendo os genes *pqqACDE* essenciais: PqqA serve de substrato para PqqE; PqqC é responsável pela etapa final da síntese do PQQ; PqqD interage fisicamente com PqqE, mas a função exata é desconhecida; PqqE cliva o radical S-adenosil-metionina (SAM) em metionina e 5'-desoxiadenosina (Oteino et al., 2015; Shen et al., 2012).

1.4 AUXINA E CITOCININA

Os fitormônios auxina (ácido indolacético, IAA ou similares) e citocinina (trans-zeatina ou similares) são cruciais ao desenvolvimento e crescimento de plantas. Não foram encontradas variedades mutantes e funcionais para estas classes de hormônios, sugerindo que sua ação é

essencial à todas as plantas. Estes hormônios são produzidos de forma relativamente contínua nos tecidos vegetais, uma vez que estes são moduladores de funções vitais como divisão e expansão celular (Taiz & Zeiger, 2010).

As auxinas são conhecidas por promoverem o alongamento dos coleótilos e segmentos caulinares e também por induzir o desenvolvimento de raízes adventícias em folhas ou caules excisados. As citocininas são conhecidas por estimularem divisão celular, retardar senescência foliar e promover a expansão de cotilédones em dicotiledôneas (Taiz & Zeiger, 2010). Uma importante aplicação biotecnológica da auxina e da citocinina está relacionada a regeneração *in vitro* de tecidos vegetais, que é modulada pelo balanço da concentração destes no meio. Quando a concentração de auxina for superior à de citocinina, haverá desenvolvimento de raízes, enquanto que na condição inversa, o desenvolvimento de parte aérea é favorecido. Em situação de equimolaridade, haverá crescimento de um calo indiferenciado (Skoog & Miller, 1957).

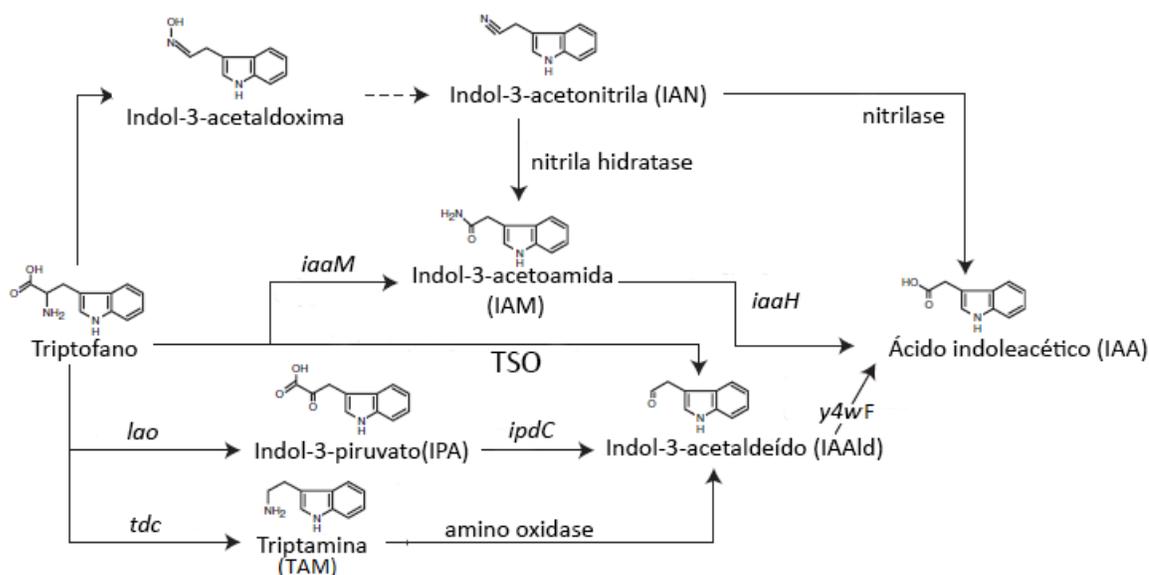


Figura 4 – Vias de biossíntese de auxina dependentes de triptofano. Os genes conhecidos envolvidos nas rotas estão em itálico. O triptofano pode ser metabolizado em IAA por cinco vias: IAM, IAN, IPA, TAM, e TSO. Na IAM, o triptofano é convertido em IAM, que é hidrolisado por *IaaH* em IAA. Na via do IAN, o triptofano é convertido em indol-3-acetaldoxima, que é convertido em IAN. O IAN pode então ser desnitrificado em IAA por nitrilase ou ser convertido em IAM por nitrila hidratase. Na IPA, o triptofano é transaminado por *LAO* em IPA, que é descarboxilado por *IpdC* em IAAld, que é desidrogenado por *Y4wF* em IAA. Na TAM, o triptofano é descarboxilado por *TDC* em triptamina, que é convertida em IAAld por uma amino oxidase. Na TSO, o triptofano é convertido diretamente em IAAld. A partir do IAAld, as vias TAM e TSO compartilham o mesmo caminho metabólico que IPA. Esquema adaptado de (Imada et al., 2017; Spaepen & Vanderleyden, 2011; Spaepen et al., 2007).

Em bactérias, as auxinas podem ser sintetizadas por diversas vias dependentes de triptofano (Figura 4), como a via do indol-3-acetoamida (IAM), indol-3-piruvato (IPA), triptamina (TAM/TRM), indol-3-acetonitrila (IAN) e via de oxidação das cadeias laterais de triptofano (TSO, *tryptophan side-chain oxidase*). Na via do IAM, o triptofano é convertido pela triptofano monooxigenase (gene *iaaM*) para IAM, que é hidrolisado em IAA e amônia pela IAM hidrolase (gene *iaaH*). Na via do IPA, o triptofano é convertido por uma transaminase (gene *lao*) em IPA, que é descarboxilado em indol-3-acetaldeído (IAAld) pela enzima indol-3-piruvato descarboxilase (gene *ipdC*). O IAAld é oxidado em IAA por uma desidrogenase (gene *y4wF*). Na via da TAM, o triptofano é descarboxilado por uma triptofano descarboxilase (gene *tdc*) em triptamina, que é convertida em IAAld por uma amino oxidase. Na via do IAN, o triptofano é convertido em indol-3-acetaldoxima, que é convertido em IAN. O IAN pode então ser convertido em IAA por nitrilase ou ser convertido em IAM por nitrila hidratase. Na TSO, o triptofano é convertido diretamente em IAAld, e a partir deste metabólito segue pela mesma via do IPA (Imada et al., 2017; Spaepen & Vanderleyden, 2011; Spaepen et al., 2007).

Diversas bactérias podem sintetizar diretamente ou estimular a produção de auxinas e citocininas nas plantas. O rizóbio produz auxinas por múltiplas rotas, possivelmente, alterando o balanço de auxina na região dos nódulos e estimulando o seu desenvolvimento (Mathesius, 2008). Além disso, PGPR com genes de síntese de IAA mutados apresentaram redução em sua capacidade de promover o enraizamento (Dobbelaere et al., 1999), sugerindo que IAA participa do estímulo ao enraizamento, em conjunto com outros fatores (Spaepen & Vanderleyden, 2011).

A citocinina bacteriana da classe trans-zeatina, produzida por *Bacillus subtilis* IB-22, possui um efeito de estímulo à produção de exsudados nas raízes (Kudoyarova et al., 2014). As citocininas zeatina e ribosil zeatina, produzidas pelas cepas *Bacillus licheniformis* Am2, *B. subtilis* BC1 e *Pseudomonas aeruginosa* E2, estimulam cotilédones de pepino, aumentando seu tamanho, massa fresca e divisão celular (Hussain & Hasnain, 2009). Uma proteína similar a LOG, enzima ativadora de citocininas identificada em arroz (Kurakawa et al., 2007), foi documentada em genomas bacterianos (Seo et al., 2016), embora sua função permaneça pouco conhecida em bactérias.

1.5 ETILENO E ACC DESAMINASE

A produção de etileno é uma resposta comum em plantas a diversas situações de estresse biótico e abiótico, como o estresse salino (Liu et al., 2017), solo contaminado por metais (Keunen et al., 2016), estresse hídrico (Narayana et al., 1991), exposição a altas temperaturas (Maduraimuthu & Prasad, 2010), infecção por nematoides (Glazer et al., 1985), fungo (Chen et al., 2003) e bactérias (Lenarcic et al., 2017). Esta resposta se dá inicialmente com a síntese de auxina, que sinaliza crescimento vegetal e produção de SAM, que é convertido pela ACC sintase em ACC, que, por sua vez, é oxidado em etileno pela ACC oxidase. O etileno inibe os fatores de resposta a auxina, freando o crescimento vegetal (Glick, 2014).

Um modelo de atividade do etileno (Glick, 2014; Pierik et al., 2006; van Loon et al., 2006) preconiza que seu efeito se dá em dois picos de produção: no primeiro pico, os genes de defesa e proteção da planta são ativados; já no segundo pico, há uma produção de etileno muito superior à do primeiro, desencadeando reações deletérias ao crescimento vegetal, como clorose e senescência.

A enzima ACC desaminase, responsável pela clivagem de ACC em α -cetobutarato e amônia (Honma & Shimomura, 1978), é encontrada em bactérias, fungos e algumas plantas (Plett et al., 2009). O ACC é o precursor direto do etileno pela ação da enzima ACC oxidase. Portanto, a enzima ACC desaminase compete pelo substrato com a enzima ACC oxidase (Figura 5). A produção de ACC desaminase se inicia com o primeiro pico de etileno, sem interferir no seu efeito de proteção (Robison et al., 2001). Isto se deve ao fato da ACC desaminase ter menor afinidade pelo substrato do que a ACC oxidase, sendo necessária uma quantidade superior de ACC desaminase para que haja escoamento do ACC. Esta quantidade necessária de ACC desaminase é alcançada apenas no segundo pico de etileno, podendo reduzir a atividade deletéria ao crescimento vegetal causada pelo etileno de 50% até 90% (Glick et al., 1998). Além disso, bactérias que sintetizam ACC desaminase podem também sintetizar auxina (Figura 5), contribuindo significativamente para a promoção do crescimento vegetal.

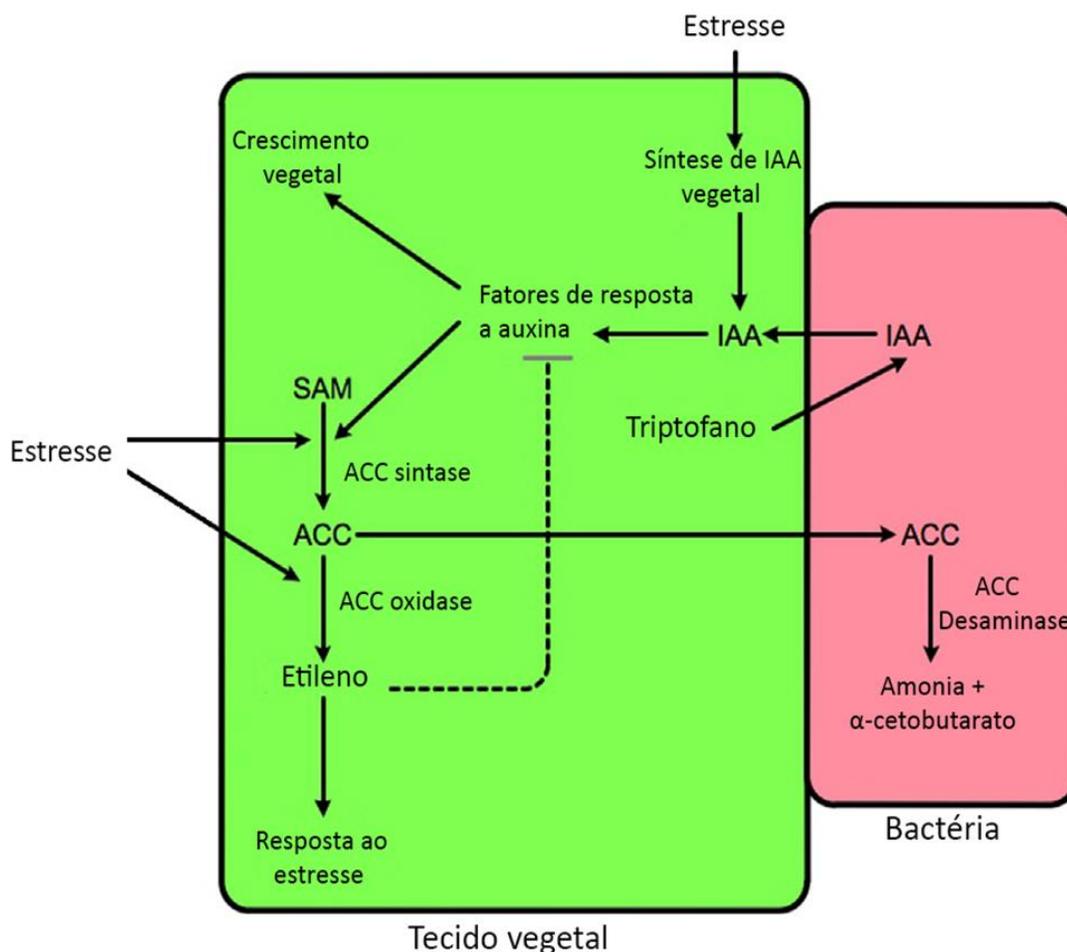


Figura 5 – Esquema de desvio da rota de etileno mediado por ACC desaminase. Em condições de estresse, o vegetal sintetiza auxina, que sinaliza a produção de SAM, que é convertido em ACC pela enzima ACC sintase. O ACC é convertido em etileno pela ACC oxidase. Neste sistema de interação planta-bactéria, a bactéria produz a enzima ACC desaminase, que cliva o ACC em α -cetobutarato e amônia, desviando a rota de síntese do etileno, ao mesmo tempo que produz auxina a partir do triptofano vegetal. Adaptado de (Glick, 2014)

1.6 CEPAS UENF114111 E UENF412522

Devido a demanda por tecnologias de fertilização de baixo impacto ambiental, foi desenvolvido um produto chamado inoculante agrícola. Os inoculantes agrícolas são formulações que contêm uma espécie ou mais de microrganismos que exercem as propriedades citadas anteriormente, como fixação de nitrogênio e solubilização de fósforo (Bashan, 1998). Estes microrganismos normalmente são isolados de sistemas solo-planta. Dentre estes sistemas, as rizosferas de fruteiras no Brasil foram pouco exploradas para prospecção de novas cepas com potencial de integrar novos inoculantes.

Nosso grupo de pesquisa (Rodrigues et al., 2016) vem se utilizando de uma abordagem (meio de cultura semi-sólido, isento de nitrogênio inorgânico ou orgânico) bastante eficiente para isolar, de fruteiras tropicais (mamão, goiaba, maracujá e abacaxi), uma série de microrganismos com potencial para fixação de nitrogênio, solubilização de fosfato, zinco e produção de auxinas. Segundo a informação do rRNA 16S, os microrganismos isolados estão distribuídos nos gêneros *Azospirillum*, *Burkholderia*, *Herbaspirillum* e *Pseudomonas*.

As cepas UENF114111 e UENF412522, isoladas por (Rodrigues et al., 2016), foram identificadas pelo rRNA16S como *Burkholderia* sp. e *Azospirillum* sp, respectivamente. Análises bioquímicas revelaram que UENF114111 e UENF412522 são capazes de metabolizar 26 e 24 diferentes fontes de carbono. Plântulas de abacaxi inoculadas com a cepa UENF114111 apresentaram aumento da massa fresca do sistema radicular em relação ao controle. Os testes de inoculação para a cepa UENF412522 em maracujá apontam aumento de comprimento, altura e massa fresca em relação ao controle. Os testes bioquímicos (tabela 1) apontam que a cepa UENF114111 solubiliza zinco e que ambas as cepas solubilizam fosfato, fixam de nitrogênio e não produzem auxinas (Rodrigues et al., 2016). Este conjunto de características de promoção do crescimento vegetal justifica então o sequenciamento dos genomas destas cepas pelo nosso grupo de pesquisa, sendo prospectados os genes responsáveis por estas e outras propriedades.

Para conduzir o sequenciamento, as amostras de DNA foram extraídas utilizando o kit Qiagen, a qualidade da extração foi verificada por eletroforese na plataforma Agilent Technologies Bioanalyzer 2100 e as amostras com qualidade suficiente foram liofilizadas no liofilizador Liotop K105. O sequenciamento foi conduzido na plataforma Illumina HiSeq2500 localizada no LaCTAD – Unicamp (Campinas, SP). Os *reads* provenientes deste sequenciamento são a base para as análises genômicas deste trabalho.

| Cepa | Identificação pelo rRNA 16s | Vegetal – local de coleta | Produção de auxina | Solubilização de fosfato | Solubilização de zinco | Fixação de nitrogênio |
|-------------|------------------------------------|----------------------------------|---------------------------|---------------------------------|-------------------------------|------------------------------|
| UENF114111 | <i>Burkholderia sp.</i> | Abacaxi – superfície do tronco | Não | Sim | Sim | Sim |
| UENF412522 | <i>Azospirillum sp.</i> | Maracujá – rizoplano | Não | Sim | Não | Sim |

Tabela 1 – Identificação, vegetal hospedeiro e propriedades bioquímica relacionadas a promoção do crescimento vegetal nas cepas UENF114111 e UENF412522 (Rodrigues et al., 2016).

2. JUSTIFICATIVA

As cepas UENF114111 e UENF412522 apresentam algumas características importantes do ponto de vista de seu potencial biotecnológico. Contudo, pouco se sabe acerca das bases genéticas destas características. Uma das maneiras de melhor explorar estas cepas de forma sistemática é o sequenciamento genômico. Desse modo, visa-se identificar não apenas os genes envolvidos com os processos identificados experimentalmente, como também prospectar novas potencialidades desconhecidas nestas cepas e que possam promover o crescimento vegetal. Este tipo de abordagem é comum em trabalhos de genômica, envolvendo a montagem de genoma, predição de genes, reconstrução filogenética e prospecção de genes envolvidos na promoção do crescimento vegetal (Duan et al., 2013; Gupta et al., 2014; Pedrosa et al., 2011).

3. OBJETIVO

O objetivo geral deste trabalho é identificar as bases genéticas responsáveis pela promoção do crescimento vegetal (tanto das características listadas nos testes bioquímicos como fenótipos ainda não verificados) e colaborar com o entendimento da fisiologia das cepas UENF114111 e UENF412522. Para isto, são propostos os seguintes objetivos específicos:

1. Montagem do genoma das cepas UENF114111 e UENF412522 a partir de dados brutos de sequenciamento;
2. Realizar predição gênica nos genomas montados;
3. Anotar funcionalmente os genes preditos;
4. Reconstruir as árvores filogenéticas das cepas UENF412522 e UENF114111;

5. Curar manualmente os genes potencialmente envolvidos com a promoção do crescimento em vegetal (PGPG, *Plant Growth-Promoting Genes*).

4. MÉTODOS

A qualidade e quantidade de *reads* provenientes do sequenciador Illumina HiSeq2500 foram avaliadas pelo software fastQC 0.11.5. Os *reads* foram selecionados e filtrados por qualidade com o software trimmomatic 0.36 (Bolger et al., 2014), alterando os parâmetros *avgqual* para 30 e *slidingwindow* para 35.

Para a montagem do genoma, foram utilizados os montadores SPADes 3.8 (Bankevich et al., 2012), IDBA 1.1.1 (Peng et al., 2010), Velvet 1.2 (Zerbino & Birney, 2008) e Minia 1.0 (Chikhi & Rizk, 2013). O principal parâmetro testado foi o valor de *kmer* (k). Nos montadores SPADes e IDBA é possível utilizar múltiplos valores de k na montagem, diferentemente dos montadores Minia e Velvet que trabalham apenas com um valor de k por montagem. Os *contigs* resultantes foram unidos em *scaffolds* pelo software SSPACE 3.0 (Boetzer et al., 2011). O resultado de cada montagem foi analisado pelo o software QUAST 3.0 (Gurevich et al., 2013). As melhores montagens tipicamente possuem menor número de contigs, além de apresentarem valores de L50 e L75 e com maiores valores de N50 e N75. As métricas L50 e L75 dizem qual o número mínimo de *scaffolds* necessário para cobrir metade (L50) ou 75% (L75) do genoma. Com os *scaffolds* ordenados do maior ao menor, as métricas N50 e N75 indicam o tamanho do *scaffold* que está na metade da cobertura do genoma (N50) e o tamanho do *scaffold* na posição correspondente a 75% da cobertura (N75). Os genomas foram inspecionados manualmente a partir de ideogramas gerados com o software CIRCOS (Krzywinski et al., 2009).

Cada cepa sequenciada foi identificada pela informação do rRNA16S recuperada diretamente dos *reads* pelo software EMIRGE (Miller et al., 2011) e por uma busca de correlação tetra contra a versão 35 do *Ensembl Bacteria* utilizando a plataforma JSpeciesWS (Richter et al., 2016). Esta análise distribui em um gráfico a frequência de 256 categorias de tetranucleotídeos (número máximo de combinação de sequências de quatro nucleotídeos), e gera uma curva de regressão. Curvas de diferentes genomas podem ser comparadas e, quanto maior a sobreposição das curvas, maior o Z-score e a similaridade dos genomas (Richter & Rossello-Mora, 2009).

As reconstruções filogenéticas das cepas UENF412522 e UENF114111 foram realizadas a partir de comparações com os genomas mais próximos disponíveis no Genbank. Estes genomas, juntamente com os genomas de UENF412522 e UENF114111, foram anotados com a ferramenta Prokka (Seemann, 2014). Os resultados obtidos foram analisados com a ferramenta Roary (Page et al., 2015). Neste software, utilizando resultados de BLASTP com identidade maior ou igual à 70%, foram gerados alinhamentos de *core genes* (genes conservados em todos os genomas comparados). Os polimorfismos de nucleotídeos únicos (SNPs) foram extraídos dos alinhamentos utilizando o software SNP-sites (Page et al., 2016). Os alinhamentos dos SNPs foram então submetidos ao software RaxML 8.2.11 (Stamatakis, 2014) para se realizar a reconstrução filogenética por máxima verossimilhança. O modelo de substituição utilizado foi ASC_GTRCAT com 100 *bootstraps*. Os grupos externos utilizados na reconstrução filogenética de *Azospirillum* e *Paraburkholderia* foram *Rhodospirillum centenum* SW (Munk et al., 2011) e *Burkholderia thailandensis* E254 (Johnson et al., 2015), respectivamente. Na comparação entre *Paraburkholderia* e *Ralstonia*, o grupo externo utilizado foi *Polynucleobacter necessarius* QLW-P1DMWA-1 (Zhu et al., 2011). Estes grupos externos foram escolhidos por apresentaram alta afinidade filogenética com os genomas de interesse.

As melhores montagens tiveram a predição automática dos genes pelo anotador Prokka (Seemann, 2014). Os genes potencialmente envolvidos com a promoção do crescimento em vegetal foram anotados manualmente a partir dos ORFs (*Open Reading Frame*) preditos pelo Prokka, com o uso da ferramenta BLASTP (Altschul et al., 1997) utilizando a base curada de sequências SwissProt (Boutet et al., 2016). A anotação de um ORF foi aceita nos casos de e-value menor que $1e^{-5}$ e conservação de domínios proteicos com a sequência de referência do SwissProt (Gupta et al., 2014). As informações disponíveis sobre as proteínas contidas no SwissProt e nas publicações referentes a estas sequências foram utilizadas para inferir a função dos ORFs das cepas UENF114111 e UENF412522. As ORFs foram mapeadas nas rotas metabólicas do KEGG utilizando a ferramenta blastKOALA (Kanehisa et al., 2016). Para a anotação funcional dos *reads*, estes foram alinhados pelo blastX a base de dados *nr* utilizando o software Diamond (Buchfink et al., 2015) e a plataforma Megan 6.6.7 (Huson et al., 2016) para posterior classificação nas categorias funcionais do SEED (Overbeek et al., 2014). Os domínios proteicos conservados foram analisados com a ferramenta CDD-NCBI (Marchler-Bauer et al., 2017).

5. RESULTADOS E DISCUSSÃO

5.1 Montagem dos genomas de UENF412522 e UENF114111

A qualidade das montagens de genomas pode variar de acordo com o genoma e o montador utilizado. Neste trabalho, metodologias de multi-k (SPAdes e IDBA) ou apenas um valor de k (Velvet e Minia) foram comparadas na montagem dos genomas das cepas UENF412522 e UENF114111.

Na montagem com o SPAdes 3.8 as melhores montagens foram obtidas utilizando valores de k 33, 45, 57, 63, 69, 73, 77, 83. Na montagem com o IDBA foram utilizados valores de k 20, 26, 32, 38, 44, 50, 56, 62, 68, 74, 80. Na montagem com Velvet foi utilizado o script VelvetOptimiser 2.2.5, apresentando melhores métricas para k = 61 e *coverage cutoff* igual a metade da cobertura do genoma. Na montagem com Minia, o melhor valor de k foi definido pelo software KMERgenie como 61 ambas as cepas.

| | SPAdes | | IDBA | | Velvet | | Minia | |
|---------------|---------------|---------------|-------|-------------|--------|--------|-------|-------|
| UENF114111 | Ctg | Scf | Ctg | Scf | Ctg | Scf | Ctg | Scf |
| #contigs | <u>137</u> | <u>104</u> | 215 | 115 | 467 | 278 | 697 | 377 |
| Contig Maior | <u>280k</u> | 737k | 187k | <u>777k</u> | 175k | 317k | 102k | 212k |
| Tamanho total | 6.71m | 6.72m | 6.68m | 6.72m | 6.69m | 6.72m | 6.63m | 6.72m |
| N50 | <u>113.2k</u> | <u>122.6k</u> | 79.1k | 116.7k | 26.5k | 50.2k | 19.5k | 34k |
| N75 | <u>64.6k</u> | <u>84k</u> | 40.8k | 71.6k | 15k | 24.6k | 9.1k | 18.3k |
| L50 | <u>20</u> | <u>16</u> | 28 | <u>16</u> | 72 | 38 | 102 | 54 |
| L75 | <u>40</u> | <u>32</u> | 57 | 34 | 153 | 86 | 225 | 121 |
| #N por 100kbp | 0.01 | 32.13 | 0 | 196 | 0.01 | 304.28 | 0 | 661 |

Tabela 2 – Métricas da montagem para cepa UENF114111 utilizando os montadores SPAdes, IDBA, Velvet e Minia. Em Ctg estão evidenciadas métricas para os *contigs* e em Scf para os *Scaffolds*. Os valores indicando as melhores métricas estão sublinhados.

As métricas de contiguidade (N50, N75, L50, L75, tabela 2 e 3) foram melhores para o montador SPAdes, seguido pelo IDBA. Isto se deve a método empregado por estes montadores, que permitem o uso de múltiplos valores de k, que são todos incrementados em uma única montagem. Montadores mais antigos, como Velvet e Minia, utilizam apenas um valor de k por montagem, o que atualmente é considerado obsoleto em montagem de genomas bacterianos.

| | SPAdes | | IDBA | | Velvet | | Minia | |
|---------------|---------------|---------------|-------|--------|--------|--------|-------|--------|
| UENF412522 | Ctg | Scf | Ctg | Scf | Ctg | Scf | Ctg | Scf |
| #contigs | <u>125</u> | <u>123</u> | 224 | 124 | 532 | 168 | 755 | 223 |
| Contig Maior | <u>474k</u> | <u>557k</u> | 374k | 431k | 91.2k | 265.4k | 77.9k | 332.4k |
| Tamanho total | 7.83m | 7.83m | 7.82m | 7.87m | 7.7m | 7.78m | 7.7m | 7.85m |
| N50 | <u>175.2k</u> | <u>179.2k</u> | 99.8k | 155.6k | 28k | 102.5k | 19k | 77.4k |
| N75 | <u>109.4k</u> | <u>111.9k</u> | 52.1k | 104.5k | 14.4k | 56.4k | 10.4k | 38.8k |
| L50 | <u>15</u> | <u>14</u> | 23 | 16 | 85 | 23 | 121 | 29 |
| L75 | <u>30</u> | <u>28</u> | 51 | 31 | 182 | 46 | 254 | 65 |
| #N por 100kbp | 0 | 2.32 | 0.73 | 200 | 0.12 | 541.6 | 0.1 | 1028 |

Tabela 3 – Métricas da montagem para cepa UENF412522 utilizando os montadores SPAdes, IDBA, Velvet e Minia. Em Ctg estão evidenciadas métricas para os *contigs* e em Scf para os *Scaffolds*. Os valores indicando as melhores métricas estão sublinhados.

5.2 Identificação filogenética das cepas UENF412522 e UENF114111

As sequências do rRNA 16S recuperadas pelo EMIRGE para a cepa UENF412522 apresentaram alta identidade (>95%) com sequências publicamente disponíveis do gênero *Azospirillum* (e.g.: *Azospirillum lipoferum* B2 e BL2, *Azospirillum melinis* LMG24250 e TMCY41, *Azospirillum oryzae* KNUC9025, além de vários *Azospirillum* sp.). Para a cepa UENF114111, a maioria das correspondências foram com *Burkholderia* (*Burkholderia silvatlantica* AB219, *Burkholderia denitrificans* UENF114111 e UENF11111, *Burkholderia mimosarum* PTK31), além de algumas *Paraburkholderia* (*Paraburkholderia ferrariae* FeG101, *Paraburkholderia*

silvatlantica SRMrh-20). Recentemente, o gênero *Burkholderia* foi desmembrado em dois: *Burkholderia*, maioria constituintes de bactérias patogênicas e *Paraburkholderia*, com bactérias ambientais (Sawana et al., 2014). Conforme esperado, os bancos de dados ainda não foram completamente atualizados (e.g. *B. silvatlantica*, que foi reclassificada para *Paraburkholderia* (Eberl & Vandamme, 2016)).

A correlação tetra obtida com o JSpeciesWS para a cepa UENF412522 apresentou Z-scores acima do *cutoff* de 0,989 para *A. lipoferum* 4B, *Azospirillum tiophilum* DSM21654 e *Azospirillum sp.* B510, evidenciando maior grau de parentesco com estas cepas do que as de *Azospirillum brasilense* Sp245 e *A. brasilense* Az39. Na filogenia de *Azospirillum* (Figura 6), a cepa UENF412522 forma um grupo irmão com um clado contendo *Azospirillum sp.* B2, *Azospirillum sp.* B506, *A. oryzae* A2P, *Azospirillum humicireducens* SgZ5 e *A. lipoferum* 4B, o que indica que UENF412522 pode ser tanto um isolado de alguma destas espécies, ou mesmo uma nova espécie. Foi encontrado um único clado monofilético representativo que contém apenas uma espécie: *A. brasilense*. De maneira geral, existem poucos genomas de *Azospirillum* depositados (24, incluindo UENF412522), sendo necessário um esforço de sequenciamento para elucidar a filogenia desta espécie.

A correlação tetra para a cepa UENF114111 não apresentou nenhuma correspondência com Z-score acima do *cutoff*: *Paraburkholderia sacchari* 19450 (Z-score de 0,98) foi a que mais se aproximou do *cutoff*, seguida por *Paraburkholderia glathei* DSM50014 (0,96) e *Candidatus Paraburkholderia kirkii* UZHbot2 (0,96). Este dado sugere que a cepa UENF114111 se trata de uma *Paraburkholderia*, sendo necessário revisar a entrada no GenBank (KU836603.1) referente ao 16S desta sequência, que está classificada como *B. denitrificans*.

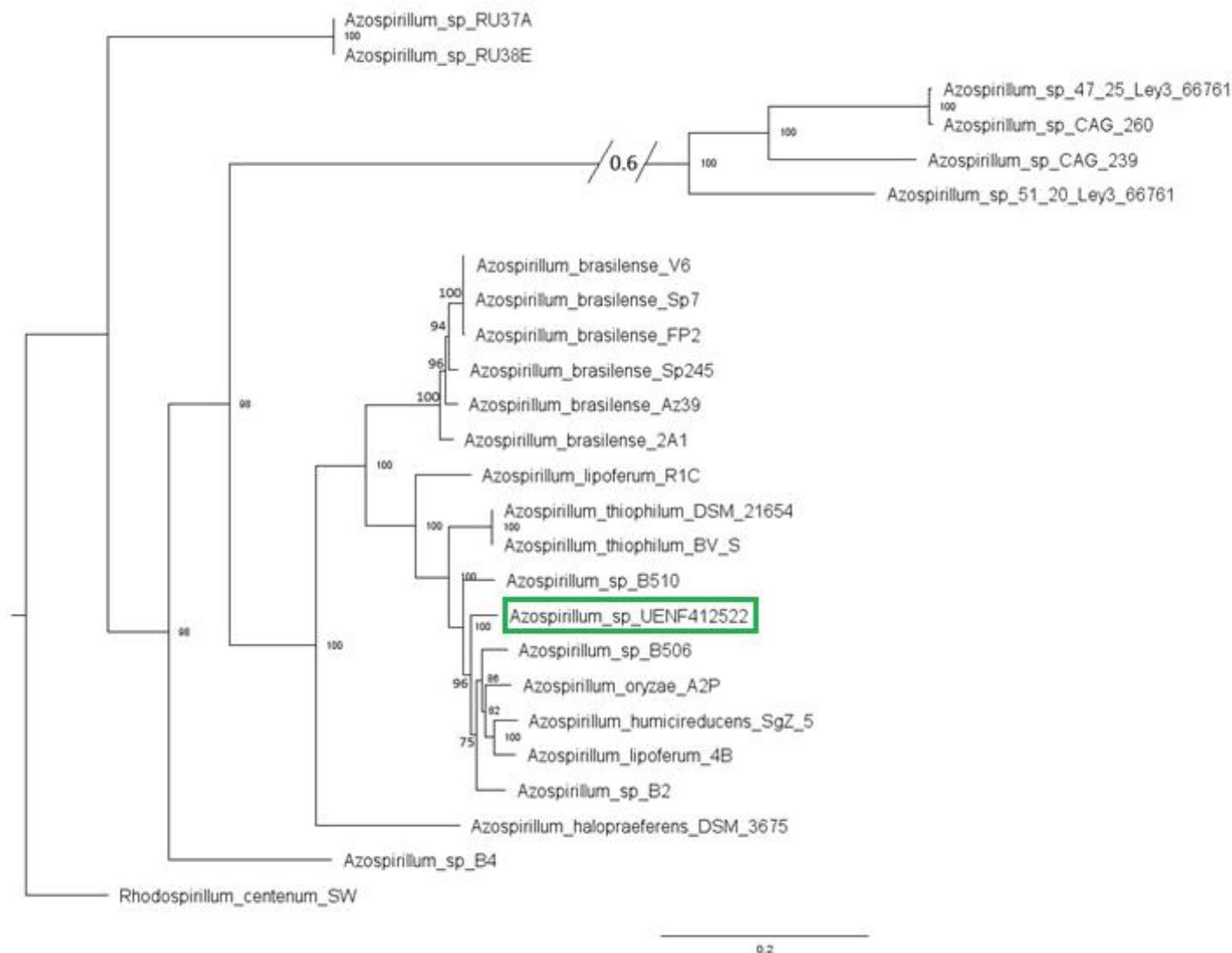


Figura 6 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos *core genes* para as espécies de *Azospirillum*. *R. centenum* SW foi definido como grupo externo. A cepa UENF412522 (destacada em verde) formou um grupo irmão com um clado contendo *Azospirillum sp.* B2, *Azospirillum sp.* B506, *A. oryzae* A2P, *A. humicireducens* SgZ5 e *A. lipoferum* 4B. A única espécie deste gênero contida em um clado monofilético representativo é *A. brasilense*.

Na filogenia de *Paraburkholderia* (Figura 7), a cepa UENF114111 se posicionou como grupo irmão de um clado contendo *Paraburkholderia nodosa*, *Paraburkholderia heleia* e *Paraburkholderia mimosarum*. Sendo assim, é possível que a cepa UENF114111 pertença à alguma dessas três espécies ou mesmo que se trate de uma nova espécie. É possível observar um ramo longo com baixo suporte no interior da filogenia, contendo *Paraburkholderia rhizoxinica* HKI454. Esta cepa foi caracterizada, originalmente, como *Burkholderia rhizoxinica* (Lackner et al., 2011), sendo remanejada para *Paraburkholderia*. A correlação tetra para esta cepa mostrou

uma mistura de correspondências abaixo do *cutoff* de 0,989 para o gênero *Paraburkholderia* e *Ralstonia*, ambas pertencentes à família Burkholderiaceae. Em uma segunda filogenia contendo espécies de *Ralstonia* e *Paraburkholderia* (Figura 8), observa-se que esta cepa é a espécie mais basal dentro de *Paraburkholderia*, seguida por *Paraburkholderia caryophylli* como a segunda espécie mais basal dentro de *Paraburkholderia*. A partir dela, o gênero desmembra-se em dois grupos irmãos: um contendo espécies determinadas como *Candidatus Paraburkholderia* e outro contendo as demais espécies do gênero.

5.3 Predição e anotação gênica

A predição gênica automática para a cepa UENF114111 revelou 5960 ORFs, das quais 4561 (~76,53%) foram anotadas. Já para a cepa UENF412522 foram preditas 6892 ORFs, tendo sido 5266 (~76,41%) anotadas. A anotação funcional dos *reads* para a cepa UENF114111 obteve 40% dos *reads* classificados em 46 categorias funcionais do SEED pela plataforma MEGAN, enquanto que para a cepa UENF412522 obteve-se 37% dos *reads* classificados em 45 categorias funcionais. A anotação manual das ORFs (utilizando a base Swissprot) ajudou a inferir a função de 3592 (~60,27%) e 3825 (~55,50%) genes de UENF114111 e UENF412522, respectivamente. O número de ORFs anotadas por esta abordagem é menor do que as automáticas. Contudo, a qualidade da anotação manual é maior por permitir curadoria com dados da literatura e comparação com proteínas de referência de maior qualidade, disponíveis no SwissProt.



Figura 7 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos *core genes* para as espécies de *Paraburkholderia*. A cepa UENF114111 (destacada em verde) esta como grupo irmão de um clado contendo *P. nodosa*, *P. heleia* e *P. mimosarum*, sem se agrupar diretamente com nenhuma outra cepa. Observa-se um ramo longo no interior da filogenia com baixo suporte contendo *P. rhizoxinica* HKI454. Desconsiderando as relações com suporte menor do que 50% (ramos marcados com triângulo vermelho), temos uma politomia que desmembra em *P. rhizoxinica* HKI454, *P. caryophylli* Ballard720, um clado contendo as espécies com sufixo candidatos (porção inferior) e o clado contendo as demais espécies. A cepa *B. thalandensis* E254 foi definida como grupo externo.

5.4 Genes relacionados à promoção do crescimento vegetal e outros processos ecofisiológicos

Ambas as cepas apresentaram fenótipos relacionados a biofertilização, embora pouco se saiba sobre as suas possíveis bases genéticas relacionadas a promoção do crescimento vegetal. Após a identificação destes genes, seria possível transferi-los para outras bactérias, concedendo-as o respectivo fenótipo. Um exemplo de sucesso desta abordagem foi a transferência do complexo da nitrogenase para *Escherichia coli*, tornando-a uma fixadora de nitrogênio (Vicente & Dean, 2017). Além disto, a investigação genômica pode revelar propriedades de promoção do crescimento vegetal não averiguadas por experimentos.

Para a cepa UENF114111 e UENF412522 foram encontrados diversos genes *nif*, que integram o complexo da nitrogenase, um gene homólogo a LOG, que codifica uma proteína participante da etapa final de ativação de citocininas (Kurakawa et al., 2007), além do gene *acdS* (Grichko & Glick, 2000), traduzido na enzima ACC desaminase.

A cepa UENF412522 apresentou oito operons relacionados a FBN: *nifABSTZZ*, *nifDHNEX*, *nifPSUVW*, *nifQ*, *nifS* (duas regiões), *nifJ* e *sufBCDS* (Figura 9a). Os genes responsáveis pela tradução das unidades catalíticas da nitrogenase (*nifDHN*) estão na região *nifDHNEX* e junto a este operon há reguladores típicos da nitrogenase: *draG* (promotor) e *draT* (repressor). O gene *nifJ* está associado a transferência de elétrons para o complexo da nitrogenase. Os demais genes estão envolvidos em disponibilização de enxofre, síntese de grupamentos Fe-S e do cofator FeMo (*sufBCDS*, *nifS*, *nifPSUVW*, *nifQ*, *nifABSTZZ*). Já na cepa UENF114111, os genes da FBN estavam todos concentrados em quatro operons na mesma região: *modABCE*, *nifKDHQXEN*+ferrodoxina, *nifABZZST*+ferrodoxina e *nifVW+iscAS*+ferrodoxina (Figura 9b). O operon *mod* está envolvido com a absorção de molibdato (Perinet et al., 2016), sendo fonte de molibdênio para síntese do cofator FeMo. O gene *nifW*, possivelmente, age como um protetor da nitrogenase contra oxigênio (Kim & Burgess, 1996). *iscAS* estão envolvidos na síntese de

grupamentos Fe-S (Ding & Clark, 2004; Urbina et al., 2001) e as ferredoxinas, possivelmente, cumprem o papel de *nifJ* na transferência de elétrons. No meio deste cluster há um operon contendo genes de função desconhecida. Este *cluster* de operons da FBN, por estar centralizado, pode ser utilizado para transformar bactérias não-diazotróficas, uma vez que possui todos os genes essenciais deste processo (Vicente & Dean, 2017).

Na cepa UENF114111, o gene *gabY*, possivelmente, está envolvido na solubilização de fosfato, uma vez que o *knock out* deste gene em *E. coli* cessou a solubilização de fosfato em consequência da não produção de ácido glucônico (Babu-Khan et al., 1995). Para a cepa UENF412522, o operon *pqq* (Felder et al., 2000), provavelmente, está envolvido solubilização de fosfato.

A cepa UENF412522 possui o gene *iaaH*, responsável pela etapa final da síntese da auxina ácido índol-3-acético (Spaepen & Vanderleyden, 2011). A presença apenas do gene final da via do pode atuar no *pool* de IAM na planta, deslocando o equilíbrio para o IAA. A citocinina LOG também pode agir por um mecanismo similar a este, ativando o *pool* de citocininas inativas vegetais. Não foram encontrados genes relacionados à síntese de IAA na cepa UENF114111.

Figura 8 – Filogenia de máxima verossimilhança de SNPs extraídos do alinhamento dos core genes para os gêneros *Paraburkholderia* e *Ralstonia*. *P. necessarius* QLV-P1DMWA-1 foi definido como grupo externo. A cepa *P. rhizoxinica* HKI454 mostrou-se um grupo basal dentro de *Paraburkholderia*, sendo a espécie mais distinta deste clado.

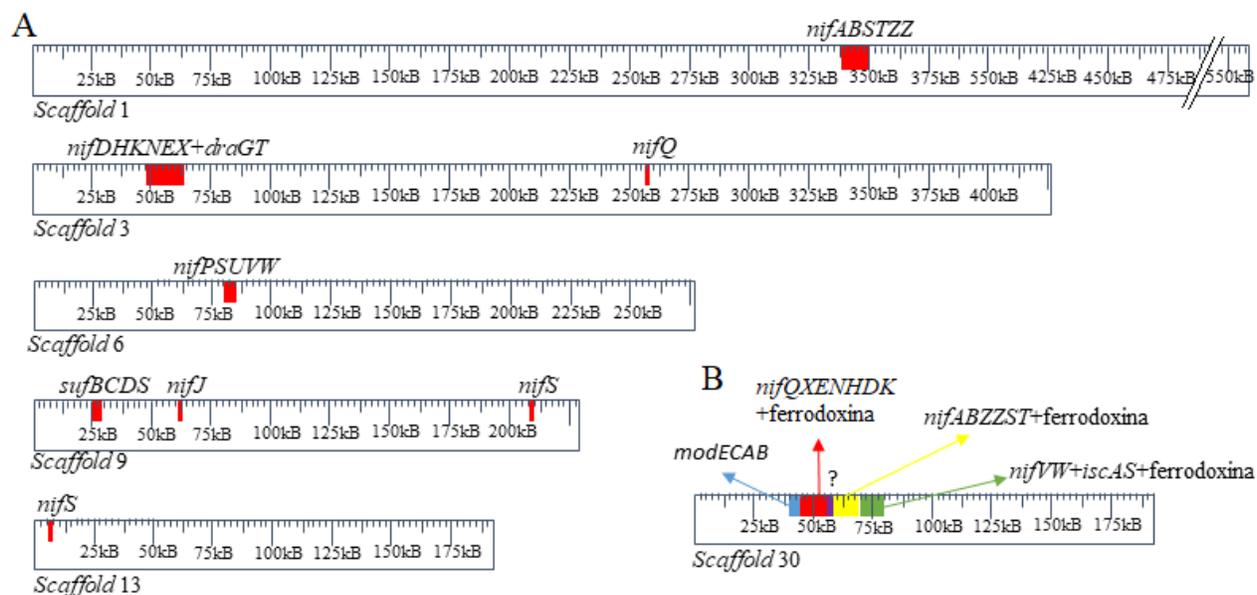


Figura 9 – Distribuição dos genes *nif* posicionados em representações dos *scaffolds* das cepas UENF41522 (A) e UENF114111 (B). Na cepa UENF41522, os genes da FBN estão em oito operons dispersos em 5 *scaffolds*. Na cepa UENF114111, quatro operons contendo os genes da FBN estão adjacentes no mesmo *scaffold*.

Em relação aos genes envolvidos em outros processos ecofisiológicos foram encontrados diversos genes de resistência a antibióticos, a cepa UENF114111 apresentou os genes de resistência a fosmidomicina (*fsr*) (Fujisaki et al., 1996), metilenomicina A (*mmr*) (Neal & Chater, 1987), tabtoxina (*ttr*) (Anzai et al., 1990), biciclomicina (*bcr*) (Bentley et al., 1993), polimixina (*ArnA/pmr*) (Gatzeva-Topalova et al., 2004; Nummala et al., 1995), ácido fusárico (*fusAC*) (Utsumi et al., 1991), bacitracina (*uppP/BacA*) (El Ghachi et al., 2004) e a toxina ratA (*pasT*) (Norton & Mulvey, 2012), além de genes envolvidos na resistência a múltiplas drogas como *MdtABCEG* (Nagakubo et al., 2002; Nishino & Yamaguchi, 2002). A cepa 412522 apresentou genes de resistência a fosmidomicina (*fsr*), biciclomicina (*bcr*), bacitracina (*uppP/BacA*), tabtoxina (*ttr*), fenazina (*ehpR*) (Giddens et al., 2002), tetraciclina (*tetA*) (Waters et al., 1983), resistência a múltiplas drogas *MdtABCE*, *AcrA* (Ma et al., 1993) e *NorM* (Brown et al., 1999). A maior parte dos antibióticos citados foram isolados originalmente do gênero *Streptomyces*, extremamente representativo na composição do solo (total de 576 espécies descritas) (Labeda, 2011). A

resistência às toxinas produzidas por este gênero é desejável para um melhor estabelecimento de inoculantes no solo.

A síntese de trealose bacteriana pode conferir diversas propriedades osmoprotetoras aos vegetais (Duan et al., 2013). Na cepa UENF412522 foram encontradas três vias de síntese para trealose: *treS*, *treYZ* e *otsAB*. *otsAB* é conhecido na forma de um operon, porém nesta cepa estão distantes no genoma: *ostA* está associado a uma glicoquinase e *ostB* a uma frutoquinase (Figura 12), possivelmente, sendo co-expressos para converter UDP-glicose em trealose (Kaasen et al., 1994). O operon *treYZ* apresenta um gene *treZ* fundido com um domínio de maltoquinase (Figura 10), forma não reportada na literatura e compartilhada por outras cepas de Rhodospirillaceae (Figura 11). Na cepa UENF114111, os genes *treS* e *treYZ* estão a poucas ORFs de distância e no mesmo lado da fita de DNA, possivelmente pertencendo ao mesmo operon.

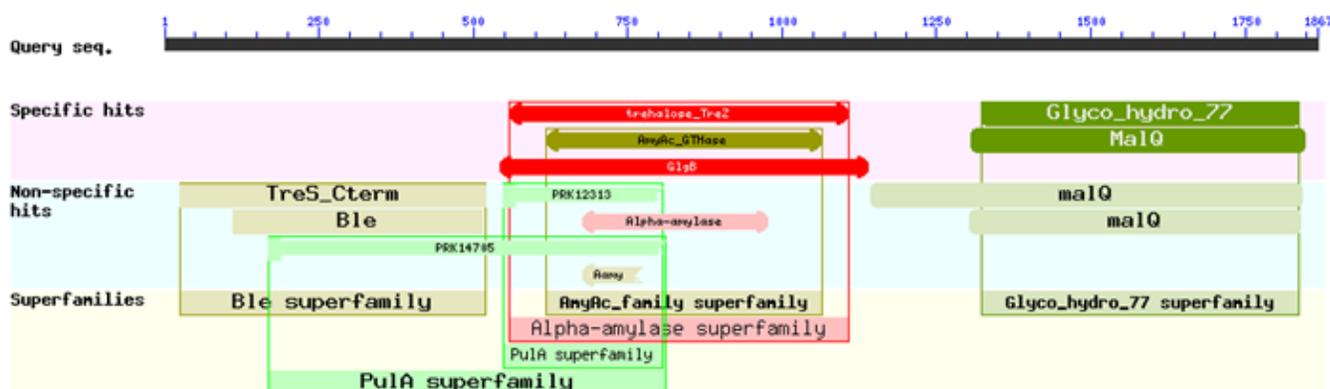


Figura 10 – Análise dos domínios de *treZ* da cepa UENF412522 pelo CDD-NCBI. Há um domínio de maltoquinase (*malQ*) anexado ao restante da estrutura de *treZ*.

A cepa UENF412522 possivelmente fixa carbono através do ciclo de Calvin-Benson-Bassham (CBB). Recentemente os constituintes desta via foram identificados experimentalmente no gênero *Azospirillum* (Orlova et al., 2016). Diversos genes CBB (Rasigraf et al., 2014) foram encontrados nesta cepa, como o gene da subunidade maior da enzima RubisCO (*cbbL*) e subunidade menor (*cbbS*) (Schwedock et al., 2004), o regulador transcricional do operon da RubisCO (*cbbR*) (Gibson & Tabita, 1993), RubisCO ativase (*cbbX*) (Mueller-Cajar et al., 2011), fosfoglicerato quinase (*pgk*) (Alefounder & Perham, 1989), gliceraldeído-3-fosfato desidrogenase (*gapB*) (Brinkmann et al., 1989), frutose bisfosfato aldolase (*cbbA*) (Hayashi et al., 2000), ribulose fosfato 3 epimerase (*cbbE*) e transcetolase (*cbbT*) (van den Bergh et al., 1996), fosfoglicolato fosfatase (*cbbZ*) (Schaferjohann et al., 1993), frutose-1,6-bisfosfatase (*fpb*) (Hamilton et al., 1988)

e fosforibuloquinase (*cfxP*) (Kossmann et al., 1989). O CBB possibilita que esta cepa de *Azospirillum* fixe carbono do ar em condições de baixa disponibilidade de carboidratos, o que contribui para a sobrevivência desta cepa em condições de campo.

Outros importantes genes estão presentes em ambas as cepas para auxiliar no combate ao estresse oxidativo causado pela planta durante a colonização das bactérias (Sousa & Olivares, 2016), como o gene da catalase *cat* (Henkle-Duhrsen et al., 1998) e os genes da superóxido dismutase *sodB* (Laudenbach et al., 1989) e *sodC2* (Figueroa-Bossi & Bossi, 1999).

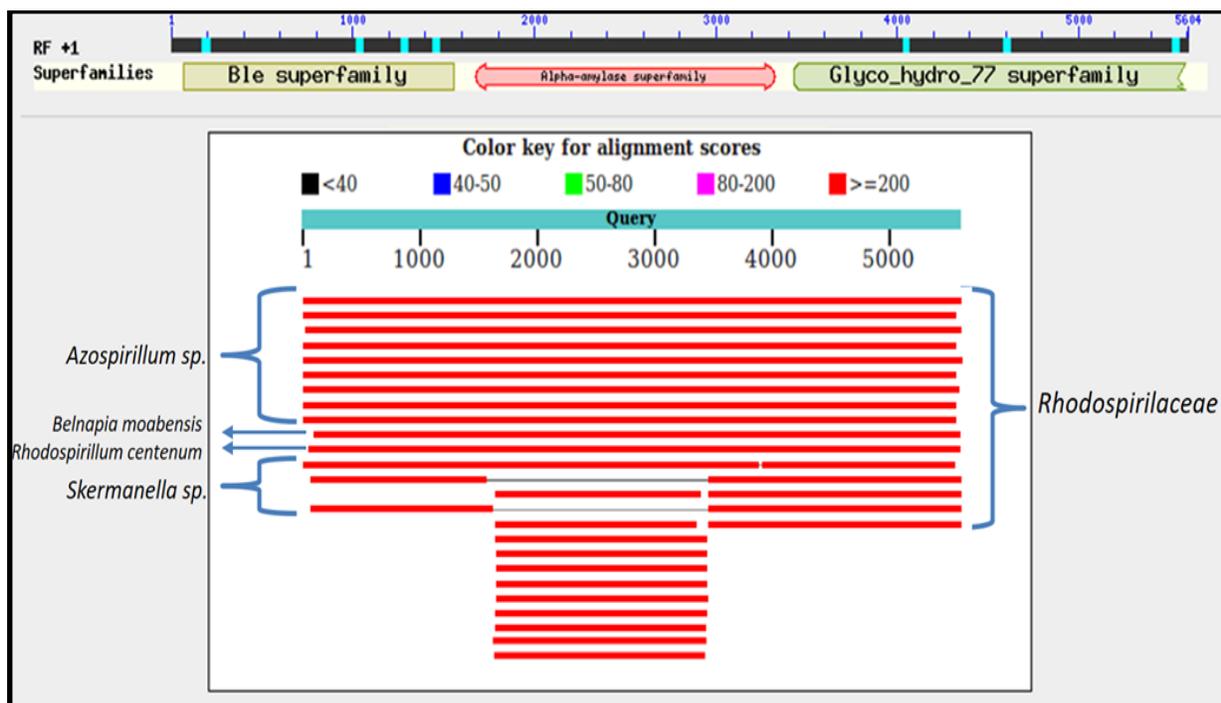


Figura 11 – blastP contra a base nr para treZ de UENF412522. Observa-se a mesma forma de treZ+malQ em nove espécies de *Azospirillum*, em *Belnapia moabensis* e *Rhodospirillum centenum*. Esta fusão aparenta ser específica de alguns membros da família Rhodospirilaceae, uma vez que as cepas de *Skermanella* não possuem esta estrutura.

Nas visualizações circulares dos genomas pelo CIRCOS (Figuras 11 e 12), foi possível representar 31 regiões genômicas de UENF412522 e 30 de UENF114111 divididas em três categorias: genes de promoção do crescimento vegetal, resistência a antibióticos e metabolismo de carbono. Os *scaffolds* inferiores a 80kb foram concatenados para melhor visualização. A distribuição dos genes foi relativamente homogênea, não sendo observadas posições preferenciais de genes.

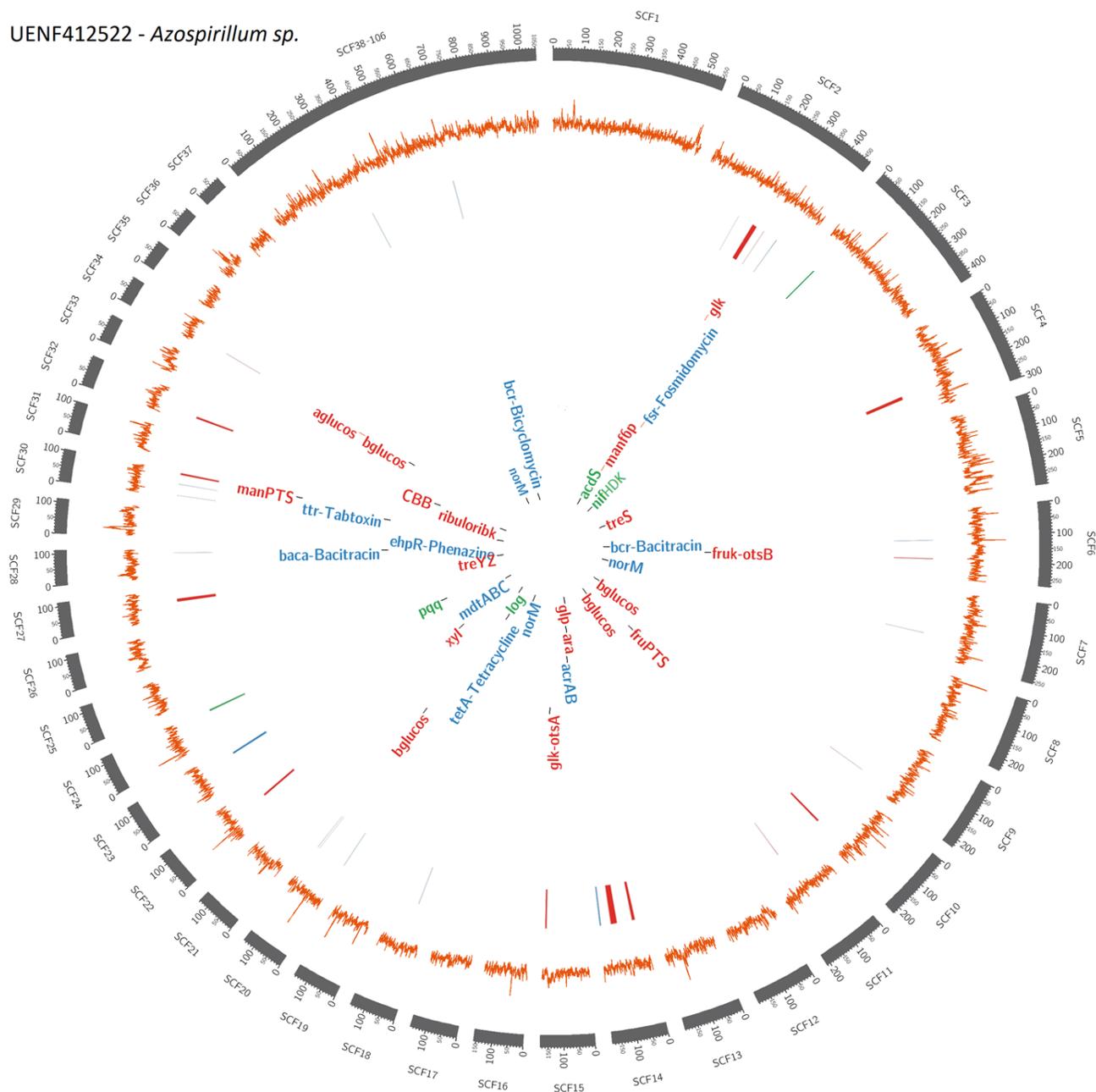
UENF412522 - *Azospirillum* sp.

Figura 12 – Visualização circular do genoma da cepa UENF412522. Nos anéis mais externos (cinza) estão plotados os scaffolds, seguido por porcentagem de GC (laranja), posição dos genes e suas respectivas identificações. Em azul os genes de resistência à antibióticos: fosmidomicina (*fsm*), biciclomicina (*bcr*), bacitracina (*uppP/BacA*), tabtoquina (*ttr*), fenazina (*ehpR*), tetraciclina (*tetA*), resistência a multiplas drogas *MdtABCE*, *AcrAB* e *NorM*. Em vermelho os genes de metabolismo de carbono: importação-fosforilação de manose (*manPTS*) e frutose (*fruPTS*), α -glucosidase (*aglucos*), β -glucosidase (*bglucos*), ciclo de Calvin-Benson-Brassham (*CBB*), ribulose quinase + riboquinase (*ribuloribk*), trehalose (*treYZ*, *treS*, *otsB*+frutoquinase, *otsA*+glicoquinase), operon da xilose (*xyl*), operon do glicerol (*glp*), operon da arabinose (*ara*), glicoquinase (*glk*), frutoquinase-6 + manitol-2-desidrogenase (*manf6p*). Em verde, os genes diretamente relacionados a promoção do crescimento vegetal: síntese de ácido glucônico (*pqq*), citocinina (*log*), ACC desaminase (*acds*) e nitrogenase (*nifHDK*).

5.5 Genes de metabolismo de carbono

Diversos genes envolvidos no metabolismo de carbono foram identificados, principalmente com auxílio dos mapas metabólicos do Kegg. Na cepa UENF412522, foram encontrados sistemas de importação-fosforilação de manose (manPTS) e frutose (fruPTS); para quebrar polissacarídeos, esta cepa possui uma α -glucosidase e uma β -glucosidase, adjacentes no genoma (Figura 12), embora em fitas diferentes. Os genes da ribulose quinase e riboquinase estão no mesmo operon. Há ainda a presença do operon da xilose, contendo importadores (*xylHF*) (Erbeznik et al., 2004), xilose isomerase (*xylA*) e xilulose quinase (*xylB*) (Takeda et al., 1998); operon do glicerol, contendo os genes de importação *glpPQSTV* (Eiglmeier et al., 1987) em uma região diferente dos genes glicerol quinase (*glpK*) (Pettigrew et al., 1988) e glicerol-3-fosfato desidrogenase (*glpD*) (Austin & Larson, 1991); os genes de metabolismo da arabinose estão em localizados em dois operons adjacentes: *araAC* e *araBDE*, sendo que cada um dos operons possui importadores ABC para açúcar próprios, não-homólogos ao já descrito *araXYZ* (Watanabe et al., 2006); há um gene com função glicoquinase, sorbitol desidrogenase (conversão sorbose em sorbitol (*polS*)) (Schauder et al., 1995), e além dos sistemas PTS para manose e frutose, há um operon que contém frutoquinase-6 e manitol-2-desidrogenase (*mtlK*) (Schneider et al., 1993). O teste de fontes de carbono API 50 CH/E detectou atividade para glicogênese nesta cepa (tabela 4), sendo o gene codificador da frutose-1,6-bisfosfatase (*fpb*) um dos envolvidos neste processo.

Na cepa UENF114111 há um operon da xilose *xylAB* e outro contendo os transportadores *xylFGH*; operon do ribitol, contendo ribitol-2-desidrogenase (*rbtD*) e uma ribuloquinase; para quebrar polissacarídeos há uma β -glucosidase (curiosamente, nenhuma α -glucosidase foi encontrada); para frutose, manitol-2-desidrogenase (*dald*, conversão de manitol e frutose) (Heuel et al., 1998) e frutoquinase (*mak*) (Sproul et al., 2001). Após fosforilada pode seguir por outras vias como manose-6-fosfato isomerase e glutamina-frutose-6-fosfato aminotransferase (três cópias em posições distintas, Figura 13); operon da ramnose+glicerol, contendo genes como glicerol quinase (*glpK*), glicerol-3-fosfato desidrogenase (*glpD*), L-ramnose-1-desidrogenase *dhg2* (Koivistoinen et al., 2008). Os transportadores *rhaSPQT* estão em um operon adjacente; glicerol quinase (*glp*); todos os genes referentes à arabinose estão presentes nos operons *araFGHB*, *araHGFADCE* e o gene *araB* isolado. Os genes *araEFGH* são transportadores, *araA* arabinose isomerase, *araB* ribuloquinase, *araC* regulador e *araD* L-ribulose-fosfato espimerase (Schleif, 2000); os genes de metabolismo da ribose foram *rbsBCA*, *rbsCK*, *rbsRK*, sendo *rbsABC*

transportadores, *rbsK* riboquinase e *rbsR* o repressor; uma glicoquinase e sorbitol desidrogenase isoladas.

| Composto | 114111 | 412522 | Genes |
|-----------------------------------|---------------|---------------|--|
| Glicerol | + | + | <i>glp</i> |
| Eritritol | + | - | ? |
| D-adonitol | + | - | <i>rbt</i> |
| Dulcitol | + | - | ? |
| D-manitol | + | + | <i>mtlK, dalD</i> |
| D-sorbitol | + | - | <i>polS</i> |
| L-arabinose | + | + | <i>ara</i> |
| D-ribose | + | + | <i>rbs</i> |
| D-xilose | + | + | <i>xyl</i> |
| Metil- βD- Xilopiranosido | + | + | ? |
| D-galactose | + | + | <i>araA</i> |
| D-glicose | + | + | <i>glk</i> |
| D-frutose | + | + | <i>fru-PTS, mak, mtlK, dalD</i> |
| D-manose | +? | + | <i>manPTS</i> |
| L-sorbose | + | + | <i>polS</i> |
| L-ramnose | + | - | <i>dhg2</i> |
| Metil- αD- Glucopiranosido | + | - | ? |
| D-trealose | + | + | <i>treS, treY, otsAB, glucosidases</i> |
| D-maltose | + | + | <i>treS, glucosidases</i> |
| D-arabinose | - | + | ? |
| Glicogenese | - | + | <i>fbp</i> |
| Gentiobiose | - | + | <i>glucosidases</i> |
| D-Tagatose | - | + | |
| Esculina citrato de ferro | + | + | |
| Salicina | - | + | |
| Amigdalina | - | + | |
| D-celobiose | - | + | |
| D-lactose (origem bovina) | + | + | |
| D-melibiose | + | + | |
| D-sacarose | + | + | |
| Amido | + | - | |
| D-Rafinose | + | - | |
| D-Turanose | + | - | |

Tabela 4 – Teste para fontes de carbono API 50 CH/E com seus respectivos possíveis genes de metabolismo. O sinal de + indica habilidade de metabolizar a fonte, e - inabilidade de metabolizar, e ? gene não encontrado.

O uso do teste para fontes de carbono API 50 CH/E é uma ferramenta utilizada na identificação de microrganismos de acordo com as fontes de carbono que estes podem metabolizar. Neste trabalho, os genes relacionados as fontes de carbono testadas foram sistematizados (tabela 4). Para alguns compostos (eritritol, dulcitol, metil- β D-xilopiranosido, metil- α D-glucopiranosido) não foi possível encontrar genes diretamente relacionados com seu metabolismo. É possível que as enzimas mencionadas anteriormente atuem em outras fontes de carbono além das relatadas na literatura, devido considerável similaridade na estrutura química de alguns carboidratos.

7. CONCLUSÃO

As montagens de genomas apresentaram melhores métricas de contiguidade para o montador SPADes 3.0, o que indica que este montador foi o mais adequado para este conjunto de dados. As cepas UENF412522 e UENF114111 foram caracterizadas como *Azospirillum sp.* e *Paraburkholderia sp.* e foi possível predizer 6892 e 5960 genes respectivamente. A anotação automática pelo Prokka inferiu a função de aproximadamente 76,5% dos genes, enquanto que a anotação feita por blast contra o SwissProt inferiu a função de 55,5% dos genes para UENF412522 e 60% para UENF114111.

As cepas *Azospirillum sp.* UENF412522 e *Paraburkholderia sp.* UENF114111 apresentaram diversos genes interessantes no contexto de sua possível aplicação na formulação de inoculantes agrícolas, como resistência a antibióticos da rizosfera, produção de fitormônios vegetais, solubilização de fosfato e FBN. Além disto, este trabalho serve como base para análises de expressão gênica (e.g. para a cepa UENF412522 verificar a presença transcritos contendo a fusão de *treY* e *malQ* e se há expressão do CBB na ausência de fontes de carbono), além da verificação da resistência aos antibióticos citados.

É importante que comunidade científica se atente aos estudos genômicos de bactérias do solo. Para efeito de comparação, *Paraburkholderia* (ambiental) possui 69 genomas depositados e *Burkholderia* (majoritariamente isolados clínicos), em torno de 1900. O gênero *Azospirillum* também possui uma baixa representatividade em termo de genomas sequenciados (24 genomas). Isto reflete o baixo número de estudos genômicos de bactérias do solo em contraste com a área clínica. A dificuldade no cultivo destes microrganismos contribui ainda mais para este hiato. Em particular, espera-se que as novas técnicas de sequenciamento possam auxiliar na melhor

caracterização de rizobactérias, como exemplificado no presente trabalho. É possível que bactérias com novas propriedades de promoção do crescimento vegetal estejam deixando de ser investigadas, o que indiretamente incentiva a manutenção dos sistemas agrícolas dependentes de insumos químicos.

8. REFERÊNCIAS

- Adnan, M., Shah, Z., Fahad, S., Arif, M., Alam, M., Khan, I. A., Mian, I. A., Basir, A., Ullah, H., Arshad, M., Rehman, I. U., Saud, S., Ihsan, M. Z., Jamal, Y., Amanullah, Hammad, H. M., & Nasim, W. (2017). Phosphate-Solubilizing Bacteria Nullify the Antagonistic Effect of Soil Calcification on Bioavailability of Phosphorus in Alkaline Soils. *Sci Rep*, 7(1), 16131.
- Alefunder, P. R., & Perham, R. N. (1989). Identification, molecular cloning and sequence analysis of a gene cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-phosphoglycerate kinase and a putative second glyceraldehyde 3-phosphate dehydrogenase of *Escherichia coli*. *Mol Microbiol*, 3(6), 723-732.
- Alori, E. T., Glick, B. R., & Babalola, O. O. (2017). Microbial Phosphorus Solubilization and Its Potential for Use in Sustainable Agriculture. *Front Microbiol*, 8, 971.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25(17), 3389-3402.
- Anzai, H., Yoneyama, K., & Yamaguchi, I. (1990). The nucleotide sequence of tabtoxin resistance gene (*ttr*) of *Pseudomonas syringae* pv. *tabaci*. *Nucleic Acids Res*, 18(7), 1890.
- Arshad, M., & Frankenberger, W. T. (1991). Microbial production of plant hormones. *Plant and Soil*, 133(1), 1-8.
- Asea, P. E. A., Kucey, R. M. N., & Stewart, J. W. B. (1988). Inorganic phosphate solubilization by two *Penicillium* species in solution culture and soil. *Soil Biology and Biochemistry*, 20(4), 459-464.
- Austin, D., & Larson, T. J. (1991). Nucleotide sequence of the *glpD* gene encoding aerobic sn-glycerol 3-phosphate dehydrogenase of *Escherichia coli* K-12. *J Bacteriol*, 173(1), 101-107.
- Aziz, M., Nadipalli, R. K., Xie, X., Sun, Y., Surowiec, K., Zhang, J. L., & Pare, P. W. (2016). Augmenting Sulfur Metabolism and Herbivore Defense in *Arabidopsis* by Bacterial Volatile Signaling. *Front Plant Sci*, 7, 458.
- Babu-Khan, S., Yeo, T. C., Martin, W. L., Duron, M. R., Rogers, R. D., & Goldstein, A. H. (1995). Cloning of a mineral phosphate-solubilizing gene from *Pseudomonas cepacia*. *Appl Environ Microbiol*, 61(3), 972-978.
- Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S., & Vivanco, J. M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol*, 57, 233-266.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*, 19(5), 455-477.
- Bashan, Y. (1998). Inoculants of plant growth-promoting bacteria for use in agriculture. *Biotechnology Advances*, 16(4), 729-770.
- Bentley, J., Hyatt, L. S., Ainley, K., Parish, J. H., Herbert, R. B., & White, G. R. (1993). Cloning and sequence analysis of an *Escherichia coli* gene conferring bicyclomycin resistance. *Gene*, 127(1), 117-120.

- Biondi, E. G., Toro, N., Bazzicalupo, M., & Martínez-Abarca, F. (2011). Spread of the group II intron RmlInt1 and its insertion sequence target sites in the plant endosymbiont *Sinorhizobium meliloti*. *Mobile Genetic Elements*, *1*(1), 2-7.
- Bodirsky, B. L., Popp, A., Lotze-Campen, H., Dietrich, J. P., Rolinski, S., Weindl, I., Schmitz, C., Müller, C., Bonsch, M., Humpenoder, F., Biewald, A., & Stevanovic, M. (2014). Reactive nitrogen requirements to feed the world in 2050 and potential to mitigate nitrogen pollution. *Nat Commun*, *5*, 3858.
- Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., & Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, *27*(4), 578-579.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120.
- Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., Bansal, P., Bridge, A. J., Poux, S., Bougueleret, L., & Xenarios, I. (2016). UniProtKB/Swiss-Prot, the Manually Annotated Section of the UniProt KnowledgeBase: How to Use the Entry View. *Methods Mol Biol*, *1374*, 23-54.
- Brinkmann, H., Cerff, R., Salomon, M., & Soll, J. (1989). Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. *Plant Mol Biol*, *13*(1), 81-94.
- Brown, M. H., Paulsen, I. T., & Skurray, R. A. (1999). The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol Microbiol*, *31*(1), 394-395.
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat Methods*, *12*(1), 59-60.
- Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, *8*(3), 559-568.
- Carvalho, T. L., Balsemao-Pires, E., Saraiva, R. M., Ferreira, P. C., & Hemerly, A. S. (2014). Nitrogen signalling in plant interactions with associative and endophytic diazotrophic bacteria. *J Exp Bot*, *65*(19), 5631-5642.
- Chang, W.-S., Lee, H.-I., & Hungria, M. (2015). Soybean Production in the Americas. In B. Lugtenberg (Ed.), *Principles of Plant-Microbe Interactions: Microbes for Sustainable Agriculture* (pp. 393-400). Cham: Springer International Publishing.
- Chauhan, A., Guleria, S., Balgir, P. P., Walia, A., Mahajan, R., Mehta, P., & Shirkot, C. K. (2017). Tricalcium phosphate solubilization and nitrogen fixation by newly isolated *Aneurinibacillus aneurinilyticus* CKMV1 from rhizosphere of *Valeriana jatamansi* and its growth promotional effect. *Braz J Microbiol*, *48*(2), 294-304.
- Chen, N., Goodwin, P. H., & Hsiang, T. (2003). The role of ethylene during the infection of *Nicotiana tabacum* by *Colletotrichum destructivum*. *J Exp Bot*, *54*(392), 2449-2456.
- Chikhi, R., & Rizk, G. (2013). Space-efficient and exact de Bruijn graph representation based on a Bloom filter. *Algorithms Mol Biol*, *8*(1), 22.
- Chung, J. H., Song, G. C., & Ryu, C. M. (2016). Sweet scents from good bacteria: Case studies on bacterial volatile compounds for plant growth and immunity. *Plant Mol Biol*, *90*(6), 677-687.
- Danyal, K., Shaw, S., Page, T. R., Duval, S., Horitani, M., Marts, A. R., Lukoyanov, D., Dean, D. R., Raugei, S., Hoffman, B. M., Seefeldt, L. C., & Antony, E. (2016). Negative cooperativity in the nitrogenase Fe protein electron delivery cycle. *Proc Natl Acad Sci U S A*, *113*(40), E5783-E5791.
- de Almeida, C. V., Andreote, F. D., Yara, R., Tanaka, F. A. O., Azevedo, J. L., & de Almeida, M. (2009). Bacteriosomes in axenic plants: endophytes as stable endosymbionts. *World Journal of Microbiology and Biotechnology*, *25*(10), 1757-1764.
- Ding, H., & Clark, R. J. (2004). Characterization of iron binding in IscA, an ancient iron-sulphur cluster assembly protein. *Biochem J*, *379*(Pt 2), 433-440.

- Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol*, 2(8), 621-631.
- Dobbelaere, S., Croonenborghs, A., Thys, A., Vande Broek, A., & Vanderleyden, J. (1999). Phytostimulatory effect of *Azospirillum brasilense* wild type and mutant strains altered in IAA production on wheat. *Plant and Soil*, 212(2), 153-162.
- Duan, J., Jiang, W., Cheng, Z., Heikkila, J. J., & Glick, B. R. (2013). The complete genome sequence of the plant growth-promoting bacterium *Pseudomonas* sp. UW4. *PLoS One*, 8(3), e58640.
- Duijff, B. J., Gianinazzi-Pearson, V., & Lemanceau, P. (1997). Involvement of the outer membrane lipopolysaccharides in the endophytic colonization of tomato roots by biocontrol *Pseudomonas fluorescens* strain WCS417r. *New Phytologist*, 135(2), 325-334.
- Eberl, L., & Vandamme, P. (2016). *Members of the genus Burkholderia: good and bad guys [version 1; referees: 3 approved]* (Vol. 5).
- Eiglmeier, K., Boos, W., & Cole, S. T. (1987). Nucleotide sequence and transcriptional startpoint of the *glpT* gene of *Escherichia coli*: extensive sequence homology of the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system. *Mol Microbiol*, 1(3), 251-258.
- Einsle, O., Tezcan, F. A., Andrade, S. L., Schmid, B., Yoshida, M., Howard, J. B., & Rees, D. C. (2002). Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. *Science*, 297(5587), 1696-1700.
- El Ghachi, M., Bouhss, A., Blanot, D., & Mengin-Lecreulx, D. (2004). The *bacA* gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. *J Biol Chem*, 279(29), 30106-30113.
- Elmerich, C. (2015). One Hundred Years Discovery of Nitrogen-Fixing Rhizobacteria *Biological Nitrogen Fixation* (pp. 897-912): John Wiley & Sons, Inc.
- EMBRAPA. Fixação Biológica de Nitrogênio Retrieved 15/02, 2017, from <https://www.embrapa.br/tema-fixacao-biologica-de-nitrogenio/nota-tecnica>
- Erbezniak, M., Hudson, S. E., Herrman, A. B., & Strobel, H. J. (2004). Molecular analysis of the *xylFGH* operon, coding for xylose ABC transport, in *Thermoanaerobacter ethanolicus*. *Curr Microbiol*, 48(4), 295-299.
- Evenson, R. E., & Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300(5620), 758-762.
- Felder, M., Gupta, A., Verma, V., Kumar, A., Qazi, G. N., & Cullum, J. (2000). The pyrroloquinoline quinone synthesis genes of *Gluconobacter oxydans*. *FEMS Microbiol Lett*, 193(2), 231-236.
- Figueroa-Bossi, N., & Bossi, L. (1999). Inducible prophages contribute to *Salmonella* virulence in mice. *Mol Microbiol*, 33(1), 167-176.
- Fujisaki, S., Ohnuma, S., Horiuchi, T., Takahashi, I., Tsukui, S., Nishimura, Y., Nishino, T., Kitabatake, M., & Inokuchi, H. (1996). Cloning of a gene from *Escherichia coli* that confers resistance to fosmidomycin as a consequence of amplification. *Gene*, 175(1-2), 83-87.
- Garnerone, A. M., Cabanes, D., Foussard, M., Boistard, P., & Batut, J. (1999). Inhibition of the FixL sensor kinase by the FixT protein in *Sinorhizobium meliloti*. *J Biol Chem*, 274(45), 32500-32506.
- Gatzeva-Topalova, P. Z., May, A. P., & Sousa, M. C. (2004). Crystal structure of *Escherichia coli* ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry*, 43(42), 13370-13379.
- Gibson, J. L., & Tabita, F. R. (1993). Nucleotide sequence and functional analysis of *cbbR*, a positive regulator of the Calvin cycle operons of *Rhodobacter sphaeroides*. *J Bacteriol*, 175(18), 5778-5784.
- Giddens, S. R., Feng, Y., & Mahanty, H. K. (2002). Characterization of a novel phenazine antibiotic gene cluster in *Erwinia herbicola* Eh1087. *Mol Microbiol*, 45(3), 769-783.

- Glazer, I., Orion, D., & Apelbaum, A. (1985). Ethylene Production by Meloidogyne spp.-Infected Plants. *J Nematol*, 17(1), 61-63.
- Glick, B. R. (2014). Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiol Res*, 169(1), 30-39.
- Glick, B. R. (2012). Plant Growth-Promoting Bacteria: Mechanisms and Applications. *Scientifica*, 2012, 15.
- Glick, B. R., Penrose, D. M., & Li, J. (1998). A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *J Theor Biol*, 190(1), 63-68.
- Grichko, V. P., & Glick, B. R. (2000). Identification of DNA sequences that regulate the expression of the Enterobacter cloacae UW4 1-aminocyclopropane-1-carboxylic acid deaminase gene. *Can J Microbiol*, 46(12), 1159-1165.
- Gupta, A., Gopal, M., Thomas, G. V., Manikandan, V., Gajewski, J., Thomas, G., Seshagiri, S., Schuster, S. C., Rajesh, P., & Gupta, R. (2014). Whole genome sequencing and analysis of plant growth promoting bacteria isolated from the rhizosphere of plantation crops coconut, cocoa and arecanut. *PLoS One*, 9(8), e104259.
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075.
- Hamilton, W. D., Harrison, D. A., & Dyer, T. A. (1988). Sequence of the Escherichia coli fructose-1,6-bisphosphatase gene. *Nucleic Acids Res*, 16(17), 8707.
- Hayashi, N. R., Terazono, K., Kodama, T., & Igarashi, Y. (2000). Structure of ribulose 1,5-bisphosphate carboxylase/oxygenase gene cluster from a thermophilic hydrogen-oxidizing bacterium, Hydrogenophilus thermoluteolus, and phylogeny of the fructose 1,6-bisphosphate aldolase encoded by cbbA in the cluster. *Biosci Biotechnol Biochem*, 64(1), 61-71.
- Henkle-Duhrsen, K., Eckelt, V. H., Wildenburg, G., Blaxter, M., & Walter, R. D. (1998). Gene structure, activity and localization of a catalase from intracellular bacteria in Onchocerca volvulus. *Mol Biochem Parasitol*, 96(1-2), 69-81.
- Heuel, H., Shakeri-Garakani, A., Turgut, S., & Lengeler, J. W. (1998). Genes for D-arabinitol and ribitol catabolism from Klebsiella pneumoniae. *Microbiology*, 144 (Pt 6), 1631-1639.
- Honma, M., & Shimomura, T. (1978). Metabolism of 1-Aminocyclopropane-1-carboxylic Acid. *Agricultural and Biological Chemistry*, 42(10), 1825-1831.
- Howard, J. B., & Rees, D. C. (2006). How many metals does it take to fix N₂? A mechanistic overview of biological nitrogen fixation. *Proc Natl Acad Sci U S A*, 103(46), 17088-17093.
- Huson, D. H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H. J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol*, 12(6), e1004957.
- Hussain, A., & Hasnain, S. (2009). Cytokinin production by some bacteria: Its impact on cell division in cucumber cotyledons. *African Journal of Microbiology Research*, 3(11), 704-712.
- Illmer, P., & Schinner, F. (1995). Solubilization of inorganic calcium phosphates—Solubilization mechanisms. *Soil Biology and Biochemistry*, 27(3), 257-263.
- Imada, E. L., Rolla Dos Santos, A. A., Oliveira, A. L., Hungria, M., & Rodrigues, E. P. (2017). Indole-3-acetic acid production via the indole-3-pyruvate pathway by plant growth promoter Rhizobium tropici CIAT 899 is strongly inhibited by ammonium. *Res Microbiol*, 168(3), 283-292.
- Intorne, A. C., de Oliveira, M. V., Lima, M. L., da Silva, J. F., Olivares, F. L., & de Souza Filho, G. A. (2009). Identification and characterization of Gluconacetobacter diazotrophicus mutants defective in the solubilization of phosphorus and zinc. *Arch Microbiol*, 191(5), 477-483.
- Iyer, B., Rajput, M. S., & Rajkumar, S. (2017). Effect of succinate on phosphate solubilization in nitrogen fixing bacteria harbouring chick pea and their effect on plant growth. *Microbiol Res*, 202, 43-50.

- James, E. K., & Olivares, F. L. (1998). Infection and Colonization of Sugar Cane and Other Gramineous Plants by Endophytic Diazotrophs. *Critical Reviews in Plant Sciences*, 17(1), 77-119.
- Johnson, S. L., Bishop-Lilly, K. A., Ladner, J. T., Daligault, H. E., Davenport, K. W., Jaissle, J., Frey, K. G., Koroleva, G. I., Bruce, D. C., Coyne, S. R., Broomall, S. M., Li, P. E., Teshima, H., Gibbons, H. S., Palacios, G. F., Rosenzweig, C. N., Redden, C. L., Xu, Y., Minogue, T. D., & Chain, P. S. (2015). Complete genome sequences for 59 burkholderia isolates, both pathogenic and near neighbor. *Genome Announc*, 3(2).
- Jones, D. L. (1998). Organic acids in the rhizosphere – a critical review. *Plant and Soil*, 205(1), 25-44.
- Jorquera, M. A., Crowley, D. E., Marschner, P., Greiner, R., Fernandez, M. T., Romero, D., Menezes-Blackburn, D., & De La Luz Mora, M. (2011). Identification of beta-propeller phytase-encoding genes in culturable Paenibacillus and Bacillus spp. from the rhizosphere of pasture plants on volcanic soils. *FEMS Microbiol Ecol*, 75(1), 163-172.
- Kaasen, I., McDougall, J., & Strom, A. R. (1994). Analysis of the otsBA operon for osmoregulatory trehalose synthesis in Escherichia coli and homology of the OtsA and OtsB proteins to the yeast trehalose-6-phosphate synthase/phosphatase complex. *Gene*, 145(1), 9-15.
- Kaminski, P. A., Mandon, K., Arigoni, F., Desnoues, N., & Elmerich, C. (1991). Regulation of nitrogen fixation in Azorhizobium caulinodans: identification of a fixK-like gene, a positive regulator of nifA. *Mol Microbiol*, 5(8), 1983-1991.
- Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J Mol Biol*, 428(4), 726-731.
- Keunen, E., Schellingen, K., Vangronsveld, J., & Cuypers, A. (2016). Ethylene and Metal Stress: Small Molecule, Big Impact. *Front Plant Sci*, 7, 23.
- Khan, M. S., Zaidi, A., Ahemad, M., Oves, M., & Wani, P. A. (2010). Plant growth promotion by phosphate solubilizing fungi – current perspective. *Archives of Agronomy and Soil Science*, 56(1), 73-98.
- Khush, G. S. (2001). Green revolution: the way forward. *Nature Reviews Genetics*, 2, 815.
- Kim, S., & Burgess, B. K. (1996). Evidence for the direct interaction of the nifW gene product with the MoFe protein. *J Biol Chem*, 271(16), 9764-9770.
- Koivistoinen, O. M., Hilditch, S., Voutilainen, S. P., Boer, H., Penttila, M., & Richard, P. (2008). Identification in the yeast Pichia stipitis of the first L-rhamnose-1-dehydrogenase gene. *FEBS J*, 275(10), 2482-2488.
- Kossmann, J., Klintworth, R., & Bowien, B. (1989). Sequence analysis of the chromosomal and plasmid genes encoding phosphoribulokinase from Alcaligenes eutrophus. *Gene*, 85(1), 247-252.
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res*, 19(9), 1639-1645.
- Kuan, K. B., Othman, R., Abdul Rahim, K., & Shamsuddin, Z. H. (2016). Plant Growth-Promoting Rhizobacteria Inoculation to Enhance Vegetative Growth, Nitrogen Fixation and Nitrogen Remobilisation of Maize under Greenhouse Conditions. *PLoS One*, 11(3), e0152478.
- Kudoyarova, G. R., Melentiev, A. I., Martynenko, E. V., Timergalina, L. N., Arkhipova, T. N., Shendel, G. V., Kuz'mina, L. Y., Dodd, I. C., & Veselov, S. Y. (2014). Cytokinin producing bacteria stimulate amino acid deposition by wheat roots. *Plant Physiol Biochem*, 83, 285-291.
- Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., Sakakibara, H., & Kyojuka, J. (2007). Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature*, 445(7128), 652-655.
- Labeda, D. P. (2011). Multilocus sequence analysis of phytopathogenic species of the genus Streptomyces. *Int J Syst Evol Microbiol*, 61(Pt 10), 2525-2531.

- Lackner, G., Moebius, N., Partida-Martinez, L., & Hertweck, C. (2011). Complete genome sequence of *Burkholderia rhizoxinica*, an Endosymbiont of *Rhizopus microsporus*. *J Bacteriol*, *193*(3), 783-784.
- Laudenbach, D. E., Trick, C. G., & Straus, N. A. (1989). Cloning and characterization of an *Anacystis nidulans* R2 superoxide dismutase gene. *Mol Gen Genet*, *216*(2-3), 455-461.
- Lenarcic, T., Albert, I., Bohm, H., Hodnik, V., Pirc, K., Zavec, A. B., Podobnik, M., Pahovnik, D., Zagar, E., Pruitt, R., Greimel, P., Yamaji-Hasegawa, A., Kobayashi, T., Zienkiewicz, A., Gomann, J., Mortimer, J. C., Fang, L., Mamode-Cassim, A., Deleu, M., Lins, L., Oecking, C., Feussner, I., Mongrand, S., Anderluh, G., & Nurnberger, T. (2017). Eudicot plant-specific sphingolipids determine host selectivity of microbial NLP cytolysins. *Science*, *358*(6369), 1431-1434.
- Liu, Y., Cao, L., Tan, H., & Zhang, R. (2017). Surface display of ACC deaminase on endophytic Enterobacteriaceae strains to increase saline resistance of host rice sprouts by regulating plant ethylene synthesis. *Microb Cell Fact*, *16*(1), 214.
- Ma, D., Cook, D. N., Alberti, M., Pon, N. G., Nikaido, H., & Hearst, J. E. (1993). Molecular cloning and characterization of *acrA* and *acrE* genes of *Escherichia coli*. *J Bacteriol*, *175*(19), 6299-6313.
- Mabilde, L., De Neve, S., & Sleutel, S. (2017). Regional analysis of groundwater phosphate concentrations under acidic sandy soils: Edaphic factors and water table strongly mediate the soil P-groundwater P relation. *J Environ Manage*, *203*(Pt 1), 429-438.
- Maduraimuthu, D., & Prasad, P. V. V. (2010). *Ethylene production under high temperature stress causes premature leaf senescence in soybean* (Vol. 37).
- Mandal, S. M., Chakraborty, D., & Dey, S. (2010). Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signal Behav*, *5*(4), 359-368.
- Mansouri, K., & Piepersberg, W. (1991). Genetics of streptomycin production in *Streptomyces griseus*: nucleotide sequence of five genes, *strFGHIK*, including a phosphatase gene. *Mol Gen Genet*, *228*(3), 459-469.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., & Bryant, S. H. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*, *45*(D1), D200-D203.
- Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., & Dixon, R. (2004). The NifL-NifA System: a multidomain transcriptional regulatory complex that integrates environmental signals. *J Bacteriol*, *186*(3), 601-610.
- Mathesius, U. (2008). Goldacre paper: Auxin: at the root of nodule development? *Functional Plant Biology*, *35*(8), 651-668.
- McFadden, G. I. (1999). Endosymbiosis and evolution of the plant cell. *Curr Opin Plant Biol*, *2*(6), 513-519.
- McNear Jr., D. H. (2013). The Rhizosphere - Roots, Soil and Everything In Between. *Nature Education Knowledge*, *4*(3), 1. Retrieved from
- Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W., & Banfield, J. F. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol*, *12*(5), R44.
- Mueller-Cajar, O., Stotz, M., Wendler, P., Hartl, F. U., Bracher, A., & Hayer-Hartl, M. (2011). Structure and function of the AAA+ protein CbbX, a red-type Rubisco activase. *Nature*, *479*(7372), 194-199.
- Munk, A. C., Copeland, A., Lucas, S., Lapidus, A., Del Rio, T. G., Barry, K., Detter, J. C., Hammon, N., Israni, S., Pitluck, S., Brettin, T., Bruce, D., Han, C., Tapia, R., Gilna, P., Schmutz, J., Larimer, F., Land, M., Kyrpides, N. C., Mavromatis, K., Richardson, P., Rohde, M., Goker, M., Klenk, H. P., Zhang, Y.,

- Roberts, G. P., Reslewic, S., & Schwartz, D. C. (2011). Complete genome sequence of *Rhodospirillum rubrum* type strain (S1). *Stand Genomic Sci*, 4(3), 293-302.
- Nagakubo, S., Nishino, K., Hirata, T., & Yamaguchi, A. (2002). The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. *J Bacteriol*, 184(15), 4161-4167.
- Narayana, I., Lalonde, S., & Saini, H. S. (1991). Water-stress-induced ethylene production in wheat : a fact or artifact? *Plant Physiol*, 96(2), 406-410.
- Neal, R. J., & Chater, K. F. (1987). Nucleotide sequence analysis reveals similarities between proteins determining methylenomycin A resistance in *Streptomyces* and tetracycline resistance in eubacteria. *Gene*, 58(2-3), 229-241.
- Nishino, K., & Yamaguchi, A. (2002). EvgA of the two-component signal transduction system modulates production of the yhiUV multidrug transporter in *Escherichia coli*. *J Bacteriol*, 184(8), 2319-2323.
- Norton, J. P., & Mulvey, M. A. (2012). Toxin-antitoxin systems are important for niche-specific colonization and stress resistance of uropathogenic *Escherichia coli*. *PLoS Pathog*, 8(10), e1002954.
- Nummila, K., Kilpelainen, I., Zahringer, U., Vaara, M., & Helander, I. M. (1995). Lipopolysaccharides of polymyxin B-resistant mutants of *Escherichia coli* are extensively substituted by 2-aminoethyl pyrophosphate and contain aminoarabinose in lipid A. *Mol Microbiol*, 16(2), 271-278.
- Orlova, M. V., Tarlachkov, S. V., Dubinina, G. A., Belousova, E. V., Tutukina, M. N., & Grabovich, M. Y. (2016). Genomic insights into metabolic versatility of a lithotrophic sulfur-oxidizing diazotrophic Alphaproteobacterium *Azospirillum thiophilum*. *FEMS Microbiol Ecol*, 92(12).
- Oteino, N., Lally, R. D., Kiwanuka, S., Lloyd, A., Ryan, D., Germaine, K. J., & Dowling, D. N. (2015). Plant growth promotion induced by phosphate solubilizing endophytic *Pseudomonas* isolates. *Front Microbiol*, 6, 745.
- Ott, T., van Dongen, J. T., Gunther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P., & Udvardi, M. K. (2005). Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Curr Biol*, 15(6), 531-535.
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res*, 42(Database issue), D206-214.
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., Fookes, M., Falush, D., Keane, J. A., & Parkhill, J. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31(22), 3691-3693.
- Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., & Harris, S. R. (2016). SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom*, 2(4), e000056.
- Pedrosa, F. O., Monteiro, R. A., Wassem, R., Cruz, L. M., Ayub, R. A., Colauto, N. B., Fernandez, M. A., Fungaro, M. H., Grisard, E. C., Hungria, M., Madeira, H. M., Nodari, R. O., Osaku, C. A., Petzl-Erler, M. L., Terenzi, H., Vieira, L. G., Steffens, M. B., Weiss, V. A., Pereira, L. F., Almeida, M. I., Alves, L. R., Marin, A., Araujo, L. M., Balsanelli, E., Baura, V. A., Chubatsu, L. S., Faoro, H., Favetti, A., Friedermann, G., Glienke, C., Karp, S., Kava-Cordeiro, V., Raittz, R. T., Ramos, H. J., Ribeiro, E. M., Rigo, L. U., Rocha, S. N., Schwab, S., Silva, A. G., Souza, E. M., Tadra-Sfeir, M. Z., Torres, R. A., Dabul, A. N., Soares, M. A., Gasques, L. S., Gimenes, C. C., Valle, J. S., Ciferri, R. R., Correa, L. C., Murace, N. K., Pamphile, J. A., Patussi, E. V., Prioli, A. J., Prioli, S. M., Rocha, C. L., Arantes, O. M., Furlaneto, M. C., Godoy, L. P., Oliveira, C. E., Satori, D., Vilas-Boas, L. A., Watanabe, M. A., Dambros, B. P., Guerra, M. P., Mathioni, S. M., Santos, K. L., Steindel, M., Vernal, J., Barcellos, F. G., Campo, R. J., Chueire, L. M., Nicolas, M. F., Pereira-Ferrari, L., Silva, J. L., Gioppo, N. M.,

- Margarido, V. P., Menck-Soares, M. A., Pinto, F. G., Simao Rde, C., Takahashi, E. K., Yates, M. G., & Souza, E. M. (2011). Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. *PLoS Genet*, *7*(5), e1002064.
- Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2010). IDBA – A Practical Iterative de Bruijn Graph De Novo Assembler. In B. Berger (Ed.), *Research in Computational Molecular Biology: 14th Annual International Conference, RECOMB 2010, Lisbon, Portugal, April 25-28, 2010. Proceedings* (pp. 426-440). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Perinet, S., Jeukens, J., Kukavica-Ibrulj, I., Ouellet, M. M., Charette, S. J., & Levesque, R. C. (2016). Molybdate transporter ModABC is important for *Pseudomonas aeruginosa* chronic lung infection. *BMC Res Notes*, *9*, 23.
- Pettigrew, D. W., Ma, D. P., Conrad, C. A., & Johnson, J. R. (1988). *Escherichia coli* glycerol kinase. Cloning and sequencing of the glpK gene and the primary structure of the enzyme. *J Biol Chem*, *263*(1), 135-139.
- Pierik, R., Tholen, D., Poorter, H., Visser, E. J., & Voesenek, L. A. (2006). The Janus face of ethylene: growth inhibition and stimulation. *Trends Plant Sci*, *11*(4), 176-183.
- Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proc Natl Acad Sci U S A*, *109*(31), 12302-12308.
- Plett, J. M., McDonnell, L., & Regan, S. (2009). Plant encoded 1-aminocyclopropane-1-carboxylic acid deaminase activity implicated in different aspects of plant development. *Plant Signal Behav*, *4*(12), 1186-1189.
- Raddadi, N., Cherif, A., Ouzari, H., Marzorati, M., Brusetti, L., Boudabous, A., & Daffonchio, D. (2007). *Bacillus thuringiensis* beyond insect biocontrol: plant growth promotion and biosafety of polyvalent strains. *Annals of Microbiology*, *57*(4), 481-494.
- Radzki, W., Gutierrez Manero, F. J., Algar, E., Lucas Garcia, J. A., Garcia-Villaraco, A., & Ramos Solano, B. (2013). Bacterial siderophores efficiently provide iron to iron-starved tomato plants in hydroponics culture. *Antonie Van Leeuwenhoek*, *104*(3), 321-330.
- Rasigraf, O., Kool, D. M., Jetten, M. S., Sinninghe Damste, J. S., & Ettwig, K. F. (2014). Autotrophic carbon dioxide fixation via the Calvin-Benson-Bassham cycle by the denitrifying methanotroph "Candidatus *Methyloirabilis oxyfera*". *Appl Environ Microbiol*, *80*(8), 2451-2460.
- Ravishankara, A. R., Daniel, J. S., & Portmann, R. W. (2009). Nitrous oxide (N₂O): the dominant ozone-depleting substance emitted in the 21st century. *Science*, *326*(5949), 123-125.
- Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One*, *8*(6), e66428.
- Reddy, C. A., & Saravanan, R. S. (2013). Polymicrobial Multi-functional Approach for Enhancement of Crop Productivity. *Adv Appl Microbiol*, *82*, 53-113.
- Richardson, A. E., & Simpson, R. J. (2011). Soil microorganisms mediating phosphorus availability update on microbial phosphorus. *Plant Physiol*, *156*(3), 989-996.
- Richter, M., & Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A*, *106*(45), 19126-19131.
- Richter, M., Rossello-Mora, R., Oliver Glockner, F., & Peplies, J. (2016). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*, *32*(6), 929-931.
- Robison, M. M., Griffith, M., Pauls, K. P., & Glick, B. R. (2001). Dual role for ethylene in susceptibility of tomato to *Verticillium* wilt. *Journal of Phytopathology-Phytopathologische Zeitschrift*, *149*(7-8), 385-388.
- Rodrigues, P. S. L., Santos, S. T., Duarte, N. S., Intorne, A. C., Matteoli, F. P., Venâncio, T. M., Araújo, J. L. S., Baldani, J. I., Reis, J. F. B., & Olivares, F. L. (2016). *Molecular and physiological*

- characterization of plant growth-promoting bacteria associated to tropical fruits crops*. Tese de doutorado Tese de doutorado, UENF.
- Rodriguez, H., & Fraga, R. (1999). Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol Adv*, 17(4-5), 319-339.
- Sawana, A., Adeolu, M., & Gupta, R. S. (2014). Molecular signatures and phylogenomic analysis of the genus Burkholderia: proposal for division of this genus into the emended genus Burkholderia containing pathogenic organisms and a new genus Paraburkholderia gen. nov. harboring environmental species. *Front Genet*, 5, 429.
- Schaferjohann, J., Yoo, J. G., Kusian, B., & Bowien, B. (1993). The cbb operons of the facultative chemoautotroph *Alcaligenes eutrophus* encode phosphoglycolate phosphatase. *J Bacteriol*, 175(22), 7329-7340.
- Schauder, S., Schneider, K. H., & Giffhorn, F. (1995). Polyol metabolism of *Rhodobacter sphaeroides*: biochemical characterization of a short-chain sorbitol dehydrogenase. *Microbiology*, 141 (Pt 8), 1857-1863.
- Schleif, R. (2000). Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet*, 16(12), 559-565.
- Schneider, K. H., Giffhorn, F., & Kaplan, S. (1993). Cloning, nucleotide sequence and characterization of the mannitol dehydrogenase gene from *Rhodobacter sphaeroides*. *J Gen Microbiol*, 139(10), 2475-2484.
- Schulz, S., & Dickschat, J. S. (2007). Bacterial volatiles: the smell of small organisms. *Nat Prod Rep*, 24(4), 814-842.
- Schwartz, D., Berger, S., Heinzelmann, E., Muschko, K., Welzel, K., & Wohlleben, W. (2004). Biosynthetic gene cluster of the herbicide phosphinothricin tripeptide from *Streptomyces viridochromogenes* Tu494. *Appl Environ Microbiol*, 70(12), 7093-7102.
- Schwedock, J., Harmer, T. L., Scott, K. M., Hektor, H. J., Seitz, A. P., Fontana, M. C., Distel, D. L., & Cavanaugh, C. M. (2004). Characterization and expression of genes from the RubisCO gene cluster of the chemoautotrophic symbiont of *Solemya velum*: cbbLSQO. *Arch Microbiol*, 182(1), 18-29.
- Seefeldt, L. C., Dance, I. G., & Dean, D. R. (2004). Substrate interactions with nitrogenase: Fe versus Mo. *Biochemistry*, 43(6), 1401-1409.
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068-2069.
- Seo, H., Kim, S., Sagong, H. Y., Son, H. F., Jin, K. S., Kim, I. K., & Kim, K. J. (2016). Structural basis for cytokinin production by LOG from *Corynebacterium glutamicum*. *Sci Rep*, 6, 31390.
- Sharma, S. B., Sayyed, R. Z., Trivedi, M. H., & Gobi, T. A. (2013). Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. *Springerplus*, 2, 587.
- Shen, Y. Q., Bonnot, F., Imsand, E. M., RoseFigura, J. M., Sjolander, K., & Klinman, J. P. (2012). Distribution and properties of the genes encoding the biosynthesis of the bacterial cofactor, pyrroloquinoline quinone. *Biochemistry*, 51(11), 2265-2275.
- Simonetti, E., Roberts, I. N., Montecchia, M. S., Gutierrez-Boem, F. H., Gomez, F. M., & Ruiz, J. A. (2018). A novel *Burkholderia ambifaria* strain able to degrade the mycotoxin fusaric acid and to inhibit *Fusarium* spp. growth. *Microbiol Res*, 206, 50-59.
- Skoog, F., & Miller, C. O. (1957). Chemical regulation of growth and organ formation in plant tissues cultured in vitro. *Symp Soc Exp Biol*, 11, 118-130.
- Sousa, J. A. d. J., & Olivares, F. L. (2016). Plant growth promotion by streptomycetes: ecophysiology, mechanisms and applications. *Chemical and Biological Technologies in Agriculture*, 3(1), 24.
- Spaepen, S., & Vanderleyden, J. (2011). Auxin and plant-microbe interactions. *Cold Spring Harb Perspect Biol*, 3(4).

- Spaepen, S., Vanderleyden, J., & Remans, R. (2007). Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol Rev*, *31*(4), 425-448.
- Sproul, A. A., Lambourne, L. T., Jean-Jacques, D. J., & Kornberg, H. L. (2001). Genetic control of manno(fructo)kinase activity in *Escherichia coli*. *Proc Natl Acad Sci U S A*, *98*(26), 15257-15259.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.
- Taiz, L., & Zeiger, E. (2010). *Plant Physiology* (5th ed.): Sinauer Associates, Inc.
- Takeda, Y., Takase, K., Yamato, I., & Abe, K. (1998). Sequencing and characterization of the xyl operon of a gram-positive bacterium, *Tetragenococcus halophila*. *Appl Environ Microbiol*, *64*(7), 2513-2519.
- Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, *108*(50), 20260-20264.
- Tuckerman, J. R., Gonzalez, G., Dioum, E. M., & Gilles-Gonzalez, M. A. (2002). Ligand and oxidation-state specific regulation of the heme-based oxygen sensor FixL from *Sinorhizobium meliloti*. *Biochemistry*, *41*(19), 6170-6177.
- Turner, B. L., Paphazy, M. J., Haygarth, P. M., & McKelvie, I. D. (2002). Inositol phosphates in the environment. *Philos Trans R Soc Lond B Biol Sci*, *357*(1420), 449-469.
- Urbina, H. D., Silberg, J. J., Hoff, K. G., & Vickery, L. E. (2001). Transfer of sulfur from IscS to IscU during Fe/S cluster assembly. *J Biol Chem*, *276*(48), 44521-44526.
- Utsumi, R., Yagi, T., Katayama, S., Katsuragi, K., Tachibana, K., Toyoda, H., Ouchi, S., Obata, K., Shibano, Y., & Noda, M. (1991). Molecular cloning and characterization of the fusaric acid-resistance gene from *Pseudomonas cepacia*. *Agric Biol Chem*, *55*(7), 1913-1918.
- Van Bavel, J. (2013). The world population explosion: causes, backgrounds and -projections for the future. *Facts Views Vis Obgyn*, *5*(4), 281-291.
- van den Bergh, E. R., Baker, S. C., Riggers, R. J., Terpstra, P., Woudstra, E. C., Dijkhuizen, L., & Meijer, W. G. (1996). Primary structure and phylogeny of the Calvin cycle enzymes transketolase and fructosebiphosphate aldolase of *Xanthobacter flavus*. *J Bacteriol*, *178*(3), 888-893.
- van Loon, L. C., Geraats, B. P., & Linthorst, H. J. (2006). Ethylene as a modulator of disease resistance in plants. *Trends Plant Sci*, *11*(4), 184-191.
- Vespermann, A., Kai, M., & Piechulla, B. (2007). Rhizobacterial volatiles affect the growth of fungi and *Arabidopsis thaliana*. *Appl Environ Microbiol*, *73*(17), 5639-5641.
- Vicente, E. J., & Dean, D. R. (2017). Keeping the nitrogen-fixation dream alive. *Proc Natl Acad Sci U S A*, *114*(12), 3009-3011.
- Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H., & Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecological Applications*, *7*(3), 737-750.
- Wang, D., Yang, S., Tang, F., & Zhu, H. (2012). Symbiosis specificity in the legume: rhizobial mutualism. *Cell Microbiol*, *14*(3), 334-342.
- Watanabe, S., Kodaki, T., & Makino, K. (2006). Cloning, expression, and characterization of bacterial L-arabinose 1-dehydrogenase involved in an alternative pathway of L-arabinose metabolism. *J Biol Chem*, *281*(5), 2612-2623.
- Waters, S. H., Rogowsky, P., Grinsted, J., Altenbuchner, J., & Schmitt, R. (1983). The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence analysis. *Nucleic Acids Res*, *11*(17), 6089-6105.
- Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*, *18*(5), 821-829.

- Zhang, Y., Pohlmann, E. L., & Roberts, G. P. (2005). GlnD is essential for NifA activation, NtrB/NtrC-regulated gene expression, and posttranslational regulation of nitrogenase activity in the photosynthetic, nitrogen-fixing bacterium *Rhodospirillum rubrum*. *J Bacteriol*, *187*(4), 1254-1265.
- Zhu, B., Zhou, S., Lou, M., Zhu, J., Li, B., Xie, G., Jin, G., & De Mot, R. (2011). Characterization and inference of gene gain/loss along burkholderia evolutionary history. *Evol Bioinform Online*, *7*, 191-200.
- Zhulin, I. B., Taylor, B. L., & Dixon, R. (1997). PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox. *Trends Biochem Sci*, *22*(9), 331-333.